CS 194-198 (Network Thy)

Joon Kim

OH Th 1:15 - 2
cory 271 M

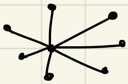
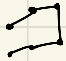
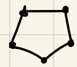
Networks

ex) Internet, social networks, gene regulatory networks, etc.

Review of Basic Graph Theory:

Graph $G=(V,E)$. Neighborhood $N_G(v) := \{u \mid uv \in E\}$.

Adjacency Matrix $A_{uv} = \mathbb{1}\{uv \in E\}$. Adjacency List also used.

ex) Star S_n.  Line L_n.  Circle C_n. 

Complete Graph K_n. $\binom{n}{2}$ edges.

Subgraph $G' = (V', E') \subseteq G$ where $V' \subseteq V$, $E' \subseteq E$, $\forall uv \in E'$, $u, v \in V'$.

Spanning Subgraph G' where $V' = V$.

Induced Subgraph $G[V']$ where $E' := \{uv \mid u, v \in V'\}$ (all possible edges)

↳ also equivalent to test $uv \in E' \Leftrightarrow uv \in E$ $\forall u, v \in V'$.

Walk $w := (v_1, v_2, \dots, v_k)$ s.t. $v_i v_{i+1} \in E$ $\forall i \in [k-1]$. $|w| = (k-1)$.

Path is a walk s.t. $v_i \neq v_j$ $\forall i \neq j$.

Distance $d_G(x,y) :=$ length of shortest path from x to y. ∞ if no such path.

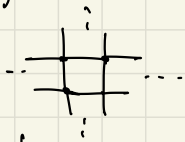
Diameter $D_G := \max_{x,y \in V} \{d_G(x,y)\}$. G is connected if D_G is bounded.

Connected Component $C \subseteq G$ where C is an induced subgraph that is connected; C is maximal s.t. any $C' \supseteq C$ with same properties is always C.

Tree T is a connected graph without cycles.

Forest F is a graph without cycles (or, a bunch of trees)

Lattice $(\mathbb{Z}^d, \mathbb{E}_d)$ looks like a jungle gym in d dimensions.



Torus of length L is a lattice that "wraps around" mod L.

bit distance

Hypercube has $V = \{0,1\}^n$ and $E = \{uv \mid d(u,v) = 1\}$.

Spanning Tree is a spanning subgraph which is a tree.

Thm) G is connected $\Leftrightarrow$ G has a spanning tree.

Proof: $\Leftarrow$) obvious, move on the spanning tree to visit any vertex.

$\Rightarrow$) remove edges in cycles until none exist. we get a spanning tree.

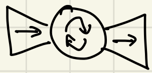
↳ a more elegant way is to do a breadth-first search (BFS).

Cayley's Formula: # of spanning trees on n vertices is $n^{(n-2)}$.

↳ Proof follows from a more general formula involving multinomials

Observed Properties of Graphs

- Short diameter of human connections, "6 degrees of seperation"
- Power-law degree distribution
- High clustering
- Connectivity

- Undirected: connected, or one giant component 
- Directed: one large WCC, giant SCCs, bow-tie structure

Probability Review

sample space   event space $\ni A \subseteq \Omega$, "questions"   probability measure: $\mathcal{F} \to [0,1]$

Probability Space: Tuple $(\Omega, \mathcal{F}, \Pr[\cdot])$ s.t.:

- $\Pr[\Omega] = 1$, $\Pr[\emptyset] = 0$, $\Pr[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} \Pr[A_i]$ if A_i are disjoint.
- $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$ (PIE)
- $\Pr[\bigcup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} \Pr[A_i]$ (Union Bound)
- $\Pr[A] + \Pr[A^c] = 1$ (Excluded Middle)

A happens <u>almost surely</u> if $\Pr[A] = 1$. ex) $\Omega = [0,1]$, $A = [0,\tfrac{1}{2}),(\tfrac{1}{2},1]$. (assume uniform)

Random Variables: Function $X : \Omega \to \mathbb{R}$ s.t. $\{\omega \in \Omega \mid X(\omega) \in I\} \in \mathcal{F}$.

↳ View as a realization of the underlying Ω

ex) $\Omega :=$ all infinite sequences $\omega \in \{0,1\}^{\infty}$. $X :=$ # of Heads until time t.

$Y :=$ first time we see a Head. ...

ex2) $\Omega :=$ all graphs on $[n]$. $X :=$ degree of vertex 1. $Y :=$ average degree.

$Z :=$ size of largest component. ...

Indicator RV: $\mathbb{1}_A\{\omega\} := \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{o.w.} \end{cases}$

Distribution on RV X: $P_X[A] := \Pr[\{\omega \mid X(\omega) \in A\}]$.

ex) $\Omega := [0,1]$, $\Pr$ is uniform. $X(\omega) := \mathbb{1}\{\omega \leq p\}$. $\rightarrow X \sim \mathrm{Ber}(p)$.

$Y(\omega) = 2\omega$. $\rightarrow Y \sim \mathrm{Uni}(0,2)$.

Discrete RV: $\exists$ countable $S := \{n_1, n_2, \dots\}$. $\Pr[X = n_i] = p_i$.

Probability Mass Function: $P_X(n_i) := p_i = \Pr[X = n_i]$.

Continuous RV: $\exists \underset{\text{PDF}}{\varphi_X} : \mathbb{R} \rightarrow \mathbb{R}_+$ s.t. $\Pr[X \in A] = \int_A \varphi_X(x)\,dx$.

Mixed RV: $\Pr[X \in A] = \sum_i P_X(n_i) + \int_A \varphi_X(x)\,dx$

Expectations: $E[X] := \int x\,\varphi_X(x)\,dx$, or for discrete RV, $\sum_i n_i P_X(n_i)$

Composition: $Y := h(X)$, then $E[Y] = \int \varphi_X(x)\,h(x)\,dx$ (LOTUS)

Linearity: $E[\sum_i a_i X_i] = \sum_i a_i E[X_i]$

Monotonicity: $X \leq Y \Rightarrow E[X] \leq E[Y]$

Variance: $\mathrm{Var}(X) := E[X^2] - E[X]^2 = E[(X - E[X])^2] = \sigma^2 \geq 0$.

PGF, MGF give higher order moments through derivatives.

ex) $X \sim \mathrm{Ber}(p)$. $E[X] = p$, $\mathrm{Var}(X) = p(1-p)$.

$X \sim \mathrm{Bi}(n,p)$. $E[X] = np$, $\Pr[X = k] = \binom{n}{k} p^k (1-p)^{(n-k)}$, $\mathrm{Var}(X) = np(1-p)$.

$X \sim \mathrm{Geo}(p)$. $\Pr[X = k] = (1-p)^{(k-1)} p$. $E[X] = \frac{1}{p}$.

$X \sim Po(c)$. $Pr[X=k] = \frac{c^k}{k!} e^{-c}$, $E[X] = c$.

$X \sim Uni(a,b)$. $E[X] = \frac{a+b}{2}$. $\varphi_X(x) = \frac{1}{a-b} \mathbb{1}_{x \in [a,b]}$

$X \sim Exp(\lambda)$. $\varphi_X(x) = \lambda e^{-\lambda x}$. $E[X] = \frac{1}{\lambda}$.

$X \sim N(\mu, \sigma)$. $\varphi_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$. $E[X] = \mu$, $Var(X) = \sigma^2$.

Joint Distributions: 2 RVs X, Y. Ω: $X(\omega)$, $Y(\omega)$. JD of X & Y is $D_{X,Y}(A) = Pr[(X(\omega), Y(\omega)) \in A]$. If X, Y are discrete, $X, Y \in \mathbb{N}_0$, then $Pr[X=k, Y=l] = P_{X,Y}(k,l)$. If X, Y are continuous, $X, Y \in \mathbb{R}$, then $Pr[(X,Y) \in A] = \int_A dx\,dy\, \rho_{X,Y}(x,y)$.

Marginals of $D_{X,Y}$: $D_X(A) = Pr[X(\omega) \in A] = Pr[X(\omega) \in A, Y(\omega) \in \mathbb{R}]$
$= D_{X,Y}(A \times \mathbb{R})$. Similarly, $D_Y(A) = D_{X,Y}(\mathbb{R} \times A)$.
$\rho_X(x) = \int dy\, \rho_{X,Y}(x,y)$, $P_X(k) = \sum_l P_{X,Y}(k,l)$.

Independence: Two events A, B are independent if $Pr[A \cap B] = Pr[A] \cdot Pr[B]$.
$A_1, \dots, A_n$ are pairwise indep. $\Leftrightarrow A_i \perp\!\!\!\perp A_j \ \forall i \neq j$.
Mutually indep. $\Leftrightarrow Pr[\bigcap_{i \in K} A_i] = \prod_{i=k} A_i \ \forall K \subseteq [n]$.
Independent RVs: $X \perp\!\!\!\perp Y \Leftrightarrow A = \{X \leq \lambda\}, B = \{Y \leq \eta\}, A \perp\!\!\!\perp B \ \forall \lambda, \eta$.
$\Leftrightarrow P_{X,Y}(k,l) = P_X(k) P_Y(l)$ (discrete), $\varphi_{X,Y}(x,y) = \varphi_X(x)\varphi_Y(y)$.

Conditionals: $\Pr[A|B] = \frac{\Pr[A\cap B]}{\Pr[B]}$.

If $X,Y \in \mathbb{N}_0$, $\Pr[X=k|Y=l] = \frac{\Pr[X=k, Y=l]}{\Pr[Y=l]}$.

If $X,Y \in \mathbb{R}$, $\varphi_X(x|Y=y) = \frac{\varphi_{X,Y}(x,y)}{\varphi_Y(y)}$.

Lemma: $X_1,\dots,X_n$ are mutually indep. $\Leftrightarrow \Pr[X_1=k_1,\dots,X_{n-1}=k_{n-1}|X_n=k_n] = \prod_{i=1}^{n-1}\Pr[X_i=k_i]$

Concentration Bounds: Given RV X, knowing that $E[X]=\mu$, what do I know about probabilities?

Markov (1st Moment Bound): $\Pr[X \geq a] \leq \frac{E[X]}{a}$ for $X \in [0,\infty)$.

$\hookrightarrow E[X] = \sum_k P_X(k)\cdot k \geq \sum_{k\geq a} P_X(k)\cdot k \geq \sum_{k \geq a} P_X(k)\cdot a = a \cdot \Pr[X\geq a]$.

Chebyshev (2nd Moment Bounds): $\Pr[|X-E[X]| \geq \lambda] \leq \frac{Var(X)}{\lambda^2}$.

Chernoff Bounds: $\Pr[\exp(|X-E[X]|) \geq e^\lambda] \leq E[e^X]e^{E[X]-\lambda}$.

Lemma: If $X_1,\dots,X_n$ are i.i.d $Bi(n,\frac{c}{n})$ or $Po(c)$, then

$$\Pr\left[\frac{1}{n}\sum_{i=1}^n X_i \leq \lambda\right] \leq \exp\left(-\frac{(c-\lambda)^2}{2c}n\right) \text{ if } \lambda < c,$$

$$\Pr\left[\frac{1}{n}\sum_{i=1}^n X_i \geq \lambda\right] \leq \exp\left(-\frac{(c-\lambda)^2}{2\lambda}n\right) \text{ if } \lambda > c.$$

Notions of Convergence: $X_n \to X$ in distribution if $\Pr[X_n \leq \lambda] \to \Pr[X\leq\lambda]$, or if dicrete, $\Pr[X_n=k] \to \Pr[X=k]$. $X_n \to X$ in probability if $\Pr[|X_n - X| \geq \varepsilon] \to 0 \ \forall \varepsilon > 0$.

i) CIP $\Rightarrow$ CID.

ii) If $X = c$, CIP $\Leftrightarrow$ CID.

iii) $X_n \to X$ in distr. $\Leftrightarrow G_{X_n}(\lambda) \to G_X(\lambda) \ \forall \lambda \in [0,1]$

$\Leftarrow M_{X_n}(t) \to M_X(t) \ \forall t \in (-\varepsilon, \varepsilon) \ \exists \varepsilon > 0$.

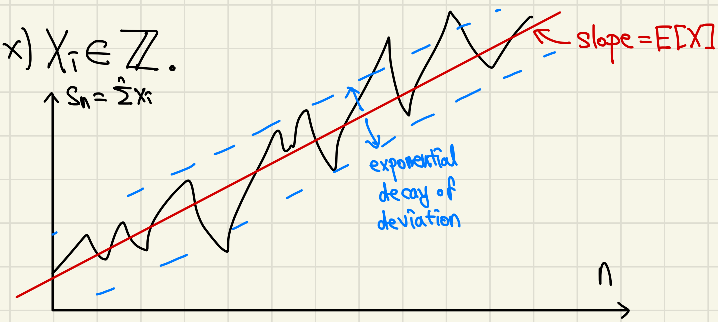
Weak Law of Large Numbers: If $X_1, \dots, X_n$ are i.i.d with $E[|X_i|] < \infty$,

$\Rightarrow \frac{1}{n} \sum X_i \to E[X_i]$ in probability.

Central Limit Theorem: If $X_1, \dots, X_n$ are indep. and $E[X_i^2] = \sigma^2 < \infty$,

$\Rightarrow \frac{1}{\sqrt{n}} \sum (X_i - E[X_i]) \to \mathcal{N}(E[X_i], \sigma^2)$.

ex) $X_i \in \mathbb{Z}$.

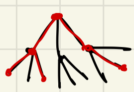


Law of Total Probability: If $B_1, \dots, B_n$ are disjoint, $\sum \Pr[B_i] = 1$.

Also, $\Pr[A] = \sum_i \Pr[A|B_i] \Pr[B_i]$. $\to E[A] = \sum_i E[A|B] \Pr[B_i]$.

Pf. $\sum_i \Pr[A|B_i]\Pr[B_i] = \sum_i \Pr[A \cap B_i] = \Pr[\bigcup (A \cap B_i)] = \Pr[A \cap \cancel{(\bigcup B_i)}^{\Omega}]$. //

Branching Processes

Consider a d-ary tree.  We have an infection at the root, which spreads w.p. p. Does the infection survive?

$R_0 = E[\#\text{ of infected children} \mid \text{parent infected}] = pd.$

General: $X \in \mathbb{N}_0$. D_x is an offspring distribution. $E[X] = c$. $\Pr[X=k] \overset{p_k.}{=}$

$\emptyset, v_1$ is the root. X_1 children. $X_1 \sim D_X$.

Inductively, if in generation $(t-1)$ we have $Z_{(t-1)}$ individuals,

then for $i = 1, \dots, Z_{(t-1)}$, each i has $X_{t,i} \sim D_x$ children (i.i.d).

The possibly infinite tree is denoted T_X.

Analysis: $\theta = \Pr[|T_X| = \infty]$? When does the tree explode?

$E[Z_t \mid Z_{t-1} = k] = E[\sum_i^k X_{t,i}] = c \cdot k.$

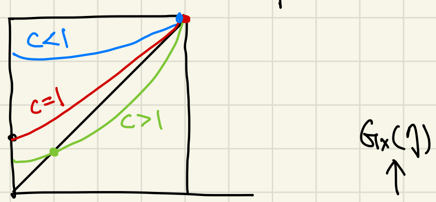
$\Rightarrow E[Z_t] = \sum_k \overbrace{E[Z_t \mid Z_{t-1} = k]}^{ck} \Pr[Z_{t-1} = k] = c \cdot E[Z_{t-1}].$

$\Rightarrow Z_0 = 1$, so $E[Z_t] = c^t$.

$$E[|T_X|] = \sum_i E[Z_i] = \begin{cases} \frac{1}{1-c} & \text{if } c<1 \text{ (subcrit.)} \\ \infty & \text{if } c=1 \text{ (crit.)} \\ \infty & \text{if } c>1 \text{ (supercrit.)} \end{cases}. \quad \underline{c<1 \Rightarrow \theta = 0.}$$

Thm) Let G_X be the PGF of X. Then $\eta := 1-\theta$ (extinction prob.)

is the <u>smallest</u> solution of $\eta = G_X(\eta)$.



$G_X(\eta) = \sum_k p_k \eta^k$. $G'_X(1) = E[X] = c$.

Obs: $\eta = \Pr[|T_X| < \infty] = \sum_k \Pr[X_1 = k] \cdot \Pr[|T_X| < \infty \mid X_1 = k] = \sum_k p_k \cdot \eta^k.$

(not proven "smallest" yet!)

Corollary: If $p_0 = 0 \Rightarrow |T_X| = \infty$ w.p. 1.

If $p_0 > 0$, $\Pr[|T_x| = \infty] = \theta > 0 \iff c > 1$.

Random Walk Representation: $|T_x|$ = size of BFS of T_x.

$Q_t :=$ Queue at step t. $D_t :=$ set of discovered v. $Y_t := |Q_t|$.

↳ Step 0: Root $= v_1$, $Q_0 = \{v_1\} = D_0$. $Y_0 = 1$.

↳ Step 1: Dequeue v_1, add children of v_1 to Q_0. # of children $X_1 \sim D_x$.

$Y_1 = Y_0 - 1 + X_1$. $Q_1 = \{v_2, \ldots, v_{1+X_1}\}$.

↳ … $Q_{t-1} = \{v_t, \ldots\}$ if $Q_{t-1} \neq \emptyset$, $Y_{t-1} > 0$. Dequeue v_t. Add $\overset{\sim D_x}{X_t}$ children.

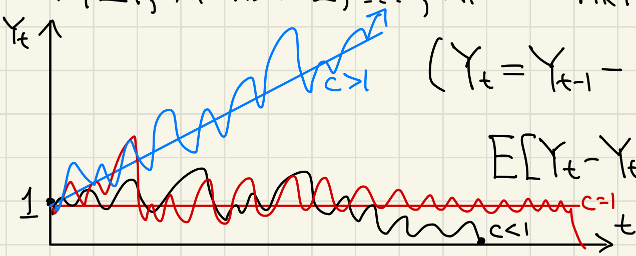
$Y_t = Y_{t-1} - 1 + X_t = 1 + X_1 + \cdots + X_t - t$.

If $Y_t > 0 \ \forall t$, $|T_x| = \infty$. O/w, let $k := \min_t \{t \mid Y_t = 0\}$.

Then, $Q_{k-1} = \{v_k\}$. $V(T) = \{v_1, \ldots, v_k\}$. $\Rightarrow |T| = k$.

The random size of T_x is $k \iff k = \min_t \{t \mid Y_t = 0\}$ by recurrunce of Y_k

$\iff X_1 \geq 1, X_1 + X_2 \geq 2, \ldots, X_1 + \cdots + X_{k-1} \geq k-1, X_1 + \cdots + X_k = k-1$.



$(Y_t = Y_{t-1} - 1 + X_t)$

$E[Y_t - Y_{t-1}] = E[X_t] - 1 = c - 1$.

→ If $c < 1$, we have a down drift. If $c = 1$, we hover around $Y_t = 1$, but as long as X is not constant, it eventually hits 0.

$\Pr[X=1] = 1$. (* in expectation it's infinite, but fluctuation kills it in finite steps)

Claim: $|T_x| = k \Leftrightarrow X_1 \ge 1, \dots, X_1 + \cdots + X_{k-1} \ge k-1, X_1 + \cdots + X_k = k-1$.

$\hookrightarrow c = E[X] < 1 \Rightarrow \Pr[|T_x| < 1] = 1?$

(a) If $|T_x| \ge k \Rightarrow X_1 + \cdots + X_k \ge k$. (b) If $|T_x| = k \Rightarrow X_1 + \cdots + X_k < k$.

Lemma: If $X \sim Poi(c)$ or $Bi(n, c/n)$, then

(a) for $c < 1$, $\Pr[|T_x| > k] \le e^{-\frac{(1-c)^2}{2}k}$,

(b) for $c > 1$, $\Pr[|T_x| = k] \le e^{-\frac{(c-1)^2}{2c}k}$. [Proofs in HW, use Chernoff]

Lemma 2: For general X, $E[X] = c > 1$, $\Pr[|T_x| = k] \le e^{-Dk}$ for some $D > 0$.

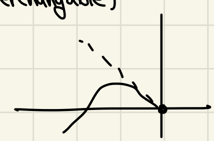
Proof: LHS $\le \Pr[X_1 + \cdots + X_k \le k] \le e^{-k I_x(1)}$ if $c > 1$.

$$I_x(1) := \sup_{t \le 0} \{\overbrace{t \cdot 1 - \log E[e^{tX}]}^{\phi(t)}\}.$$

$\phi(0) = 0 - \log 1 = 0$. $\left(\lim_{\varepsilon \to 0} \frac{E[e^{(1+\varepsilon)X} - e^{\varepsilon X}]}{\varepsilon}\right)$

$\phi'(t) = 1 - \frac{1}{E[e^{tX}]} \cdot E[Xe^{tX}] \leftarrow$ (by lin. of exp. derivative is interchangable)

$\lim_{t \to 0} \phi'(t) = 1 - E[X] < 0. \Rightarrow D = I_x(1) > 0.$ //



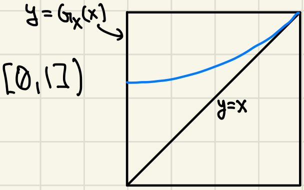
Corollary: If $c > 1$, then $\Pr[|T_x| = \infty] > 0$.

Proof: $\infty = E[|T_x|] = \infty \cdot \Pr[|T_x| = \infty] + \sum_k^{<\infty} k \cdot \overbrace{\Pr[|T_x| = k]}^{\le e^{-Dk}}$.

Assume for contradiction that $\Pr[|T_x| = \infty] = 0$.

$\rightarrow E[|T_x|] \le \sum_k k e^{-kD} < \infty$. Contradiction. $\Rightarrow \Pr[|T_x| = \infty] > 0$. //

What if $c=1$, $Pr[X=1]<1$? $(\eta := Pr[|T_x|<\infty],\ \lambda\in[0,1])$



We want η s.t. $G_X(\eta)=\eta$ where $G_X(\lambda)=\sum_k \lambda^k P_k$.

$G'_X(\lambda)=\sum_k \lambda^{(k-1)} k\cdot P_k$. $G'_X(1)=E[X]=c=1$.

Case 1: $G_X(x)$ is linear. $\to$ cannot happen because derivative <1.

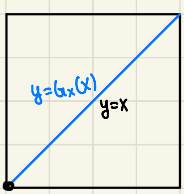
$\Rightarrow$ Must be Case 2: $G_X(\lambda)$ is strictly convex.

$\Rightarrow$ $\lambda=G_X(\lambda)$ has only one solution, $\lambda=1$. $\Rightarrow \eta=1 \Rightarrow Pr[|T_x|=\infty]=0$. //

- If $c<1$, $Pr[|T_x|=\infty]=0$.
- If $c>1$, $Pr[|T_x|=\infty]>0$.
- If $c=1$ and $P_1=Pr[X=1]<1$, $Pr[|T_x|=\infty]=0$.
- If $c=1$ and $P_1=Pr[X=1]=1$, $Pr[|T_x|=\infty]=1$.

Thm) Whatever $E[X]=c$ is, η is the first intersection of $x\mapsto x$ & $G_X(x)$.

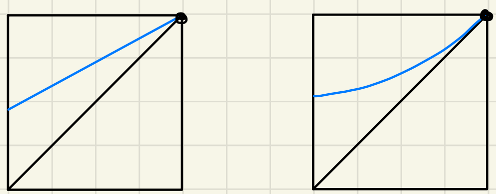
Case 1: $c=1$, $P_1=1$.



$G_X(x)=\sum_k P_k x^k = x$

$\Rightarrow \eta=0$. //

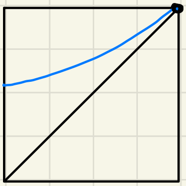
Case 2: $c<1$, $G'_X(1)=c<1$.



Only one intersection of 1,

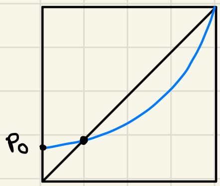
and $\eta=1$. //

Case 3: $c=1$, $p_1<1$. Case 4: $c>1 \Rightarrow G_X'(1)>1$.

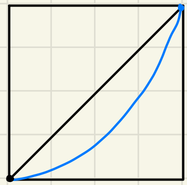


One intersection,

$\eta = 1$. //



$p_0>0$.



$p_0=0$.

if $c>1$, $\theta = 1-\eta > 0$. $\Rightarrow \eta<1$. Must be first. //

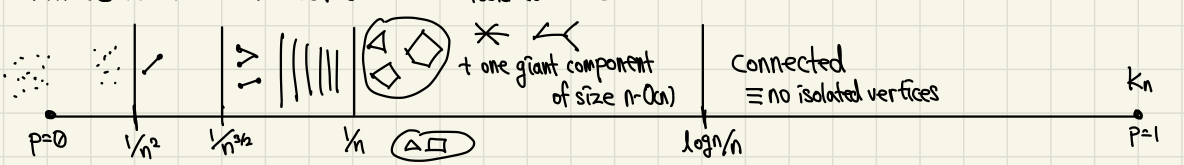
$\Rightarrow$ η is the earliest intersection of $x \mapsto x$ & $y=G_X(x)$. //

$\Rightarrow$ If $p_1 = \Pr[X=1]<1$, $\eta<1 \Leftrightarrow c>1$. *

Erdos-Renyi Random Graph [59', 60']

Def) $G(n,p)$: start with K_n and keep each edge i.i.d. w.p. p.

Threshold Behavior:



Def) Graph Property: a property P invariant of labelling, $P \subseteq \Omega$ (all $[n]$ graphs)

↳ If P is a graph property, $f(n)$ is the threshold for P if, as $n\to\infty$,

$$\Pr[P]\to 1 \text{ for } p \gg f(n),\quad \Pr[P]\to 0 \text{ for } p \ll f(n).$$

from $G(n,p)$ (or for P^c)

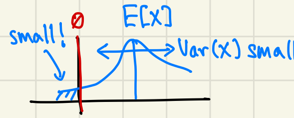
Thm) $\frac{1}{n^2}, \frac{1}{n^{k/k-1}}, \frac{1}{n}$ are thresholds for edges, k-trees, and k-cycles appearing in $G(n,p)$.

Proof Idea: First & Second Moment Method.

First Moment Method: If $X \in \mathbb{N}_0$, $\Pr[X>0] \leq E[X]$. Show $E[X] \to 0$.

Second Moment Method: Show $E[X^2] \leq E[X]^2 \cdot (1+\varepsilon_n)$ where $\varepsilon_n \to 0$.

Lemma: If $X \geq 0$, $\Pr[X>0] \geq \frac{E[X]^2}{E[X^2]}$.



Proof: $E[X] = E[X \cdot \mathbb{1}\{X>0\}] \leq E[X^2]^{1/2} \cdot E[\mathbb{1}\{X>0\}]^{1/2}$

$\to E[X]^2 \leq E[X^2] \cdot \Pr[X>0]$. //

Edges: $X :=$ # of edges in $G(n,p)$. Formally, $E :=$ edge set of $G(n,p)$,

$M := \binom{n}{2}$, # of possible edges. $X = \sum_e \mathbb{1}\{e \in E\}$.

$E[X] = \sum_e \Pr[e \in E] = M \cdot p \to 0$ when $p \ll 1/n^2$. //

For $p \gg 1/n^2$, $E[X^2] = \sum_{e,e'} E[\mathbb{1}\{e \in E\} \cdot \mathbb{1}\{e' \in E\}]$

$= \sum_{e \neq e'} E[\sim] + \sum_{e=e'} E[\sim] = M(M-1)p^2 + Mp$

$\leq M^2p^2(1+\frac{1}{Mp}) = E[X]^2(1+\varepsilon_n)$. //

$\Rightarrow \Pr[G \text{ has an edge}] = \Pr[X>0] \geq \frac{E[X]^2}{E[X^2]} \geq \frac{1}{1+\varepsilon_n} \to 1$. //

Trees: $X :=$ # of trees T of size k in $G(n,p) = \sum_{\substack{T \subseteq [n] \\ |V(T)|=k}} \mathbb{1}\{T \subseteq G(n,p)\} = \sum_T \prod_{e \in T} \mathbb{1}\{e \in E\}$.

$$E[X] = \sum_{T} E\left[\prod_{e\in T} \mathbb{1}\{e\in E\}\right] = \sum_{T} p^{(k-1)} = \sum_{S\subseteq[n],|S|=k} p^{(k-1)} \cdot N_k \text{ where } N_k := \#\text{ trees on } k \text{ vertices.}$$

$$= \binom{n}{k} N_k\, p^{(k-1)}.\quad \binom{n}{k} = \frac{n\cdots(n-k+1)}{k!} = \frac{n^k}{k!}\prod_{i=0}^{k-1}\left(1-\frac{i}{k}\right) \le \frac{n^k}{k!}.$$

$$\to E[X] \le \frac{N_k}{k!} n^k p^{k-1} \to 0 \text{ if } p \ll n^{-k/k-1} \text{ since } p n^{k/k-1} \to 0 \Rightarrow p^{k-1} n^k \to 0.$$

$$E[X^2] = \sum_{T}\sum_{T'} E[\mathbb{1}\{T\subseteq G(n,p)\}\cdot \mathbb{1}\{T'\subseteq G(n,p)\}] = \sum_{T,T'} E\left[\prod_{e\in E(T)\cup E(T')} \mathbb{1}\{e\in E\}\right]$$

$$= \sum_{T,T'} p^{|E(T)\cup E(T')|} = \sum_{T,T'} p^{|E(T)|+|E(T')|-|E(T)\cap E(T')|} = p^{2(k-1)} \sum_{T,T'} p^{-|E(T)\cap E(T')|}.$$

Assume $|E(T)\cap E(T')| = l \ge 1 \Rightarrow |V(T)\cap V(T')| \ge l+1 \ge 2$.

$$\to \le \sum_{\substack{T,T' \\ E(T)\cap E(T')=\emptyset}} p^{2(k-1)} + \sum_{s=2}^{k}\sum_{\substack{T,T' \\ |V(T)\cap V(T')|=s}} p^{2(k-1)-(s-1)} \quad [|E(T)\cap E(T')| \le |V(T)\cap V(T')|-1]$$

$$\le p^{2(k-1)}\binom{n}{k}^2\cdot N_k^2 + \underbrace{p^{2(k-1)}\sum_{s=2}^{k}\sum_{S,S'\subseteq[n],|S\cap S'|=s} p^{-(s-1)}\cdot N_k^2}$$

$\binom{n}{k}$ choices for S. $\binom{k}{s}$ choices for $S\cap S'$. $\binom{n-k}{k-s}$ choices for $S\setminus S'$.

$\binom{n-k}{k-s} \le (n-k)^{(k-s)} \le n^{(k-s)}$.

$$\Rightarrow E[X^2] \le \left[N_k\binom{n}{k}p^{k-1}\right]^2 + p^{2(k-1)}\binom{n}{k} n^k \sum_{s=2}^{k}\binom{k}{s} p^{-(s-1)} n^{s}\cdot N_k^2.$$

$$\le E[X]^2 + C(k)\, E[X]^2 \sum_{s=2}^{k}(\sim). \text{ Using } p \gg n^{k/k-1},\ p^{-(n-1)} n^{s} \ll 1,$$

the sum $\sum_{s=2}^{k}(\sim) =: \varepsilon_n \to 0$.

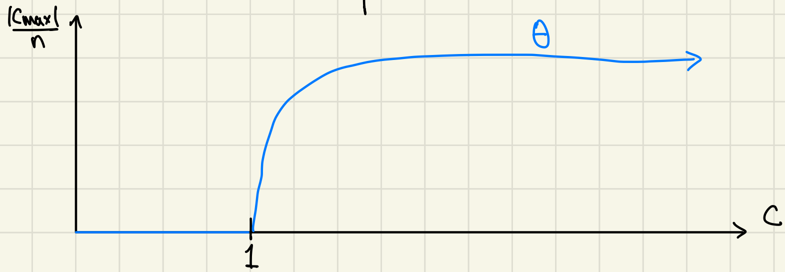
Phase Transition at $p=1/n$ (The Giant)

Let $X\in[n]$. $C(X)$ is the connected component of X.

C_{max} is the largest connected component, $\max_{X}\{C(X)\}$. $\theta = \Pr[|T_{Poi(c)}| = \infty]$.

Let $p=\frac{c}{n}$, c is constant. $\rightarrow E[\deg(x)]=(n-1)\cdot p=(n-1)\frac{c}{n}\simeq c$.

Thm) (a) If $c<1$, $|C_{max}|\in O(\log n)$ w.h.p. (b) If $c>1$, $\frac{|C_{max}|}{n}\rightarrow\theta$ in prob. as $n\rightarrow\infty$, and all other components are of size $O(\log n)$.



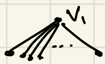
C_{max} is also called the "giant". $\rightarrow$ "The giant is unique".

Why $G(n,p)\leftrightarrow T_{Poi(c)}$?

Hint: $\deg(v)\sim Bi(n-1,\frac{c}{n})\longrightarrow Poi(c)$ in distr.

Intuition: viewing from any vertex x, $G(n,p)$ (roughly) looks like $T_{Poi(c)}$.

Explore $C(X)$ using BFS.



$x=v_1$. $Q_0=\{x\}$, $y_0=1$. $(n-1)$ possible neighbors. $\rightarrow X_1\sim Bi(n-1,p)$

$Q_1=\{v_2,\dots,v_{1+X_1}\}$. $y_1=X_1$. v_2 has $(n-1-X_1)$ possible children.

In reality, it has $X_2\sim Bi(n-1-X_1,p)$, $y_2=y_1-1+X_2$ if $y_1>0$.

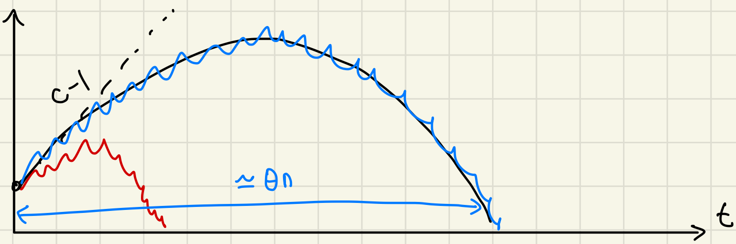
As for any branching process, we have $y_t=\begin{cases} y_{t-1}-1+X_t & \text{if } y_{t-1}>0\\ 0 & \text{if } y_{t-1}=0.\end{cases}$

$|C(X)|=k$ where $k=\min_t\{t\mid y_t=0\}\Leftrightarrow X_1+\dots+X_{k-1}\geq k-1,\ X_1+\dots+X_k=k$.

But $X_i | X_1,\dots,X_{i-1} \sim Bi(n-1-X_1-\cdots-X_{i-1}, p)$, which is not independent.

Drift of y_t? $E[y_t | \overset{(X_1,\dots,X_{t-1})}{y_{t-1}}] - y_{t-1} = E[X_t] - 1 = \frac{c}{n}(n-1-X_1-\cdots-X_{t-1}) - 1 \overset{\text{if } c<1}{<} 0$

OTOH, if $c>1$, the drift starts out $\simeq c-1$, then dies out before $t=n$.



Lemma 1 (Comparison): $\Pr[|T_{Bi(n-k,p)}| > k] \le \Pr[|C(X)| > k] \le \Pr[|T_{Bi(n,p)}| > k]$.

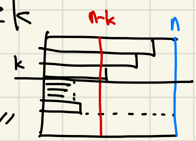
Proof: Upper bound) $x = v_1$. $X_1 = \sum_{i\ne v_1} A_{v_1,i}$ where $A_{v_1,i} \sim Ber(p)$.

Predraw n^2 RVs $B_{ij}\sim Ber(p)$. Set $X_1=\sum_{j=1}^{n-1} B_{1j}$. $X_2 = \sum_{j=1}^{n-1-X_1} B_{2j} \sim Bi(n-1-X_1,p)$.

$\Pr[|C(x)|>k] = \Pr[\overbrace{X_1\ge 1,\dots,X_1+\cdots+X_k\ge k}^{\circledast}]$. $X_i=\sum_{j=1}^{n-1-(X_1+\cdots+X_{i-1})} B_{ij} \le \sum_{j=1}^{n} B_{ij} =: X_i^+$.

$X_i^+\sim Bi(n,p)$. Since $X_i\le X_i^+$, $\circledast \Rightarrow X_1^+\ge1,\dots,X_1^++\cdots+X_k^+\ge k$

$\Leftrightarrow |T_{Bi(n,p)}|>k$. $\Rightarrow \Pr[|C(x)|>k]\le\Pr[|T_{Bi(n,p)}|>k]$. //



overcounting)

(we are sampling all possible vertices but only including unvisited ones in the Q.)

Lower bound) $\Pr[|C(x)|\le k] \le \Pr[|T_{Bi(n-k,p)}|\le k]$ (negated both sides)

$|C(x)|<k \Leftrightarrow \exists l\le k$ s.t. $\overbrace{X_1+\cdots+X_l \le l-1}^{\circledast}$. $\circledast \Rightarrow n-1-X_1-\cdots-X_l \ge n-l \ge n-k$.

$X_i = \sum_{j=1}^{n-1-(X_1+\cdots+X_{i-1})} B_{ij} \ge \sum_{j=1}^{n-k} B_{ij} =: X_i^-$. $\Rightarrow \exists l\le k$ s.t. $X_1^-+\cdots+X_l^- \le l-1 \Leftrightarrow |T_{Bi(n-k,p)}|\le k$.

$\Rightarrow \Pr[|C(x)|\le k]\le\Pr[|T_{Bi(n-k,p)}|\le k]$. // (can't use up too much coins)

Corollary 1: If $c<1$, $|C_{max}| \le \frac{2}{(1-c)^2}\log n$.

$X_i^+ \sim Bi(n,p)$

Proof: $Pr[|C(x)|>k] \le Pr[|T_{Bi(n,p)}|>k] \le Pr[X_1^+ + \cdots + X_k^+ \ge k]$

$\le \exp(-\frac{(1-c)^2}{2}k)$. If $k_0 := \lceil \frac{2}{(1-c)^2}\log n \rceil$, $Pr[|C(x)|>k_0] \le \frac{1}{n}$. Jump from 0 to $>k_0$!

$\rightarrow Pr[\exists x \text{ s.t. } |C(x)|>k_0] = Pr[N>0]$ where $N = \sum_x 1\{|C(x)|>k_0\}$.

$= Pr[N>k_0] \le \frac{E[N]}{k_0} = \frac{n \cdot Pr[|C(x)|>k_0]}{k_0} \le \frac{n \cdot 1/n}{k_0} = k_0 \rightarrow 0$ as $n \rightarrow 0$. //

Corollary 2: If $c>1$ and $k=k_n \rightarrow \infty$ where $k \in o(n)$, then

$Pr[|C(x)|>k] \rightarrow Pr[|T_{Po(c)}| = \infty] = \theta(c)$.

Proof: $Pr[|C(x)|>k] \le Pr[|T_{Bi(n,p)}|>k] = Pr[|T_{np}| = \infty] + Pr[k < |T_{np}| < \infty]$

$\rightarrow Pr[|T_{Po(c)}| = \infty] + 0 = \theta(c)$ as $np \rightarrow c$. //

$Pr[|C(x)|>k] \ge Pr[|T_{Bi(n-k,p)}|>k] = Pr[|T_{n-k,p}| = \infty] + Pr[k < |T_{n-k,p}| < \infty]$

$\ge Pr[|T_{n-k,p}| = \infty]$. $(n-k)p = \frac{n-k}{n}c \rightarrow c$ since $k \in o(n)$.

$\rightarrow Pr[|T_{Po(c)}| = \infty] = \theta(c)$. // By Squeezing, $Pr[|C(x)|>k] \rightarrow \theta(c)$. //

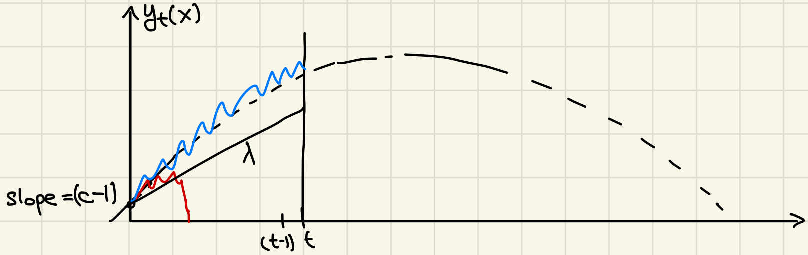
Choose some $K>0$. $V_+ := \{x \in [n] \mid |C(x)| > K\log n\}$.

$E[|V_+|] = n \cdot Pr[|C(x)| > K\log n] = n \cdot \theta(c) \rightarrow \frac{1}{n}E[|V_+|] = \theta(c)$. ...?

But we want to prove that the giant is <u>unique</u>!

Proof Strategy: All $x \in V_+$ (should) lie in one component. $Pr[|\frac{|V_+|}{n} - \frac{E[|V_+|]}{n}| \le \varepsilon] \rightarrow 1$?

Main Proof: 1) Prove that if $x \in V_+$, BFS queue stays "large" up to time $n^{2/3}$.



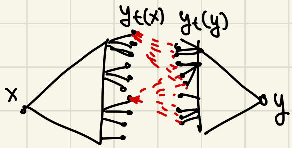
Lemma 2: If $c>1$, $1 \leq \lambda+1 < C$, and k/n is small enough, then

$$\Pr[y_{k-1} \geq 0,\ y_k \leq \lambda k] \leq \exp\left(-k\frac{(c-1-\lambda)^2}{2c} + o\left(\frac{k^2}{n}\right)\right). \quad \left[\exp\left(-k\frac{(c-(\lambda+1))^2}{4c}\right)\right]$$

Lemma 3: Let $t=n^{2/3}$. $\Pr[y_t(x) \geq \frac{c-1}{2}t,\ y_t(y) \geq \frac{c-1}{2}t,\ y \notin C(x)]$

$$\leq (1-p)^{\left(\frac{c-1}{2}n^{2/3}\right)^2}.$$

Proof of L3: BFS for x,y.



Case 1: Before reaching t, I discover a vertex already explored. $\Rightarrow y \in C(x)$.

Case 2: By time t, both BFS trees don't intersect. But we have at least $\left(\frac{c-1}{2}n^{2/3}\right)^2$ edges between $y_t(x)$ & $y_t(y)$. $\rightarrow y \notin C(x)$ is extremely unlikely, $(1-p)^{(\sim)^2}$. //

Proof of L2: $\Pr[y_{k-1} > 0,\ y_k \leq \lambda k]$. If $y_{k-1} > 0$, $y_k = 1 + X_1 + \cdots + X_k - k$.

$\Rightarrow 1 + X_1 + \cdots + X_k \leq (\lambda+1)k \Rightarrow \forall i \in [k-1],\ n-1-X_1 \cdots - X_i \geq \overbrace{n-(\lambda+1)k}^{n'}$,

$X_i^- \sim Bi(n', p)$, so $\forall i \in [k]\ X_i \geq X_i^-$.

LHS $\leq \Pr[X_1^- + \cdots + X_k^- \leq k(\lambda+1)]$. If $c_- = E[X_i^-] = cn' = c - \frac{(\lambda+1)k}{n}$,

LHS $\leq \exp\left(-k\frac{(\lambda-c_-)^2}{2c_-}\right)$ (when k/n is small enough.)

$$= \exp\left(-k\frac{(\lambda-c)^2}{2c} + o\left(\frac{k^2}{n}\right)\right). //$$

Lemma 2, in exact form is:

If $0 \le \lambda < c-1$ and $t \le \frac{c-1-\lambda}{4c^2}$, $\Pr[y_{t-1} > 0, y_t \le \lambda t] \le \exp\left(-t\frac{(c-1-\lambda)^2}{4c}\right)$.

Corollary 4: Let $k_- := \frac{4c}{(c-1)^2}\log n$, $k_+ := \frac{c-1}{4c^2}n$. Then,

(a) $\Pr[k_- \le |C(x)| \le k_+] = O(1/n)$,

(b) If $4k_- \le t \le k_+$, $\Pr[|C(x)| > k_-, y_t \le \frac{c-1}{2}t] = O(1/n)$.

Proof: (a) $\text{LHS} \le \sum_{t=k_-}^{k_+} \Pr[|C(x)| = t] = \sum_{t=k_-}^{k_+} \Pr[y_t = 0, y_{t-1} > 0]$

$\le \sum_{t=k_-}^{k_+} \exp\left(-\frac{(c-1)^2}{4c}t\right) = O\left(\exp\left(-\frac{(c-1)^2}{4c}k_-\right)\right) = O(1/n)$. //

(b) $\text{LHS} \le \Pr[k_- \le |C(x)| \le t] + \Pr[|C(x)| \ge t, y_t \le \frac{c-1}{2}t]$

$\le O(1/n) + \Pr[y_{t-1} > 0, y_t \le \frac{c-1}{2}t]$. $[\lambda = \frac{c-1}{2} \Rightarrow c-(\lambda+1) = \frac{c-1}{2}]$

$\le O(1/n) + \exp\left(-t\frac{(c-1)^2}{16}\right) = O(1/n)$. //

Corollary 5: Let $k_- := \frac{4c}{(c-1)^2}\log n$, $V_+ := \{x \mid |C(x)| \ge k_-\}$.

If $4k_- \le t = o(n)$, then w.h.p., $y_t(x) \ge \frac{c-1}{2}t \ \forall x \in V_+$.

Proof: Let $\text{Bad} := \{x \in V_+ \mid y_t(x) \le \frac{c-1}{2}t\}$.

$\Pr[\text{Bad} \ne \emptyset] = \Pr[|\text{Bad}| \ge k_-] \le \frac{1}{k_-}E[|\text{Bad}|] = \frac{1}{k_-}n \cdot o(1/n) \to 0$. //

6 Degree of Seperation in $G(n,p)$

$G(n,p) \to C_{max} \longrightarrow x,y \overset{\text{u.a.r.}}{\in}_R$. $D := \text{dist}(x,y)$.

Claim: $D \simeq \log n$.

Thm) $\frac{D}{\log_c n} \longrightarrow 1$ as $n \to \infty$ in probability.

Proof Strategy: Need to show that

(a) $\Pr[D \leq (1-\varepsilon)\log_c n] \to 0 \quad \forall \varepsilon > 0$,

(b) $\Pr[D \geq (1+\varepsilon)\log_c n] \to 0 \quad \forall \varepsilon > 0$.

Proof of (a): Let $l_n := (1-\varepsilon)\log_c n$.

Claim: for $x \neq y$ fixed, $\Pr_{G(n,p)}[d(x,y) \leq (1-\varepsilon)\log_c n] \leq \frac{1}{n}\frac{c^{l_n}}{1-1/c}$.

Define $X_l :=$ # of paths $\omega = (x, x_1, \dots, x_{l-1}, y)$ of length l from x to y.

$\Pr[\omega \text{ is a path in } G(n,p)] = p^l. \to E[X_l] = \sum_{\omega: x \to y} p^l \leq p^l \cdot n^{l-1} = \frac{1}{n} \cdot c^l$.

$\to$ LHS $= \Pr[\exists \omega \mid x \to y, |\omega| \leq l_n] \leq E[\# \text{of } \omega \mid x \to y, |\omega| \leq l_n]$

$\leq \sum_{l=0}^{l_n} E[X_n] \leq \frac{1}{n}\sum_{l=0}^{l_n} c^l \leq \frac{1}{n}\frac{c^{l_n}}{1-1/c}$. //

$\Pr[D \leq l_n] = E[\mathbb{1}\{D \leq l_n\}] = E_{G(n,p)}\left[\frac{1}{|C_{max}|^2}\sum_{x,y \in C_{max}} \mathbb{1}\{d(x,y) \leq l_n\}\right]$

$= E_{G(n,p)}\left[\left(\frac{1}{|C_{max}|^2}\sum \sim\right) \cdot \mathbb{1}\{|C_{max}| \geq \frac{\theta(c)}{2} n\}\right] + E_{G(n,p)}\left[\left(\frac{1}{|C_{max}|^2}\sum \sim\right) \cdot \mathbb{1}\{|C_{max}| \leq \frac{\theta(c)}{2}\}\right]$.

Recall that $\frac{|C_{max}|}{n} \to \theta(c)$ in prob. $\Rightarrow \Pr\left[\frac{|C_{max}|}{n} \leq \frac{\theta(c)}{2}\right] = \varepsilon_n \to 0$.

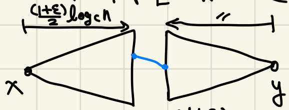
$\to \Pr[D \le l_n] \le E_{G(n,p)}\left[\left(\frac{1}{|C_{max}|^2}\sum_{x,y} 1\{d(x,y)\le l_n\}\right)\cdot 1\{|C_{max}| \ge \frac{\theta(c)}{2}n\}\right] + \varepsilon_n$

$\le \left(\frac{2}{\theta(c)n}\right)^2 \cdot \left[\cancel{1\{|C_{max}|\ge \frac{\theta(c)}{2}n\}}\cdot \sum_{x,y\in[n]} 1\{d(x,y)\le l_n\}\right] + \varepsilon_n$

$= \left(\frac{4}{n^2\theta^2}\right)\sum_{x,y\in[n]} \Pr[d(x,y)\le l_n] + \varepsilon_n.$

Using the above claim, LHS $\le \frac{4}{\theta^2}\cdot\frac{1}{n}\cdot\frac{\overset{n^{(1-\varepsilon)}}{c^{l_n}}}{1-1/c} + \frac{4}{n^2\theta^2}n + \varepsilon_n \to 0$ as $n\to\infty$.

Heuristic of (b): Recall that for BFS, $y_t :=$ length of Q at time t, and $Z_k :=$ # of individuals in generation k, then for $T_{Bi(n,p)}$, $E[Z_k] = c^k$, and that $\Pr[Z_k \le (c+\delta)^k] \to 0$. Now construct a double-sided BFS.



$y_{t_k} \simeq c^{\frac{(1+\varepsilon)}{2}\log_c n} = n^{\frac{(1+\varepsilon)}{2}}$. $\to \Pr[\text{no blue edge exists}]$

$\le (1-p)^{n^{(1+\varepsilon)}} \le e^{-pn^{(1+\varepsilon)}} = e^{-cn^{\varepsilon}} \to 0.$

Proof of (b): 1) Recall that if $t \le \frac{c-1-\lambda}{c^2}n$, $\Pr[y_{t-1}>0, y_t \le \lambda t] \le e^{-\frac{(c-1-\lambda)^2}{4c}t}$.

Setting $\lambda := c-1-\delta$ gives $\Pr[y_{t-1}>0, y_t \le (c-1-\delta)t] \le e^{-\frac{\delta^2}{4c}t}$.

2) How is y_t and Z_k related? When I empty Q of generation $k-1$, $y_{t(k)} = Z_k$.

$t(1) = 1$. $t(k+1) = t(k) + Z_k = t(k) + y_{t(k)}$.

Using the lemma, $t(k+1) \ge t(k) + (c-1-\delta)t(k)$ w.p. $1 - e^{-\frac{\delta^2}{4c}t}$.

$\to t(k+1) \ge t(k)(c-\delta)$. [$t$ grows "exponentially" w.r.t. k]

Lemma: w.h.p., $t(k) \le t_+ = n^{(1-\delta)}$ for $k \le (1-2\delta)\log_c n$.

Proof: $t(k) = t(k-1) + Z_{k-1} = 1 + Z_1 + \cdots + Z_{k-1}$. $E[t(k)] = \sum_{l=0}^{k-1} c^l \in O(c^k)$.

$\Pr[t(k) \geq n^{(1-\delta)}] \leq \frac{1}{n^{(1-\delta)}} E[t(k)] = \frac{1}{n^{(1-\delta)}} O(c^{(1-2\delta)\log_c n}) = O(n^{-\delta}) \to 0.$ //

Observe that $t(k) \geq k$, and that if $x \in C_{max}$, whp $|C(x)| \geq n^{(1-\delta)}$.

$\to$ If $x \in C_{max}$, whp, $\forall k \in [k_-, k_+]$, $k_- = t_- := \delta \log_c n$, $k_+ := (1-2\delta)\log_c n$,

$t(k) = t(k-1) + y_{t(k-1)} \geq t(k) + \ \_ \ \_ \ \_$ [finish later]

Power Laws

$Poi(c)$ achieves exponential decay in degrees. We want a power law.

Model 1 [Chung, Lu '02]: Fix expected degree sequence $\tilde{\omega} = (\omega_1, \dots, \omega_n)$.

Connect (i,j) w.p. $P_{ij} = \frac{\omega_i \omega_j}{l_n}$ where $l_n := \sum_{k=1}^{n} \omega_k$ if $i \neq j$, $\frac{\omega_i^2}{2 l_n}$ if $i = j$.

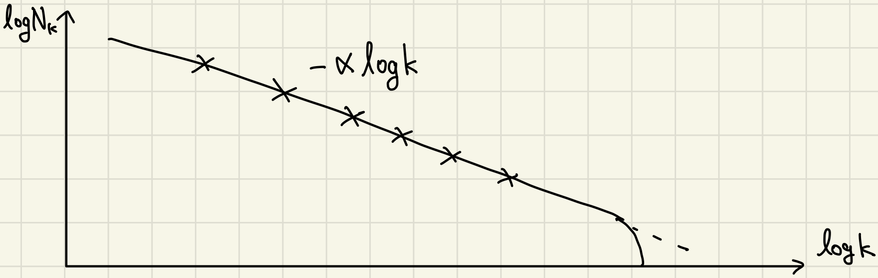
(we assume $\max_i \omega_i \leq \sqrt{l_n}$)

$\rightarrow$ Expected degree $E[d_i] = 2 A_{ii} + \sum_j A_{ij} = \sum_j \frac{\omega_i \omega_j}{l_n} = \omega_i$.

$\rightarrow$ Distribution of d_i: If we assume that $\frac{\omega_i}{\sqrt{l_n}} \to 0$, HW $\Rightarrow d_i \sim Po(\omega_i)$ as $n \to \infty$.

($\because d_i = \sum X_{ij}$ where $X_{ij} \sim Be\left(\frac{\omega_i \omega_j}{l_n}\right) \Rightarrow d_i \xrightarrow{d} Po(\omega_i)$)

Fact: If $N_k := (\#\text{ of } i \text{ s.t. } \omega_i \geq k) \sim n k^{-\alpha}$, $\tilde{N}_k := (\#\text{ of } d_i \text{ s.t. } d_i \geq k) \sim n k^{-\alpha}$.



Thm) If $\hat{\omega}$ is "nice" $\left(\frac{1}{n} \sum \omega_i \to c,\ \tilde{c} = \frac{1}{l_n} \sum \omega_i^2 < \infty\right)$, then:

i) $\exists$ a unique giant, ii) $\operatorname{diam} G \sim \log_{\tilde{c}} n$. [Proof omitted]

Model 2 [Molloy-Reed '95]: Fix a degree sequence $\underset{\sim}{d} = (d_1, \dots, d_n)$.

Define $G \sim UG(\underset{\sim}{d})$ uniformly among all graphs G with those degrees.

Q: Does there exist even one such G for some $\underset{\sim}{d}$?

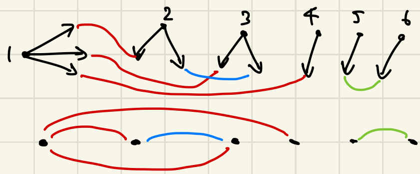
ex1) $\sum d_i = \text{odd} \Rightarrow$ DNE due to handshake lemma.

ex2) $d_1 = d_2 = 3$, $n = 2$. $\Rightarrow$ only multigraph! 

Thm) $\exists$ such a graph if the following algorithm does not output FAIL.

Havel–Hakimi–Aly ['55]:

ex) $\underset{\sim}{d} = (3, 2, 2, 1, 1, 1)$



- Start from $d_1 \geq d_2 \geq \cdots \geq d_n$.
- Set $E(G) \leftarrow \emptyset$.

LOOP:

- If $\underset{\sim}{d} = \emptyset$, output G.
- If $n = 1$ or $d_1 \geq n$, declare FAIL.
- Else:

  Pair v_1 to $v_2, \ldots, v_{d_1+1}$.

  Decrease d_i by 1 $\forall i = 2, \ldots, d_{i+1}$, and d_1 by d_1.

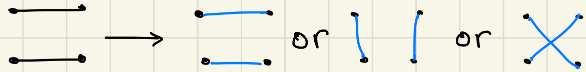
  Remove all vertices with $d_i = 0$ from $\underset{\sim}{d}$.

  Relabel d_i to be contiguous starting from $i = 1$.

GOTO LOOP

[Proof of correctness in HW]

Q: How to sample $G \sim UG(\underset{\sim}{d})$??

Subroutine: Given G with degree seq. $\underset{\sim}{d}$, choose 2 random disjoint edges and randomly rematch.  $=:$ SWAP.

Algorithm: Input $(\underset{\sim}{d}, t)$. Run H-H. If successful, run SWAP t times.

Thm*) a) If $t \to \infty$, we get the correct distribution.

b) If we want precision ε, $t = \mathrm{poly}(n, |\log(\varepsilon)|)$ suffices. (If $\max_i d_i \leq \sqrt{n}$)

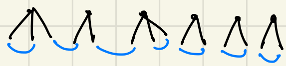
Q: What do we mean by precision ε?

↳ $d_{TV}(P, \tilde{P}) := \max_{A \in F} (|P(A) - \tilde{P}(A)|)$ (Total Variation Distance)

Equivantly, $d_{TV}(P, \tilde{P}) = \frac{1}{2} \sum_{w \in \Omega} |P(w) - \tilde{P}(w)|$.

Provisos*: 1) we need to modify SWAP to FLIP [Feder et al. '06]

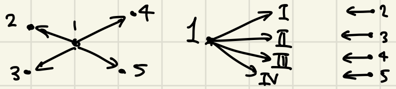
2) $t \in O(\log 1/\varepsilon + n^{41})$. 3) Bad in practice.

Configuration Model $CM(\underset{\sim}{d})$: Draw stubs. 

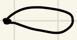
- Consider the $\ell_n = \sum d_i$ half-edges H. [ℓ_n should be even]
- Choose a random matching $\mathcal{M}$ for H.
- Output the multi-graph $G(\mathcal{M})$.

Analysis: 1) $\Pr[\mathcal{M}] = \frac{1}{(\ell_n - 1)(\ell_n - 3) \cdots 3 \cdot 1} = \frac{1}{(\ell_n - 1)!!}$.

2) $\Pr[\text{Multi-graph } G] = \frac{1}{(\ell_n - 1)!!} \times \mathcal{N}(G)$, $\mathcal{N}(G) := (\#\text{ of matchings yielding } G(\mathcal{M}) = G)$

ex) G is simple.  $\Rightarrow 4!$.

In general, we have $\mathcal{N}(G) = \prod_{i=1}^{n} \frac{d_i!}{A_{ii}!\,2^{A_{ii}}} \prod_{i<j} \frac{1}{A_{ij}!}$.

ex)  1 graph, 1 matching. $A_{ii}=1$, $A_{ij}=0$.

ex)  $d_1=d_2=4$. $A_{11}=0$, $A_{12}=4 \to \mathcal{N}(G) = \frac{d_1!\,d_2!}{A_{12}!} = 4!$

Obs 1) All simple graphs with the same degree distribution have the same weight $\frac{\prod(d_i!)}{(\ell_n-1)!!} \Rightarrow$ Conditioned on G being simple, uniform sample

↳ we can do rejection sampling!

Obs 2) Not all multigraphs have the same weight.

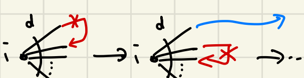
Thm) Under suitable assumptions*, # loops $\to Po\left(\frac{\nu}{2}\right)$, # multiedges $\to Po\left(\frac{\nu^2}{4}\right)$,

$\Pr[G \text{ is simple}] \to e^{-\left(\frac{\nu}{2}+\frac{\nu^2}{4}\right)}$ where $\nu := \frac{E[D(D-1)]}{E[D]}$ and with assumptions,

(D) $\frac{1}{n}\sum_i \mathbb{1}\{d_i=k\} \to \Pr[D=k]$ for some RV $D \in \mathbb{N}$.

(M1) $\frac{1}{n}\sum_i d_i \to E[D] < \infty$. (Notion of Empirical Distribution)

(M2) $\frac{1}{n}\sum_i d_i^2 \to E[D^2] \leq \infty$. $\to$ ($\Pr[G \text{ is simple}] > 0$ iff $E[D^2] < \infty$)

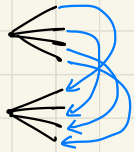
Proof Heuristic: $M_{ii} :=$ # of loops at i. $\Pr[M_{ii}=0] =$ 

$\left(1-\frac{d_i-1}{\ell_n-1}\right)\left(1-\frac{d_i-2}{\ell_n-3}\right)\cdots = \prod_{k=1}^{d_i-1}\left(1-\frac{d_i-k}{\ell_n-2k+1}\right) \overset{d_i \ll \ell_n}{\simeq} \prod_{k=1}^{d_i-1}\left(1-\frac{d_i-k}{\ell_n}\right) \simeq \exp\left(-\sum_{k=1}^{d_i-1}\frac{d_i-k}{\ell_n}\right) = e^{-\frac{(d_i-1)d_i}{2\ell_n}}$.

Now we cheat a little bit. $\Pr[M_{ii}=0\ \forall i] \simeq \prod_i e^{-\frac{d_i(d_i-1)}{2\ell_n}} = \exp\left(-\sum_i \frac{d_i(d_i-1)}{2\ell_n}\right)$

$\to \frac{\sum_i d_i(d_i-1)}{\ell_n} = \frac{\sum_i d_i(d_i-1)}{\sum_i d_i} = \frac{\frac{1}{n}\sum_i d_i(d_i-1)}{\frac{1}{n}\sum_i d_i} \longrightarrow \frac{E[D^2]-E[D]}{E[D]}$.

Where are we cheating?



1) different is are not independent.

2) ℓ_n in the denominator keeps going down, until $\ell_n - (\ell_n - 2) = 2$. [found in vHof, Vol 1]

$\rightarrow$ We need to control the dependence through unlikeliness of hitting exactly one i.

Properties of $CM(\underline{d})$ [vHof, Vol 2]

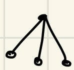
Thm 2) If (D) & (M1) hold, and $d_{min} \geq 3$, w.h.p. $G \sim CM(\underline{d})$ is connected.

Thm 3) If (D), (M1), & (M2) hold and $\Pr[D=2] < 1$, $\frac{|C_{max}|}{n} \rightarrow \theta$, $\theta > 0 \Leftrightarrow \nu > 1$.

Thm 4) If // , if $\nu > 1$ and $x, y \underset{uar}{\in} C_{max}$, then $\frac{dist(x,y)}{\log_\nu n} \xrightarrow{P} 1$ provided $\nu < \infty$. If $\nu = \infty$, and $\Pr[D=k] \sim k^{-\alpha}$ where $\alpha \in (2,3)$, $dist(x,y) \simeq \log \log n$.

Discussion: Let $\Pr[D=k] = k^{-\alpha}$. If $\alpha = 3$, $E[D] \sim \sum_k k \cdot k^{-3} = \sum_k k^{-2} < \infty$. But $E[D^2] \sim \sum_k k^2 \cdot k^{-3} = \sum_k k^{-1} \sim \log n$. If $\alpha < 3$, $\frac{|C_{max}|}{n} \rightarrow \theta > 0$, avg. dist $\sim \log \log n$. If $\alpha > 3$, $\nu \gtrsim 1$ depends on details, but if $\nu > 1$, then avg. dist $\sim \log n$.

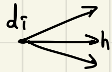
Consider BFS. What is the # of children in later generations of BFS?

For $G(n,p)$,  $X_1 \sim Bi(n-1, p)$. $\rightarrow \sim Bi(n-1-X_1, p) \simeq Po(c)$.

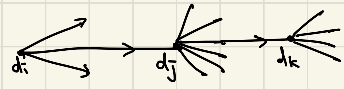
For BP, X_i is i.i.d. $\to \sim Po(c)$.

For d-regular graphs,  eg. if $d=3$, forward degree is $d-1=2$ (except root)

So what happens in CM?

Select i u.a.r. It has degree d_i. 

half-edge h matches u.a.r. to some h'. $\to Pr[h'] \simeq \frac{1}{\ell_n}$. $Pr[h'$ belongs to some $j \in [n]] \simeq \frac{d_j}{\ell_n}$.

How many children does j have? $(d_j - 1)$. And so on. 

Root has k children w.p. $\frac{1}{n}\sum_i \mathbb{1}\{d_i = k\} = P_k$.

Later vertices have degree $(k+1)$ and k children w.p. $P_k^* \sim \sum_j \mathbb{1}\{d_j = k+1\}\frac{d_j}{\ell_n}$

$= \frac{(k+1)P_{k+1}}{\ell_n} = \frac{(k+1)P_{k+1}}{\bar{d}n}$.

We get a 2-stage BP with offspring distribution:

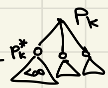
P_k for root, $P_k^* = \frac{(k+1)P_{k+1}}{\ell_n}$ for later generations.

Def) T_{D,D^*} 2-Stage BP: $D \in \mathbb{N}$, $Pr[D^* = k] = \frac{(k+1)Pr[D=k+1]}{E[D]}$.

$\to E[D^*] = \sum_k k\,Pr[D^* = k] = \frac{1}{E[D]}\sum_k k(k+1)Pr[D=k+1] = \frac{1}{E[D]}\sum_k (k-1)k\,Pr[D=k]$

$= \frac{1}{E[D]}E[D(D-1)]$, and this is just ν.

$\eta = G(\eta^*)$ ↰ $\sum_k P_k \eta^{*k}$

$\to Pr[|T_{D,D^*}| < \infty] =$ 

$= \sum_k P_k \underbrace{Pr[|T_{D,D^*}| < \infty]}_{\eta^*} = \sum_k P_k \eta^*$ where $\eta^* = G_{D^*}(\eta^*)$.

$\Rightarrow \frac{|C_{max}|}{n} \to (1-\eta)$.

Preferential Attachment

[Barabasi, Albert, '99] Models for Internet

[Yules, 1925] Model for evolution of species

[Zipf, 1929] Frequency of k'th most frequent word is $\propto 1/k$

[Simons, '55] First mathematical model, rigorous proof of power law

Model of random, growing graph G_t:

Start with some finite graph G_{t_0}.

Given G_t, a new vertex v "arrives" and sends m edges in G_t,

$(vu_1, \ldots, vu_m)$ where $\Pr[v \text{ chooses } u_i] := P_u(G_t) = \dfrac{d_u(t)}{2|E(G_t)|}$.

3 Versions:

Ind.: $u_1, \ldots, u_m$ are i.i.d. (hypergraphs allowed)

Cond.: $u_1, \ldots, u_m$ are i.i.d., conditioned on being pairwise different

Seq.: Update the degrees in between

$d_i(t+1) - d_i(t) \sim \mathrm{Bi}\left(m, \underbrace{\dfrac{d_i(t)}{2|E(G_t)|}}_{}\right)$. Given "appropriate" start, $|V_t| = t+1$, $|E_t| = mt$.

$\dfrac{d_i(t)}{2mt}$

"Appropriate" start is $t_1 = t_2 = 1$, $t_3 = 2$, $t_i = i-1$ in general.

$$E[d_i(t+1)\mid G_t] = d_i(t) + m\frac{d_i(t)}{2mt} = d_i(t)\left(1+\frac{1}{2t}\right).$$

$\rightarrow E[d_i(t)] = m\prod_{l=t_i}^{t-1}\left(1+\frac{1}{2l}\right)$ where t_i is the time vertex i is born (w/ degree m)

$$\rightarrow E[d_i(t)] \simeq m\prod_{l=t_i}^{t-1} e^{1/2l} = m\exp\left(\sum_{l=t_i}^{t-1}\frac{1}{2l}\right) \simeq m\exp\left(\int_{t_i}^{t}\frac{1}{2l}\right) = me^{\frac{1}{2}(\log t - \log t_i)} = m\sqrt{\frac{t}{t_i}}.$$

Thm) $\exists$ RV $X_n \in \mathbb{R}_+$ s.t. $\frac{d_i}{\sqrt{t}} \rightarrow X_i$.

Proof Sketch: a) $\frac{E[d_i(t)]}{\sqrt{t}}$ converges. b) $M_t = \frac{d_i(t)}{E[d_i(t)]}$ is a martingale.

(Recall that $E[M_{t+1}\mid M_1,\dots,M_t] = M_t$ for a martingale)

Largest Degree: in expectation, $\max_i E[d_i(t)] = m\sqrt{t}$.

Thm) $\exists$ RV $X_{max} \in \mathbb{R}_+$ s.t. $\frac{\max_i d_i(t)}{\sqrt{t}} \longrightarrow X_{max}$.

Degree Distribution: what is the degree of a random $i \in [t+1]$? (Heuristic)

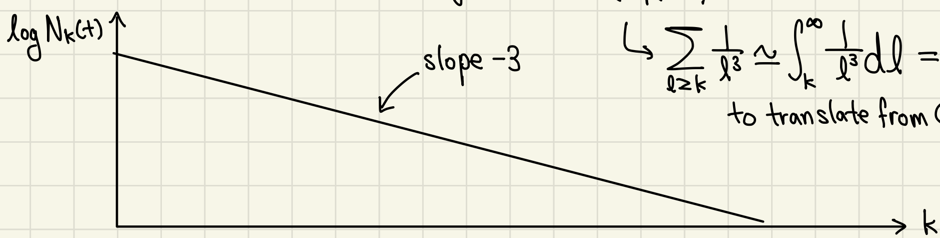
$i \simeq Ut$ where $U \overset{uar}{\sim} [0,1]$. $E[d_i(t)] = m\sqrt{\frac{t}{t_i}}$. As $t\to\infty$, $t_i \simeq i \simeq Ut$.

$\rightarrow E[d_i(t)] \simeq m\sqrt{\frac{1}{U}}$. $Pr[d_i \geq k] = Pr_U\left[m\sqrt{\frac{1}{U}} \geq k\right] = Pr_U\left[U \leq \frac{m^2}{k^2}\right] = \frac{m^2}{k^2}$.

Let $N_k(t) :=$ # of vertices of degree $k \simeq \left(\frac{2m^2}{k^3}\right)t$.

$\hookrightarrow \sum_{l\geq k}\frac{1}{l^3} \simeq \int_k^\infty \frac{1}{l^3}dl = \frac{2}{k^2}$

to translate from CDF



Thm) $\frac{N_k(t)}{t} \xrightarrow[Pr]{} \rho_t := \frac{2m(m+1)}{k(k+1)(k+2)}$ (Exact formula for the heuristic above)

Proof: $Z_k(t) := E[N_k(t)]$. Recall that $d_k(t+1) = d_k(t) + Bi\left(m, \frac{d_k(t)}{2mt}\right)$.

$$Z_k(t+1) = Z_k(t) + Z_{k-1}(t)\cdot \underbrace{\frac{k-1}{2t}}_{Pr[d=k-1 \text{ moves up to } k]} - Z_k(t)\underbrace{\frac{k}{2t}}_{Pr[d=k \text{ moves up from } k]} + \overbrace{\mathbb{1}\{m=k\}}^{\text{new vertex starts at } m} + \underbrace{O(1/t^2)}_{\text{lower order terms, can ignore}}$$

$\to$ jumps multiple buckets

Let's try $Z_k = P_k t$. $P_k(t+1) = P_k t + P_{k-1}\cancel{t}\frac{(k-1)}{2\cancel{t}} - P_k\cancel{t}\frac{k}{2\cancel{t}} + \delta_{k,m}$.

If $k>m$, $P_k = P_{k-1}\frac{(k-1)}{2} - P_k\frac{k}{2} + \delta_{k,m} \to P_k\left(1+\frac{k}{2}\right) = P_{k-1}\frac{(k-1)}{2} + \delta_{k,m}$.

$\to P_m = \frac{1}{1+m/2} = \frac{2}{2+m}$, $P_{k+1} = \left(\frac{k-1}{2+k}\right)P_k \to P_k = \prod_{l=m+1}^{k}\left(\frac{l-1}{2+l}\right)\cdot\frac{2}{2+m} = \frac{2m(m+1)}{k(k+1)(k+2)}$.

... Technically, we only prove that P_k is that term IF the limit exists.

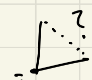
For completeness, we must show that $Z_k = t(P_k + \delta_k(t))$ and $\delta_k(t) \to 0$.

Also, we must prove concentration of $N_k(t)$ around $Z_k(t)$.

This takes some work, but can be proven. [Full proof in vHof] //

Community Structures

Local Communities: Define two clustering coefficients,



$$C_{loc} := \frac{1}{n}\sum_{i\in V} Pr[\text{2 random friends of } i \text{ are friends}] = \frac{1}{n}\sum_i \frac{\#\{\text{triangles} \ni i\}}{\binom{d_i}{2}} = \frac{1}{n}\sum_i C_i$$

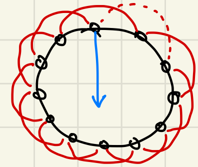
$$C_{glob} := \frac{3\cdot\#\Delta}{\sum_i\binom{d_i}{2}} = \frac{\sum_i\binom{d_i}{2}C_i}{\sum_i\binom{d_i}{2}}$$

(also average diam. $\simeq \log n$)

For $G(n,p)$, $CM(\underline{d})$, and PA, C_{loc} & $C_{glob} \to 0$ as $n\to\infty$. Not realistic!

Recipe 1: Start with local communities and $\overline{diam} = n^\alpha$. Add long-range edges. $\overline{diam} = n^\alpha \to \log n$.

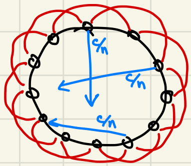
[Watts, Strogatz '98]: Start with cycle C_n. $\to$ $\overline{\text{diam}} = n/2$.



Add edges upto dist. k. $\to$ $\overline{\text{diam}} \simeq \frac{n}{2k}$. Rewire edges with prob. p, loosen 1 endpoint and randomly reattach.

$\to$ Blue (reassigned) edge graph has avg. degree pk $\to$ like $G(n, \frac{c}{n})$, $c = pk$

$\to$ Heuristically, if $k^2 p$ is constant, $\overline{\text{diam}} = \frac{\log n}{\text{const.}}$

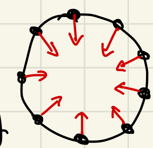


Recipe 2: Start with $\overline{\text{diam}} \simeq \log n$. Add local communities.

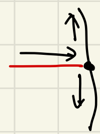
[Newman, Watts '99]: Start with cycle C_n. Add edges up to dist. k. Put $G(n, \frac{c}{n})$ on top. $\to$ long range avg. degree c.

[Barbour, Reiner '06]: $\frac{\overline{\text{diam}}}{\log_v n} \xrightarrow[\Pr]{} 1$ where $v := \frac{1}{2}\Big[\underbrace{c+1+\sqrt{(c+1)^2+4(2k-1)}}\Big]$, $\geq c+1$

(v is also the λ_{max} of $\begin{bmatrix} c & c\cdot 2k \\ 1 & 1 \end{bmatrix}$)



[Bollabas, Chung '88]: Start with cycle C_n ($k=1$). Randomly pair stubs from each vertex. $\to$ local deg. $=2$, long (global) deg. $=1$.



$\to$ If I arrive via a long edge, I must use a local edge. $\to$ $\begin{array}{cc} B & R \\ \end{array}$ $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

$\to$ $\lambda_{max} = 2$, $\frac{\overline{\text{diam}}}{\log_2 n} \xrightarrow[\Pr]{} 1$.

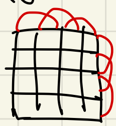
... Not a good model for social networks. Why?

- Milgram Experiment: What if we only know our immediate neighbors?

Myopic Agent: knows physical distances, knows friends and where they live.

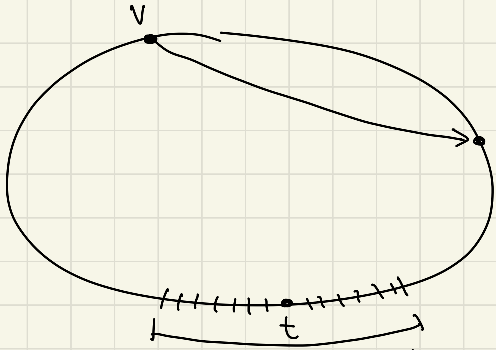
Thm) For NW model, it takes a myopic agent n^{α} time to find a random v.

[Kleinberg]: Start with a big piece of $\mathbb{Z}^d$. Add local edges up to dist. k. $\forall i \in V$, add l long range edges where the other endpoint j is chosen w.p. $\propto \frac{1}{\|i-j\|_2^{\alpha}}$ for $j \neq i$.



Thm) If $\alpha = d$, then the myopic agent find a random target in time $(\log n)^{\beta}$ where $\beta(d) < \infty$. If $\alpha \neq d$, it takes time n^{const}.

Proof Sketch:



$T = \{x \mid dist_{phy}(x,t) < \sqrt{n}\}$

$Pr[v \text{ has a neighbor in } T] \simeq \frac{1}{\sqrt{n}}$. If not, go to another neighbor. Retry.

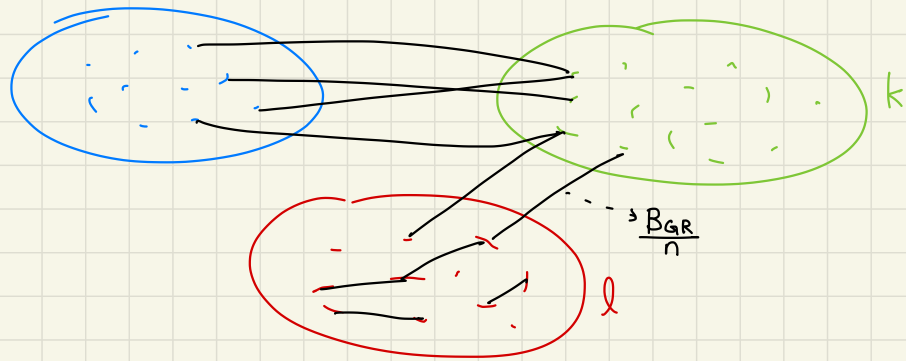
Finds T in $Geo(\frac{1}{\sqrt{n}}) \to$ Exp. time $\sqrt{n}$.

Global Communities

Stochastic Block Model: q communities. $B := q \times q$ matrix. $\alpha :=$ PD on $[q]$.

Each vertex i chooses color $k(i) \sim \alpha$ i.i.d. $V_k := \{i \in [n] \mid k(i) = k\}$.

$\to n_k := |V_k| \simeq \alpha_k \cdot n$.



Each pair (i,j) connects w.p. $p_{ij} := \frac{B_{k(i),k(j)}}{n}$.

# of edges from a vertex in k to $l = Bi(n_l, \frac{B_{k(i),k(j)}}{n}) \simeq Po(C_{kl})$

where $C_{kl} := B_{kl}\frac{n_l}{n} \simeq B_{kl}\alpha_l$.

$\rightarrow C_{kl}$ describes a BFS tree for $SBM_n(B,\alpha)$.

Def) Irreducible: B s.t. $\exists t<\infty$, $B^t_{kl} > 0$.

$\Leftrightarrow \forall k,l$, $\exists$ path $k=k_0,\dots,k_t=l$ s.t. $\prod_{i=0}^{t} B_{i-1,i} > 0$.

Thm) If B is irreducible, then:

$\lambda_{max}(C)<1 \Rightarrow |C_{max}| \in O(\log n)$.

$\lambda_{max}(C)>1 \Rightarrow \frac{|C_{max}|}{n} \longrightarrow \theta > 0$, and second largest is $O(\log n)$.

Variants:

Degree-Corrected SBM: $p_{ij} := w_i w_j \frac{B_{k(i)k(j)}}{n}$ where w_i are weights,

and it is (usually) normalized to $\frac{1}{n_k}\sum_{i\in V_k} w_i = 1$.

→ $D_k := \sum_\ell \frac{B_{k\ell} n_\ell}{n}$ is the average degree of a vertex in k. $= \sum_j p_{ij}$.

→ $p_{ij} := d_i d_j \frac{\tilde{B}_{k(i)k(j)}}{n}$ where $\hat{B}_{k\ell} := \frac{B_{k\ell}}{D_i D_j}$ and $d_i :=$ average degree of vertex i.

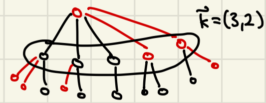
Overlapping Communities (Topic Model): Let $\Delta_q :=$ simplex in $\mathbb{R}^q$, a set of probability measures on $[q]$, i.e. $\Delta_q = \{\vec{\alpha} \in \mathbb{R}^q_+ \mid \sum_i \alpha_i = 1\}$.

Each vertex chooses some $\alpha(i) \in \Delta_q$. ["taste" of vertex i]

For $i<j$, $k_i \sim \alpha(i)$, $k_j \sim \alpha(j)$. $p_{ij} = \frac{B_{k(i),k(j)}}{n}$.

(Latent Dirichlet Models estimate the PD of $\alpha(i)$, out of scope)

Multitype BP: $Q := [q]$, the set of types (colors, labels).



RV $X_{k\ell}$:= # of children of type ℓ a parent of type k has.

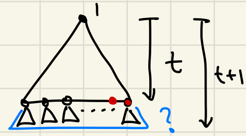
In SBM, $X_{k\ell} \sim Po(C_{k\ell})$. (*B-C model had $[X_{k\ell}] = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$)

$C_{k\ell} = E[X_{k\ell}]$. Let $\vec{X}_i := (X_{ij})_{j\in Q}$, $p_i(\vec{k}) := \Pr[\underbrace{\vec{X}_i = \vec{k}}_{(X_{ij}=k_j\ \forall j)}]$, $\vec{k} \in \mathbb{N}_0^q$.

$G_i(\vec{\lambda}) := \sum_{\vec{k}} p_i(\vec{k}) \prod_j \lambda_j^{k_j}$ is a PGF, $[0,1]^q \to [0,1]$. → $\vec{G}(\lambda) := (G_i(\vec{\lambda}))_{i\in Q}$.

(e.g. $\vec{k}=(3,2) \to \prod_j \lambda_j^{k_j} = \lambda_1^3 \lambda_2^2$) Some nice properties.

$\vec{G}_i(\vec{0}) = \Pr[\vec{X}_i = 0] = p_i(\vec{0})$, $\frac{\partial G_i(\vec{\lambda})}{\partial \lambda_j}\Big|_{\vec{\lambda}=1} = C_{ij} = E[X_{ij}]$.



$Z_{i\to j}(t)$:= # of offspring in generation t of color j if root has color i.

$\vec{Z}_i(t) := (Z_{ij}(t))$. $E[Z_{ij}(t+1) \mid \vec{Z}_i(t) = \vec{k}] = E[\sum_{\ell=1}^{q} Z_{i\ell}(t) C_{\ell\to j}] = E[(\vec{Z}_i(t) C)_j]$

→ $E[\vec{Z}_i(t+1) \mid \vec{Z}_i(t)] = E[\vec{Z}_i(t)]C$. → $E[Z_{ij}(t)] = C^t_{ij}$.

By Spectral Thm, if C is symmetric, $\sum_j C_{ij}^t \le \lambda_{max}^t(C) \cdot \text{const}$.

Let $\eta_i := \Pr[\exists t, \vec{Z}_i(t) = \vec{0}]$, the extinction probability.

$1-\eta_i(t) := \Pr[\sum_j Z_{ij}(t) > 0] \overset{UB}{\le} \sum_j \Pr[Z_{ij}(t) \ge 1] \overset{Markov}{\le} \sum_j E[Z_{ij}(t)] \le \text{const} \cdot \lambda_{max}^t(C)$.

$\to 1-\eta_i(t)$ decays if $\lambda_{max}(C) < 1$. $\Rightarrow \eta_i = 1$.

Thm) If $\Pr[\sum_j X_{ij} = 1] \ne 1$ for some i, $\eta_i(t) < 1 \iff \lambda_{max}(C) > 1$.

Def) Singular BP: $\Pr[\sum_j X_{ij} = 1] = 1 \ \forall i$. $\to$ gives a trivial, single line.

Directed Graphs & Bow-Tie Structure

Def) Di-graph: directed graph $D = (V, E)$

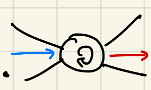
Def) DAG: no directional cycles

Def) SCC: $A \subseteq V$ is SCC if $\forall x, y \in A$, $\exists$ path $w: x \rightsquigarrow y \rightsquigarrow x$ & A is maximal.

Lemma: If we contract all SCCs into single vertices, we get a DAG.

Def) Fan-Out: $\{y \mid \exists w: A \to y\}$. Def) Fan-In: $\{y \mid \exists w: y \to A\}$.

If $A = \{x\}$, denoted as $C_+ := \{y \mid \exists w: x \to y\}$, $C_- := \{y \mid \exists w: y \to x\}$.

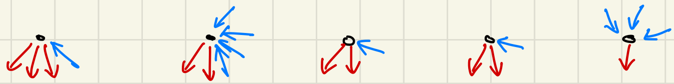


Def) Bow-Tie: If A is an SCC, $(L, A, R) := (FI(A) \setminus A, A, FO(A) \setminus A)$.

Some Models:

[Karp, '90] Random digraph $D(n,p)$: $V := [n]$, every possible edge is present i.i.d. w.p. p. (no self-loops for now)

[Cooper, Frieze, '02] Directed Configuration Model $CM(\underset{\sim}{d}^-, \underset{\sim}{d}^+)$:



$\underset{\sim}{d}^+, \underset{\sim}{d}^-$ s.t. $\sum_x \underset{\sim}{d}^+_x = \sum_x \underset{\sim}{d}^-_x = l_n$. Perform bipartite matching.

Oriented (Directed) Preferential Attachment: attach new arrivals $\propto$ degrees.

Thm) For $D(n,p)$, if $p = c/n$ and $c < 1$, $|C_\pm(x)| \in O(\log n)$ w.h.p.

If $c > 1$, $\exists$ a unique giant SCC s.t. $\frac{|SCC|}{n} \to \theta(c)^2$, $\frac{|L/R(SCC)|}{n} \to \overset{\theta(1-\theta)}{(\theta - \theta^2)}$

where $\theta := \Pr[T_{Po(c)} > \infty]$, and the rest is $O(\log n)$.

$\to \forall x \notin L \cup SCC \cup R$, $|C_+(x)|, |C_-(x)|, |SCC(x)| \in O(\log n)$.

Thm 2) The same holds for $CM(\underset{\sim}{d}^+, \underset{\sim}{d}^-)$ when it is "dense enough".

Thm 3) For directed PA, power laws for in & out degrees, correlated.

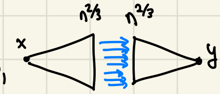
Proof of Thm 1: Use the following lemma$^{(2)}$: $\Pr[|C_+(x)| > k] \to \Pr[|T_{Po(c)}| > k]$.

Proof is corollary to another lemma$^{(1)}$: $\Pr_{D(n,p)}[|C_+(x)| > k] = \Pr_{G(n,p)}[|C(x)| > k]$.

Consider the BFS of $C_+(x)$. It's just the same procedure as BFS of $G(n,p)$.

Let $L_+ := \{x \mid |C_+(x)| > \frac{2c}{(c-1)^2} \log n\}$. $L = \{x \mid |C(x)| > \frac{2c}{(c-1)^2} \log n\}$ in $G(n,p)$

and $\frac{|L|}{n} \xrightarrow{P} \theta(c)$, thus $\underline{\frac{|L_+|}{n}, \frac{|L_-|}{n} \xrightarrow{P} \theta(c)}$ (Corollary 1).

Lemma 3: w.h.p., $\forall x \in L_+, y \in L_-, \exists w: x \rightsquigarrow y$. Proof by double BFS, 

$$\Pr[T_x, T_y \text{ are disconnected}] \le (1-p)^{n^{2/3} \cdot n^{2/3}} \le e^{-\frac{c}{n} \cdot n^{4/3}} = e^{-O(n^{1/3})} \to 0.$$ 

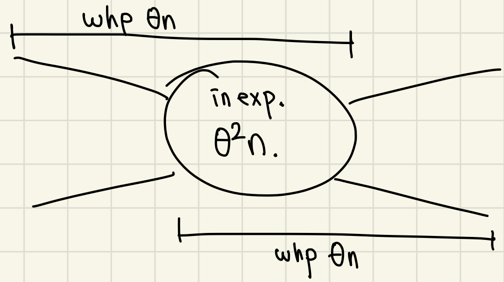
Corollary 2: w.h.p., $L_+ \cap L_-$ is SCC since $x \in L_+ \cap L_-$ must connect back to itself.

Lemma 4: $\Pr[|C_+(x)| \le k, |C_-(y)| \le k] = \Pr[|C_\pm(x)| \le k] \Pr[|C_\pm(x)| \le k](1 + O(\frac{k^2}{n}))$.

$\to \mathrm{Cov}(|C_+(x)| > k, |C_-(x)| > k) \le \frac{k^2}{n}$, which is small if k is small!

$\to E[|L_+ \cap L_-|] = \sum_x \Pr[x \in L_+ \cap L_-] = \sum_x \Pr[|C_+(x)| \ge K \log n, |C_-(x)| \ge K \log n]$

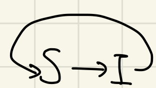
$= \sum_x \Pr[\sim] \Pr[\sim](1 + o(1)) = \sum_x \theta \cdot \theta + \text{error} = \theta^2$.



Epidemics and SIR

Compartment Model: $\underset{\text{Susceptible}}{S} \longrightarrow \underset{\text{Infected}}{I} \longrightarrow \underset{\text{Recovered}}{R}$

Homogeneous Mixing: assume anyone can infect anyone else, and everyone "behaves" in the same way.



Warm-up: SIS model. $\beta :=$ infection rate $(S \xrightarrow{\beta} I)$, $\gamma :=$ recovery rate $(I \xrightarrow{\gamma} S)$.

Initially, 1 person out of N is infected. $\to \hat{I}(0) = 1, \hat{S}(0) = n - 1$.

$I := \frac{\hat{I}}{N}$, $S := \frac{\hat{S}}{N}$. $\rightarrow S + I = 1$. $I(0) = I_0$, $S(0) = S_0$, $R_0 = \frac{\beta}{\gamma}$ (reproductive ratio)

$$\frac{dI(t)}{dt} = \beta S(t) I(t) - \gamma I(t), \quad \frac{dS(t)}{dt} = -\frac{dI(t)}{dt} = -\beta S(t) I(t) + \gamma I(t).$$

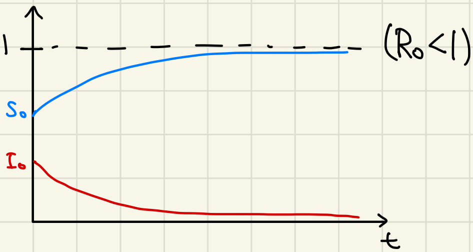
$$\hookrightarrow \frac{dI}{dt} = \beta I(1-I) - \gamma I. \quad \frac{d \log I}{dt} = \frac{1}{I}\frac{dI}{dt} = \beta(1-I) - \gamma \leq \beta - \gamma.$$

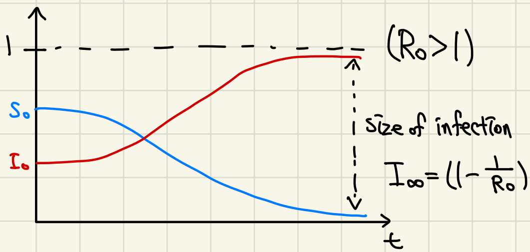
Assume $R_0 < 1$. [early stage approximation] $\rightarrow \beta - \gamma < 0$.

Take integral, $\log I - \log I_0 \leq (\beta-\gamma)t \rightarrow I(t) \leq I_0 e^{(\beta-\gamma)t}$.

Also, $\beta - \gamma = -\gamma(1-R_0)$, so $I(t) \leq I_0 e^{-\gamma(1-R_0)t}$.

$\Rightarrow R_0 < 1$ decays exponentially, $R_0 > 1$ grows exponentially.





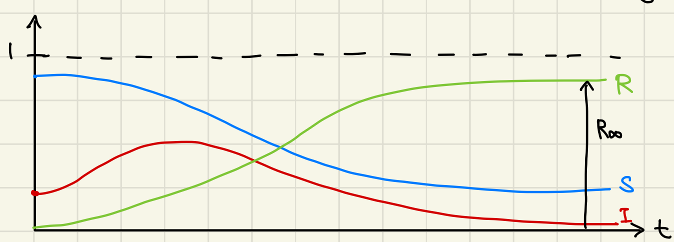
SIR: $S \xrightarrow{\beta} I \xrightarrow{\gamma} R$. $S + I + R = 1$. $I(0) = I_0$, $S(0) = 1 - I_0$, $R(0) = 0$.

$$\frac{dI}{dt} = \beta S I - \gamma I, \quad \frac{dS}{dt} = -\beta I S \leq 0, \quad \frac{dR}{dt} = \gamma I \geq 0.$$

Stationary Solution: $(S, I, R) = (x, 0, y)$ is always a solution $\forall t$. (no disease)

In general, $\frac{dS}{dt} \leq 0$, so $S \searrow S_\infty$, and $\frac{dR}{dt} \geq 0$, so $R \nearrow R_\infty$, and $I(t) \rightarrow 0$.

$R_\infty \rightarrow$ final size of infection, $(1-R_\infty) = (1-I_0)e^{-R_\infty R_0}$ yields R_∞.



Stochastic Model: Same initial conditions, but $(N+1)$ total people.

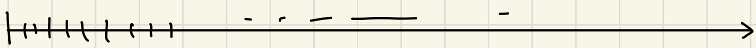
If person i is infected, they stay infected until $T_i \sim \mathrm{Exp}(\gamma)$. PDF: $\gamma e^{-\gamma x}$, $\Pr[X \geq T] = e^{-\gamma T}$, $\Pr[X < T] = 1 - e^{-\gamma T}$.

While infected, they try to infect every susceptible j at time $T_{ij} \sim \mathrm{Exp}(\beta)$.

If $T_{ij} < T_i$, j becomes infected.

$$\left[\begin{aligned}\Pr[T_{ij} < T_i] &= \int \Pr[T_i > T_{ij} = k] \Pr[T_{ij} = k]\, dk \\ &= \int e^{-\gamma k} \cdot \beta e^{-\beta k}\, dk \\ &= \frac{\beta}{-(\gamma+\beta)} e^{(-\gamma-\beta)k}\Big|_0^\infty = -\frac{\beta}{(\gamma+\beta)}[0-1] = \frac{\beta}{(\gamma+\beta)}\end{aligned}\right]$$

Rate β Poisson Process



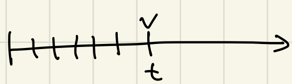
At time $t \in \varepsilon \mathbb{N}$, click w.p. $\varepsilon\beta$.

$\#\{\text{clicks in } [a,b]\} = \mathrm{Bi}\left(\frac{(b-a)}{\varepsilon}, \beta\varepsilon\right) \xrightarrow{\varepsilon \to 0} \mathrm{Po}(\beta(b-a))$

Po_β on $\mathbb{R}^+$ is a random point process.

1) $\#\{\text{clicks in interval } I\} = \mathrm{Po}(\beta|I|)$.

2) If $I_1 \cap I_2 = \emptyset$, $\#\{\text{clicks in } I_1 \cup I_2\} = \#\{\text{clicks in } I_1\} + \#\{\text{clicks in } I_2\}$.

Waiting time for first click 

$\Pr[T \geq t] = (1-\beta\varepsilon)^{t/\varepsilon} \xrightarrow{\varepsilon \to 0} e^{-\beta t}$, arrival times $T_1, T_2, \ldots \overset{iid}{\sim} \mathrm{Exp}(\beta)$.

↳ Memoryless property!

Stochastic Model of SIR

$(n+1)$ people, at time 0, n susceptible, 1 infected.

At any time t, if i is infected, it recovers at rate γ and infects susceptible individuals at rate $\beta = \gamma/n$. $\rightarrow$ $j \in S$ moves to I at rate (βI).

Recovery clock $T_i \sim \text{Exp}(\gamma)$. Infection clock $T_{ij} \sim \text{Exp}(\beta)$.

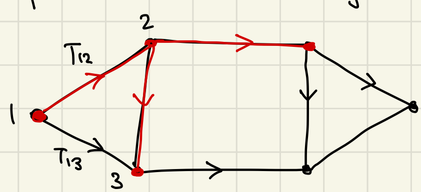
R_0, Branching Processes, Infection Digraph

On some graph $G(V,E)$.

If i gets infected, stays infected for time $T_i \sim \text{Exp}(\gamma)$.

Infects susceptible neighbors at rate β, i.e. after time $T_{ij} \sim \text{Exp}(\beta)$.

How to simulate this?



Predraw $T_i \; \forall j$, $T_{ij} \; \forall ij \in E$.

If $T_{ij} > T_i$, set $T_{ij} = \infty$, else if $T_{ij} < T_i$, set (i,j) as open.

↳ Gives a random digraph D.

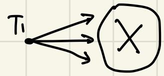
If i gets infected at time t_i, then j gets infected at time $t_j = \overbrace{\min_{w: i \to j} \sum_{e \in w} T_e}^{\text{shortest path } i \to j}$.

$\rightarrow$ Set of eventually infected nodes if epidemic starts at i: $\underset{\text{out set}}{C_+(i)}$ in D.

Infection Tree: Tree along which the infection passes, root is patient 0.

Backwards Process: People who could have infected me, $C_-(j)$ in D.

Consider again homogeneous mixing, $G = K_{n+1}$. For simplicity, $\gamma = 1$, $\beta = c/n$.

Forward Branching Process, T_{SIR}: at $t = 0$, node 1 is infected. 

$Pr[T_{1j} < T_1 | T_1] = \int_0^{T_1} e^{-\beta t} \beta \, dt = 1 - e^{-\beta T_1} =: p_{T_1}$.

$Pr[X = k | T_1] = \binom{n}{k} p_{T_1}^k (1 - p_{T_1})^{n-k} \sim Bi(n, p_{T_1})$.

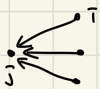
$Pr[X = k] = \int e^{-T} dT \binom{n}{k} p_{T_1}^k (1 - p_{T_1})^{n-k}$. [detail in HW, assume $\gamma = 1$]

$\rightarrow T_{SIR}$ is BP with offspring distr. given by above $Pr[X = k]$.

$E[X | T_1] = n p_{T_1}$. $E[X] = E_{T \sim Exp(1)}[n p_{T_1}] = n \frac{\beta}{\beta + 1} = n \frac{c/n}{c/n + 1} \xrightarrow{n \to \infty} c =: R_0$.

[In HW, show that $\theta_{SIR} = Pr[|T_{SIR}| = \infty] = 0$ if $c \leq 1$, $\rightarrow (1 - \frac{1}{c})$ if $c > 1$.]

[Also, if $c < 1$, $Pr[|T_{SIR}| > k] \leq \exp(-\frac{(1-c)^2}{4} k)$.]

Backwards Branching Process, T_{Back}: $Pr[T_{ij} < T_i | T_i] = p_{T_i} = 1 - e^{-\beta T_i}$. 

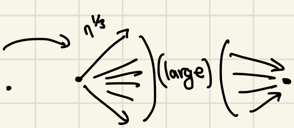
$Pr[i \text{ infects } j] = \int dT_i \, p_{T_i} = \bar{p} = \frac{\beta}{\beta + 1} = \frac{c/n}{c/n + 1}$.

$\rightarrow T_{Back}$ is BP with $Pr[X = k] = \binom{n}{k} \bar{p}^k (1 - \bar{p})^{n-k}$.

For large n, this behaves like $Poi(c)$. $\rightarrow \theta_{Back} = \theta_{Po(c)}$ is solution of $e^{-c\theta} + \theta = 1$.

Thm) If $c \leq 1$, $\underset{\lim_{t\to\infty} R(t)}{R_\infty}$ is $\overbrace{\text{bounded in probability}}^{\forall p, \exists n \text{ s.t. } Pr[R_\infty \geq n] \leq p}$. If $c > 1$, then:

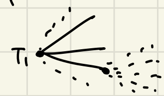
(a) $Pr[R_\infty \geq n^{1/3}] \leq 1 - 1/c$, (b) Conditioned on $R_\infty < n^{1/3}$, R_∞ is BIP, (bounded in prob.)

(c) Conditioned on $R_\infty \geq n^{1/3}$, $\frac{R_\infty}{n} \xrightarrow{Pr} \theta_{Po(c)}$. 

Proof Strategy: 1) BFS of $C_t(1)$. 2) Look at early stages, up to $n^{1/3}$ vertices. $\Pr[|C_t(1)|\le k] \simeq \Pr[|T_{SIR}|\le k]$. 3) In later stages, use deterministic approximation motivated by diff. eq.

BFS-Exploration of $C_t(1)$: Predraw $T_1, T_2, \dots \sim \mathrm{Exp}(1)$.

Explore at vertex 1, $X_1 \sim \mathrm{Bi}(n, p_{T_1})$ many children.



Then $X_2 \sim \mathrm{Bi}(n-X_1, p_{T_2})$, and so on. $X_t \sim \mathrm{Bi}(n-X_1-\dots-X_{t-1}, p_{T_t})$.

If $c \le 1$, $X_t \le X_t^+ \sim \mathrm{Bi}(n, p_{T_t}) \to \Pr[C_t(1) > k] \le \Pr[|T_{SIR}| > k] \le e^{-\frac{(c-1)^2}{4k}} \xrightarrow{k\to\infty} 0$.

Let $N_i := n - X_1 - \dots - X_i$. $\to X_i \sim \mathrm{Bi}(N_{i-1}, p_{T_i})$, $X_i^+ \sim \mathrm{Bi}(n, p_{T_i})$.

Lemma: $\Pr[|C_t(1)| \le k] = \Pr[|T_{SIR}| \le k] + O(\frac{k^2}{n})$.

Proof: If $|C_t(1)| \le k$, $N_i \ge n-k$. Then $X_i^+ - X_i \le \sum_{j=1}^{n} B_{ij} - \sum_{j=1}^{N_{i-1}} B_{ij} = \mathrm{Bi}(k, p_{T_i})$.

$\to \Pr[X_i \ne X_i^+] = \Pr[X_i^+ - X_i > 0] \le E[X_i^+ - X_i] = E_{T_i}[k \cdot p_{T_i}] = k\beta \overbrace{E[T_i]}^{\gamma=1} = k\frac{c}{n}$.

$\to \Pr[\exists i, X_i \ne X_i^+] \le \frac{k^2 c}{n}$. [Union bound over $\le k$ steps]

Corollary: 1) $\Pr[|C_t(1)| \ge n^{1/3}] = \Pr[|T_{SIR}| \ge n^{1/3}] + \overbrace{O(\frac{n^{2/3}}{n})}^{\to 0 \text{ as } n\to\infty}$

$= \Pr[|T_{SIR}| = \infty] + \overbrace{\Pr[n^{1/3} \le |T_{SIR}| < \infty]}^{\to 0 \text{ as } n\to\infty} + o(1) \longrightarrow (1-\frac{1}{c})$.

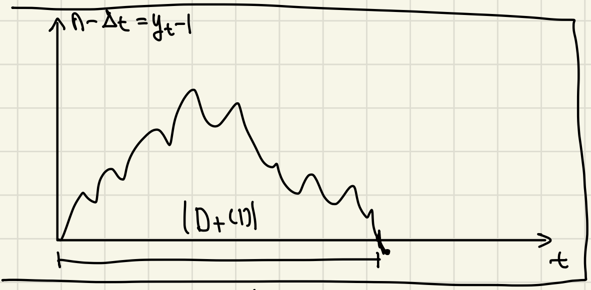
2) Conditioned on $|C_t(1)| \le n^{1/3}$, $|C_t(1)|$ is BIP.

The other side? $X_1 \sim \mathrm{Bi}(n, p_{T_1})$. $N_1 = n - X_1 \sim n - \mathrm{Bi}(n, p_{T_1}) = \mathrm{Bi}(n, 1-p_{T_1})$.

$X_2 \sim \mathrm{Bi}(N_1, p_{T_2})$. $N_2 = N_1 - X_2 \sim N_1 - \mathrm{Bi}(N_1, p_{T_2}) = \mathrm{Bi}(N_1, 1-p_{T_2})$

$= Bi(n,(1-P_{T_1})(1-P_{T_2}))$. Recall $1-P_{T_i} = e^{-\beta T_i}$. $\to N_t \sim Bi(n, \exp(-\beta\sum_{i=1}^{t} T_i))$.

$N_t = n-\Delta_t$ where $\Delta_t = X_1+\cdots+X_t$. $|D_+(1)| = \min_t\{t \mid \Delta_t - t\} = \min_t\{t \mid \overset{1+\Delta_t-t}{y_t} = 0\}$.



$\Delta_t = n - N_t \sim Bi(n, 1-\exp(-\beta\sum_{i=1}^{t} T_i))$.

Intuition: if LLN holds, $\sum_{i=1}^{t} T_i \simeq t + o(1)$.

$\hookrightarrow \Delta_t \simeq Bi(n, 1-e^{-\frac{c}{n}t}) \simeq n(1-e^{-\frac{c}{n}t})$

$\frac{\Delta_t - t}{n} = 1 - e^{\frac{ct}{n}} - \frac{t}{n}$. Let $\theta := \frac{t}{n}$. $\to \frac{\Delta_t - t}{n} = 1-e^{-c\theta}-\theta = 0 \Leftrightarrow \theta + e^{-c\theta} = 1$.

Then $\theta = \Pr[|T_{Po(c)}| = \infty]$.

Diff. Eq.: $\frac{dS}{dt} = -cSI \to S(t) = S(0)e^{-c\int I dt} = S(0)e^{-cR(t)}$. $[\frac{dR}{dt} = I]$

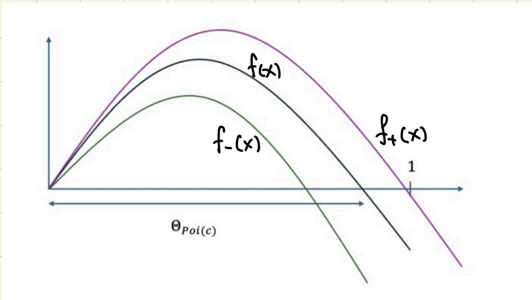
$S(t)+I(t)+R(t) = 1. \to I(t) = 1-R(t)-e^{-cR(t)}$ [for $S(0)=1-o(1)$].

Intuition: $\frac{\Delta_t - t}{n}$ has the "same form" as $I(t)$ since it is also the "ratio" of people who are currently infected if somebody recovers every timestep.

Lemma: w.p. $\geq 1-4e^{-t\varepsilon^2\cdot\frac{1}{11}}$, we can bound $\frac{\Delta_t - t}{n}$ by $f_\pm(\frac{t}{n})$ where

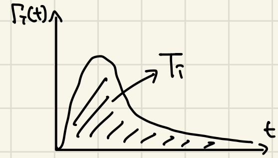
$f_\pm(x) = (1-e^{-c(1\pm\varepsilon)x})(1\pm\varepsilon) - x$. [Proof in Notes]

Proof sketch: Make the LLN claim & $\Delta_t \simeq n(1-e^{-\beta t})$ into concentration. It should work since they are binomials.



[$f(x) = 1-e^{-cx}-x$, and the first time it crosses $x=0$ is $\theta_{Po(c)}$.]

Realistic Epidemic Models



Curve of Infectiousness: β varies over my time of infection.

↳ in empirical work, fix the form $\Gamma_i(t) \simeq t^{\alpha} e^{-t\rho}$. Then we can choose parameters $A_i := \int \Gamma_i(t)\,dt$, $B_i := \int t\,\Gamma_i(t)\,dt$, kind of weight & expectation.

Contact Network: $G=(V,E)$, weights β_{ij}.

ex) Small community, $G \sim G(n,p)$. Everybody is equally likely to be friends.

Phone network (Copenhagen network).

Large populations, $CM(\underset{\sim}{d})$ where $\underset{\sim}{d} \sim \text{LogNormal}$. ← long tail distr. for superspreaders

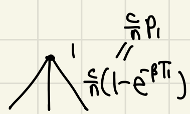
Infection Digraph: i gets infected at time t_i, infects j at rate $\beta_{ij}\Gamma_i(t-t_i)$.

→ $PPP(\beta_{ij}\Gamma_i)$, in $[a,b]$, # of clicks of Poisson clicks is $Po(\beta_{ij}\int_a^b \Gamma_i(t)\,dt)$.

→ $Pr[i \text{ can infect } j] = P_{ij} = 1-\exp(-\beta_{ij}\overbrace{\int_0^\infty \Gamma_i(t)\,dt}^{=:T_i})$, whether the clock ever clicks.

→ D_{SIR} conditioned on "T_i", $Pr[ij \in E] = P_{ij}$ are independent $\forall i \neq j$.

Example: $G \sim G(n+1, \frac{c}{n})$, $\beta_{ij}=\beta$, $T_i \sim \mu$. Look at BFS.



→ $X_1 \sim Bi(n, \frac{c}{n}P_1)$, $X_2 \sim Bi(n-X_1, \frac{c}{n}P_2)$, …. $N_1 \sim n - X_1 = Bi(n, 1-\frac{c}{n}P_1)$,

… $N_t \sim Bi(n, \prod_{i=1}^{t}(1-\frac{c}{n}P_i))$, $|D_t(1)| = \min\{t \mid \overbrace{\Delta_t}^{y_{t-1}} - t < 0\}$, where $\Delta_t = n - N_t$.

$\Delta_t - t \sim Bi(n, 1-\prod_{i=1}^{t}(1-\frac{c}{n}P_i)) - t$. Approximate $\prod_{i=1}^{t}(1-\frac{c}{n}p_i) = \exp(-\sum_{i=1}^{t} p_i \frac{c}{n} + O(\frac{1}{n^2}))$

$\simeq \exp(-t\frac{c}{n}\bar{p})$, where $\bar{p} := E[p_i] = \underset{T\sim\mu}{E}[1-e^{-\beta T}]$.

Binomials are concentrated, so $\Delta_t - t \simeq n(1-e^{-t\frac{c}{n}\bar{p}} - \frac{t}{n})$ w.h.p.,

and this mimics the curve $1-e^{-R_0 x}-x$ for $x \leftarrow \frac{t}{n}$.

Thm) Let $R_0 := c\cdot\bar{p}$, the average # of infected "children" of patient 1.

1) If $R_0<1$, R_∞ is bounded in probability.

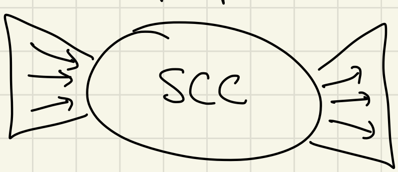
2) If $R_0>1$, then $R_\infty > n^{1/3}$ w.p. tending to P, where P is solution to the equation $(1-P = \underset{T}{E}[e^{-cPp_T}])$.

↳ Intuition: $d_i|_{T_i} \sim Bi(n, \frac{c}{n}p_i) \xrightarrow{n\to\infty} Po(cp_i)$, so $G(\lambda) = e^{-cp_i(1-\lambda)}$.

3) If $R_0>1$, conditioned on $R_\infty>n^{1/3}$, $\frac{R_\infty}{n} \xrightarrow{P} \theta$, solution of $(\theta + e^{-R_0\theta}=1)$.

→ Upshot: Details don't matter once the epidemic is large enough.

Different perspective: Forward & Backward BP. Look at D_{SIR}, a bow-tie.



$\rho \simeq \Pr[R_\infty\to\infty]$, $\theta = \lim \frac{R_\infty}{n}$, which are survival prob. of T_{SIR} (forward) & T_{Back} (Backward).

Conditioned on d_i, the distribution of FW & BW are $Bi(d_i,p_i)$, $Bi(d_i,\bar{p})$.

They have generating functions $G_d^+(\lambda) = \underset{T}{E}[1-p_T(1-\lambda)]^d$, $G_d^-(\lambda) = (1-\bar{p}(1-\lambda))^d$.

Insert $\Pr[d=k]=p_k$, then $G_d^+(\lambda) = \sum_k p_k G_k^+(\lambda) = \underset{T}{E}[\sum_k p_k(1-p_T(1-\lambda))^k]$

$G(\lambda) := \sum_k p_k \lambda^k$. Then $G_d^+(\lambda) = \mathbb{E}_T[G(1-P_T(1-\lambda))]$, $G_d^-(\lambda) = G(1-\bar{P}(1-\lambda))$.

For $G(n,\frac{c}{n})$, $d_i \sim Bi(n,\frac{c}{n}) \to Po(c)$, so $G(\lambda) = e^{-c(1-\lambda)}$.

$\to \eta = G^+(\eta) = \mathbb{E}_T[e^{-c(1-P_T(1-\eta))}] \to 1-\rho = \mathbb{E}_T[e^{-cP_T\rho}]$.

$\to \eta = G^-(\eta) = e^{-c\bar{P}(1-\eta)} \to 1-\theta = e^{-c\bar{P}\theta} = e^{-R_0\theta}$. Heuristic works!

Time Evolution of SIR

Let's start with $G(n,p)$. We can do a few tricks:

- Replace $A_{ij} \sim Ber(\frac{c}{n})$ by $A_{ij} \sim Po(\frac{c}{n}) \to \tilde{G}(n,\frac{c}{n})$.
- Explore $\tilde{G}(n,p)$ as nodes get infected
  - If v gets infected, we draw $\deg(v) \sim Po(c)$.
  - Check A_{vj} when the infection spreads.
- Look at $X :=$ # of infected half-edges, "active edges"

(actually "in expectation")

$\to$ Stochastic process, $X \to X-1$ at rate $(\gamma+\beta)X$. $\overset{\gamma\quad\beta}{\bullet\!\!\longrightarrow}$ [cheating a little in γ]

Also, at rate βX, w.p. $\frac{S(t)}{n}$ $\begin{cases} S \to S-1, \\ X-1 \to X-1+Po(c). \end{cases}$

Thm) As $n \to \infty$, $\frac{S(t)}{n} \xrightarrow{P} s(t)$, $\frac{X(t)}{n} \xrightarrow{P} x(t)$, $\frac{I(t)}{n} \xrightarrow{P} i(t)$ and obey

$$\frac{ds(t)}{dt} = -\beta x s, \quad \frac{dx(t)}{dt} = -(\beta+\gamma)x - c\frac{ds(t)}{dt}, \quad \frac{di(t)}{dt} = -\gamma i - \frac{ds}{dt}.$$

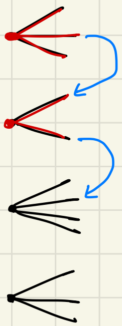
Def) Force of the infection: $-\frac{1}{s}\frac{ds}{dt} = \beta x$ (how strong infection removes an edge)

$s = s(0)e^{-\beta\int x\,dt} \Rightarrow s(\infty) = e^{-R_0(1-s(\infty))}$ [use $\frac{dx(t)}{dt} = -(\beta+\gamma)\frac{d}{dt}\int x\,dt - c\frac{ds}{dt}$]

→ extinction prob. of $Po(R_0)$, i.e. the backward process

SIR on $CM(\underset{\sim}{d})$: Consider an alternative matching algorithm,

"Match when an infection clock clicks."



- Start with $V_I \subseteq [n]$ infected nodes. $X := \sum_{v \in V_I} d_v$ active half-edges.
  - Start β clocks on all active half-edges.
  - Start γ clocks on all $v \in V_I$.
  - When β clock clicks, match the half-edge, and if it is susceptible, infect the other endpoint.

$X(t) :=$ # active half-edges, $S(t) :=$ # susc. nodes, $U(t) :=$ # unmatched half-edges

Rate of change: $U: -2\beta X$, $S_d = -\beta X \frac{dS_d}{U}$, $X: -(\beta+\gamma)X - \beta X \frac{X}{U} + \beta X \sum_d \frac{dS_d}{U}(d-1)$

where $S_d(t) :=$ # susc. nodes of degree d.

Diff Eq.: $\frac{1}{U}\frac{dU(t)}{dt} = -2\beta\frac{X}{U}$, $\frac{1}{S_d}\frac{dS_d(t)}{dt} = -\beta d\frac{X}{U}$, $\frac{1}{X}\frac{dX(t)}{dt} = -(\beta+\gamma) - \beta\frac{X}{U} + \beta\sum_d \frac{d(d-1)S_d}{U}$

$\rightarrow U(t) = \overbrace{n\bar{d}}^{U(0)=\bar{d}n} e^{-2\beta\int_0^t \frac{X(s)}{U(s)}ds}$, $S_d = S_d(0)e^{-\beta d\int\frac{X}{U}} = p_d \cdot n e^{-\beta d\int\frac{X}{U}}$. Let $\theta := e^{-\beta\int\frac{X}{U}}$.

$\rightarrow \underline{S_d = np_d\theta^d,\ U = n\bar{d}\theta^2}$.

$$\frac{\dot{X}}{X} = -(\beta+\gamma) - \beta\frac{X}{U} + \beta\sum_d d(d-1)\cdot\overbrace{\theta^{(d-2)}\cdot\frac{p_d}{\bar{d}}}^{S_d/U}.$$

But the last term is also $\beta\frac{\partial G^*(\theta)}{\partial\theta}$ where $G^*(\theta) = \sum_k p_k^*\theta^k = \sum_k \frac{(k+1)p_{k+1}}{\bar{d}}\theta^k = \sum_d \frac{d\,p_d}{\bar{d}}\theta^{(d-1)}$

$$\rightarrow \underline{\frac{\dot{X}}{X} = -(\beta+\gamma) - \beta\frac{X}{U} + \beta\frac{\partial G^*(\theta)}{\partial\theta}}.$$

Diff Eq. for θ: $\frac{d\theta}{dt} = -\beta\frac{X}{U}\cdot\theta =: -\beta\psi$.

$$\frac{d\psi}{dt} = \psi\left[\frac{\dot\theta}{\theta}+\frac{\dot X}{X}+\frac{\dot U}{U}\right] = \psi\left[\frac{\dot X}{X}-\beta\frac{X}{U}+2\beta\frac{X}{U}\right] = \psi\left[\frac{\dot X}{X}+\beta\frac{X}{U}\right]$$

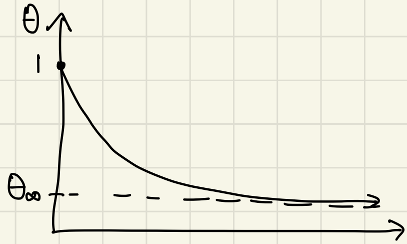
$$= -(\beta+\gamma)\psi + \beta\psi\frac{\partial}{\partial\theta}G^*(\theta) = -(\beta+\gamma)\psi - \frac{d\theta}{dt}\frac{\partial G^*(\theta)}{\partial\theta} = -(\beta+\gamma)\psi - \frac{dG^*(t)}{dt}$$

$$\Rightarrow \frac{d\psi}{dt} = \frac{\beta+\gamma}{\beta}\frac{d\theta}{dt} - \frac{d}{dt}G^*(\theta)$$

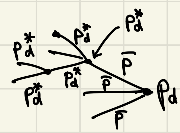
$\Rightarrow \psi = \left(\frac{\beta+\gamma}{\beta}\right)(\theta-1) - (G^*(\theta)-1)$ [using boundary condition $\psi(0)=0$]

$\frac{1}{\beta+\gamma}\frac{d\theta}{dt} = F(\theta) = (1-\theta) - \frac{\beta}{\beta+\gamma}(1-G^*(\theta))$. $\frac{\beta}{\beta+\gamma} = E[1-e^{-\beta T}] = \bar{p}$.

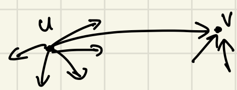
$\Rightarrow S = \sum_d p_d\theta^d = G(\theta)$, $\frac{di}{dt} = -\gamma i - \frac{ds}{dt}$, $\frac{dr}{dt} = -\gamma i$



$1-\theta_\infty = \bar{p}(1-G^*(\theta_\infty))$, $R_\infty = 1-S_\infty = 1-G(\theta_\infty)$.



Random Walks, Markov Chains, PageRank

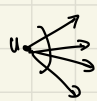


Motivation: How do we rank webpages?

Problem: 1 voter, many votes. ($d_u^{out} > 1$)

Solution: weigh by out-degree, use scaled adjacency matrix

↳ Matrix $P_{uv} := \frac{1}{d_u^{out}}A_{uv} \rightarrow \sum_v P_{uv} = 1$



→ Importance of v is $\pi_v = \sum_u \pi_u P_{uv}$ (important to important neighbors)

Def) Random Walk: Given a digraph $G(V,E)$ with A_{uv}, a random sequence $X_0, X_1, \dots$ where conditioned on X_{t-1}, X_t is a uniform out-neighbor of X_{t-1}.

↳ $\Pr[X_t = v \mid X_{t-1} = u] = P_{uv} = \frac{1}{d_u^{out}} A_{uv}$.

→ We shall study the distribution of X_t as $t \to \infty$.

Markov Chains: "State space" $V := \{1, 2, \dots, n\}$. Set of probability distributions π on V, $\Delta := \{\pi : V \to \mathbb{R}^+ \mid \sum_i \pi(i) = 1\}$. Transition matrix $M \in \mathbb{R}_+^{n\times n}$, $\sum_v M_{uv} = 1$.

↳ Homogeneous MC with TM M is a sequence $X_0, X_1, X_2, \dots$ of RVs s.t. $\Pr[X_t = u_t \mid X_s = u_s\ \forall s \in [0, t-1]] = \Pr[X_t = u_t \mid X_{t-1} = u_{t-1}] = M_{u_{t-1}u_t}$ indep. of t.

Let μ_t be the distribution of X_t.

Question: Given μ_0, what is μ_t? Does it converge?

Observe $\mu_t(v) = \sum_u \Pr[X_{t-1} = u] \Pr[X_t = v \mid X_{t-1} = u] = \sum_u \mu_{t-1}(u) M_{uv} = (\mu_{t-1} M)_v$.

$\dots = (\mu_0 M^t)_v \to \underline{\mu_t = \mu_0 M^t}$.

Def) Stationary Distribution: $\pi \in \Delta$ s.t. $\pi = \pi M$.

Questions: Does π exist? Is it unique? Does μ_t converge to π?

<u>Existence</u>: $\sum_j M_{ij} = 1$. $M_{ij} \cdot \vec{1} = \vec{1}$. → $\lambda = 1$. ⇒ $\exists u \in \mathbb{R}^n$ s.t. $uM = u$.

Normalizing u s.t. $\sum_i |U_i| = 1$, where $|U_i| = |\sum_j u_j M_{ji}| \le \sum_j |u_j| M_{ji}$.

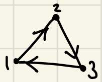
Assume $\exists i$ s.t. $|U_i| < \sum_j |u_j| M_{ji}$, $\sum_i |U_i| < \sum_{ij} |u_j| M_{ji} = \sum_j |u_j|$. ※

$\Rightarrow \forall i,\ |U_i| = \sum_i |u_j| M_{ij}$. Define $\pi_i := U_i$, then $\pi = \pi M$, $\sum_i \pi_i = 1$. // (✓)

The other two questions are not true <u>in general</u>.

ex1) [Uniqueness] RW on $\triangle \ \triangle$, $\pi = (\frac{1}{6}, \frac{1}{6}, \dots, \frac{1}{6})$. $P = \frac{1}{2}\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & & & \\ 1 & 0 & 1 & & & \\ 1 & 1 & 0 & & & \\ \hline & & & 0 & 1 & 1 \\ & & & 1 & 0 & 1 \\ & & & 1 & 1 & 0 \end{array}\right)$.

$\pi = \pi P$ is evident, but $\pi_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0)$ also gives $\pi_1 = \pi P$. (X)

ex2) [Convergence] RW on  $X_0 = 1$. $X_t = (t-1) \bmod 3$. (X)

Let D be a digraph, C be a SCC in D.

Def) Period of C: GCD(cycle length in C)

Def) Aperiodic D: Period of C $\forall$ SCC $C \subseteq D$.

Def) Transition graph D_M: given M, the graph with $uv \in E \Leftrightarrow M_{uv} > 0$.

$\rightarrow$ M is <u>irreducible</u>/<u>aperiodic</u> iff D_M is <u>SC</u>/<u>aperiodic</u>.

Thm) ① If M is irreducible, $\exists!$ stationary distribution $\pi \in \Delta$ s.t. $\pi = \pi M$.

② $\pi_i > 0$ $\forall i$ if M is irreducible. ③ If M is irreducible & aperiodic, $\forall \mu_0 \in \Delta$,

$\mu_0 M^t \longrightarrow \pi$ as $t \to \infty$.

Lemma: M is irreducible $\Leftrightarrow \forall ij, \exists t\ (M^t)_{ij} > 0$. [Proof in HW]

Remark: We can break periodicity by making M <u>lazy</u>, $M \mapsto M' := \alpha I + (1+\alpha) M$.

Then M is irred $\Leftrightarrow$ M' is irred., and $\pi M = \pi \Leftrightarrow \pi M' = \pi$.

ex) G is undirected, $M :=$ TM of RW on G; $M_{ij} = \frac{1}{d_i} A_{ij}$.

Claim: $\pi_i = \frac{1}{Z} d_i$, where $Z := 2|E|$ s.t. $\sum_i \pi_i = \frac{\sum_i d_i}{Z} = 1$.

$\hookrightarrow (\pi M)_i = \sum_j \pi_j M_{ji} = \frac{1}{Z} \sum_j A_{ij} = \frac{1}{Z} d_i = \pi_i$. //

PageRank: SD π of RW? $M_{ij} := \frac{1}{d_i^{out}} M_{ij}$. But d_i^{out} may be 0? $\rightarrow$ add self-loop.

But we can have sinks! $\rightarrow$ "Teleportation", we just randomly restart, $\pi_{sj} = \frac{1}{n}$.

$\hookrightarrow$ A robust model: $\pi_j := \alpha \frac{1}{n} + (1-\alpha) \sum_i \pi_i P_{ij} \rightarrow M_{ij} := \alpha \frac{1}{n} + (1-\alpha) P_{ij}$.

How to calculate π? 1) Solve equation $\pi(v) = \alpha \frac{1}{n} + (1-\alpha) \sum_u \pi(u) P_{uv}$, or

2) Run the MC until we find the fixed point π. Let's do 1) for now.

$\vec{1} = (1, 1, \ldots, 1)$. $\pi = \frac{\alpha}{n} \vec{1} + (1-\alpha) \pi P \rightarrow \pi(I - (1-\alpha)P) = \frac{\alpha}{n} \vec{1}$. Take inverse,

$\pi = \frac{\alpha}{n} \vec{1} (I - (1-\alpha)P)^{-1} = \frac{\alpha}{n} \vec{1} \frac{1}{I - (1-\alpha)P} = \frac{\alpha}{n} \vec{1} \sum_{k=0}^{\infty} (1-\alpha)^k P^k$ $\left[\frac{1}{1-x} = \sum_k x^k \text{ for } P \right]$

$\rightarrow \pi(v) = \frac{\alpha}{n} \sum_{k=0}^{\infty} (1-\alpha)^k \sum_u (P^k)_{uv}$.

Convergence speed: $\| \pi - \frac{\alpha}{n} \sum_{k=0}^{K-1} (1-\alpha)^k \sum_u (P^k)_{uv} \|_1 = \| \sum_{k \geq K}^{\infty} \frac{\alpha}{n} (1-\alpha)^k \sum_u (P^k)_{uv} \|_1$

$= \frac{\alpha}{n} \sum_{k \geq K} (1-\alpha)^k \overbrace{\sum_{uv} (P^k)_{uv}}^{n} = \alpha \sum_{k \geq K} (1-\alpha)^k = \alpha \frac{(1-\alpha)^K}{1-(1-\alpha)} = (1-\alpha)^K \leq e^{-\alpha K}$.

For 2), imagine we run the PageRank for t steps. $\rightarrow \mu_{t,x}(y) = (M^t)_{xv}$.

Q: How far is $\mu_{t,x}$ from π? $\rightarrow$ Use TV, $\|\mu - \nu\|_{TV} := \sup_{S \subseteq V} \{ \mu(S) - \nu(S) \}$.

Idea: After a restart, the sequence X_t "forgets" its history → converge to π!

Details: Predraw $Z_1, \ldots, Z_t \in V$, the π chain, and $Y_1, \ldots, Y_t \sim \mathrm{Ber}(\alpha)$.

Couple the chain s.t. if $Y_i = 1$, take $X_i = Z_i$, while $X_1, \ldots, X_{i-1}$ evolves from x.

$\Pr[X_t \in S] = \mathop{E}_{Y,Z}[\Pr[X_t \in S \mid Y, Z]]$, $\Pr[\tilde{X}_t \in S] = \mathop{E}_{Y,Z}[\Pr[\tilde{X}_t \in S \mid Y, Z]]$.

$\to d_{TV}(X_t, \tilde{X}_t) = \sup_{S \subseteq V}\{E[\Pr[X_t \in S] - \Pr[\tilde{X}_t \in S] \mid Y, Z]\} = (1-\alpha)^t \leq e^{-\alpha t}$.

Thm) M is irred. & aperiodic, $\|\mu_{t,x} - \pi\|_{TV} \longrightarrow 0$ as $t \to \infty$. (But how fast?)

$\to$ Let $d(t) := \max_x \{\|\mu_{t,x} - \pi\|_{TV}\}$, $\tilde{d}(t) := \max_{x,y}\{\|\mu_{t,x} - \mu_{t,y}\|_{TV}\}$.

Lemma: $d(t) \leq \tilde{d}(t) \leq 2d(t)$. Upper bound is clear by TI. [Lower in HW]

Def) Mixing Time: $\tau = \min_t \{\tilde{d}(t) \leq \frac{1}{e}\}$. Why $\frac{1}{e}$?

Lemma: $\tilde{d}(t+s) \leq \tilde{d}(t) \times \tilde{d}(s)$. [TV contraction] [Proof in HW]

Corollary: $d(tk) \leq \tilde{d}(tk) \leq \tilde{d}(t)^k \to d(\tau k) \leq \tilde{d}(\tau)^k \leq e^{-k} \to \underline{d(t) \leq e^{-\lfloor t/\tau \rfloor}}$.

Spectral Method

Consider RW on connected $G(V,E)$. $\to P_{ij} = \frac{A_{ij}}{d_i}$, $\pi_i = \frac{d_i}{2|E|}$.

Eigenvalues: $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$. Lazy Chain $M' = \frac{1}{2}(M + I)$.

Goal: τ is large $\Leftrightarrow$ G has a "bottleneck" $\Leftrightarrow$ spectral gap $(1-\lambda_2)$ is small.

Why is $|\lambda_i| \leq 1$? Look at e.v. of $\tilde{P} := D^{\frac{1}{2}} P D^{-\frac{1}{2}}$, it is same as P. ($D := \mathrm{diag}(d_i)$)

→ $\tilde{P}_{ij} = \frac{1}{\sqrt{d_i}} A_{ij} \frac{1}{\sqrt{d_j}} = P_{ji}$. By Spectral Thm, $\tilde{P} = \sum_i \lambda_i \varphi_i \varphi_i^T$ for ONB $\varphi_1, \dots, \varphi_n$.

→ If $Pv = \lambda v$, $|\lambda||v_i| = |\sum_j P_{ij} v_j| \le \max|v_j| \overbrace{\sum_j P_{ij}}^{1} \Rightarrow |\lambda| \max|v_i| \le \max|v_j| \Rightarrow |\lambda| \le 1$. //

In $G(n, c/n)$, $E[\#$ of edges between $S, S^c] = \frac{c}{n}|S||S^c| \ge \frac{c}{2}\min\{|S|,|S^c|\}$ 

Def) Expansion: $\alpha := \min_{\emptyset \subsetneq S \subsetneq V} \alpha(S)$ where $\alpha(S) := \frac{1}{\min\{|S|,|S^c|\}} \sum_{i\in S, j\notin S} A_{ij}$.

Lemma: i) If $\alpha > 0$, G is connected. ii) $\forall x,y \in V$, $\mathrm{dist}(x,y) \le 2\log_{\frac{\alpha}{d_{max}}} n$.

Def) Conductance: $\Phi := \min_{\emptyset \subsetneq S \subsetneq V} \Phi(S)$, $\Phi(S) := \frac{1}{\pi(S)\pi(S^c)} \sum_{i\in S, j\notin S} Q_{ij}$ where

$Q_{ij} := \pi_i P_{ij}$, the flow across edge ij. $\quad \hookrightarrow = \frac{1}{2|E|\pi(S)\pi(S^c)} \sum_{i,j} A_{ij}$

$\ge \frac{1}{2}\min\{\pi(S), \pi(S^c)\}$

→ $Q_{ij} = \frac{d_i}{2|E|}\cdot\frac{1}{d_i} A_{ij} = \frac{1}{2|E|}\cdot A_{ij}$, $\pi(S) = \sum_{i\in S}\frac{d_i}{2|E|}$, $\min\{\pi(S), \pi(S)\} \ge \pi(S)\cdot\pi(S^c)$

Lemma: $\frac{\alpha}{d_{max}} \le \Phi \le 2\frac{\alpha}{d_{min}}$ [We want good amount of edges <u>out</u> of S]

Recap: $P := \frac{1}{D} A$, $D := \mathrm{diag}(d_i)$, $\pi_i = \frac{d_i}{2|E|} \Rightarrow \pi P = \pi$. P has eigenvalues $1 = \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge -1$. $\Phi(S) = \frac{\Pr[X_{t-1}\in S, X_t \in S^c]}{\Pr[X_t\in S]\Pr[X_t\in S^c]} = \frac{\Pr[X_t \in S^c | X_{t-1}\in S] \swarrow \text{condition}}{\Pr[X_t\in S^c] \leftarrow \text{uncondition}}$

"Theorem": T is large $\Leftrightarrow (1-\lambda_2)$ is small $\Leftrightarrow \Phi$ is small.

Thm 1) Cheeger's Inequality: $\frac{\Phi^2}{8} \le (1-\lambda_2) \le \Phi$.

Proof Sketch of UB: $\tilde{P} := D^{-1/2} P D^{-1/2} = D^{-1/2} A D^{-1/2} = \sum_i \lambda_i\varphi_i\varphi_i^T$, $\{\varphi_i\}$ is ONB.

Lemma 1: i) $\varphi_1(v) = \sqrt{\frac{d_v}{2|E|}}$ ($\swarrow \pi(v)$) with $\lambda_1 = 1$. ii) If $f(u) := \sqrt{d_u}\left(\frac{1}{\pi(S)}\mathbb{1}\{u\in S\} - \frac{1}{\pi(S^c)}\mathbb{1}\{u\in S^c\}\right)$,

then $f\perp\varphi_1$ and $(1-\Phi(S)) = \frac{f^T\tilde{P} f}{\|f\|_2} \Leftrightarrow \Phi(S) = \frac{f^T(I-\tilde{P})f}{f^T T}$.

Corollary of i): $\tilde{P}\varphi_1 = \varphi_1$, $\|\varphi_1\|_2 = 1$. Proof: $\sum_v \varphi_1^2(v) = \sum_v \pi(v) = 1$.

Proof: $\sum_y \tilde{P}_{xy}\varphi_1(y) = \frac{1}{\sqrt{d_x}}\sum_y A_{xy}\frac{1}{\sqrt{d_y}}\sqrt{\frac{d_y}{2|E|}} = \frac{\sqrt{d_x}}{\sqrt{2|E|}} = \varphi_1(x)$.

Lemma 2: Let the Laplacian $L := D - A$. Then $f^T L f = \frac{1}{2}\sum_{i,j} A_{ij}(f(i) - f(j))^2$.

Why "Laplacian"? On $\mathbb{R}$, $f''(x) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - 2f(x) + f(x-\varepsilon)}{\varepsilon^2}$.

Discrete analogue on $\mathbb{Z}$: $\nabla^2 f(x) = [f(x-1) - 2f(x) + f(x+1)]$.

Proof: RHS $= \frac{1}{2}\sum_i f_i^2 d_i + \frac{1}{2}\sum_j f_j^2 d_j - \sum_{ij} f_i f_j A_{ij} = \sum_i f_i D_i f_i - \sum_{ij} A_{ij} f_i f_j =$ LHS.

↳ This gives proof to Lemma 1 (ii). Remark: $I - \tilde{P} = D^{-1/2}(D-A)D^{-1/2}$ [Normalized L]

Remark 2: Morally, this says $\min_S \Phi(S) = 1 - \lambda_2$.

Proof of Lemma 1 (ii): Orthogonality: $\sum_i f_i \varphi_1(i) = \frac{1\sqrt{2|E|}}{\cancel{\sqrt{2|E|}}}\sum_i \frac{d_i}{2|E|}\left(\pi(S^c)\mathbb{1}\{i\in S\} + \pi(S)\cdot\mathbb{1}\{i \in S^c\}\right)$

$= \sqrt{2|E|}\sum_i \pi(i)[\pi(S^c)\mathbb{1}\{i\in S\} - \pi(S)\mathbb{1}\{i\in S^c\}] = \sqrt{2|E|}\sum_i \pi(S^c)\pi(S) - \pi(S)\pi(S^c) = 0$.

$f^T(I - \tilde{P})f = \tilde{f}^T(D-A)\tilde{f}$ (where $\tilde{f}_i := f_i \cdot \frac{1}{\sqrt{d_i}}$) $= \frac{1}{2}\sum_{ij}|\tilde{f}_i - \tilde{f}_j|^2 A_{ij}$.

$\tilde{f}_i - \tilde{f}_j = [\pi(S^c)(\mathbb{1}\{i\in S\} - \mathbb{1}\{j \in S\}) - \pi(S)(\mathbb{1}\{i\in S^c\} - \mathbb{1}\{j\in S^c\})]$

$$= \begin{cases} 0 \text{ if } i,j\in S \text{ or } i,j\in S^c \\ \pi(S)+\pi(S^c) = 1 \text{ if } i\in S, j\in S^c \text{ or } i\in S^c, j \in S^c. \end{cases}$$

$\Rightarrow \tilde{f}^T(I-\tilde{P})\tilde{f} = \sum_{i\in S, j\in S^c} A_{ij}$. Recall that $\Phi(S) = \frac{\sum_{i\in S, j\in S^c} A_{ij}}{2|E|\pi(S)\pi(S^c)}$. is this $\tilde{f}^T\tilde{f}$?

$\tilde{f}^T\tilde{f} = \sum_i d_i[\pi(S^c)^2\cdot\mathbb{1}\{i\in S\} + \pi(S)^2\cdot\mathbb{1}\{i\in S^c\}] = 2|E|\sum_i \pi(i)[\sim]$

$= 2|E|(\pi(S^c)\pi(S) + \pi(S)\pi(S^c)) = 2|E|\pi(S)\pi(S^c)$. ✓

Now we have the tools to prove the UB of Cheeger's Inequality.

Proof of Thm1 [Upper Bound]: $\tilde{P} = D^{-1/2} A D^{-1/2}$ has same λ as $P = D^{-1/2} \tilde{P} D^{1/2}$.

By Spectral Thm, $\tilde{P}_{xy} = \sum_{i\geq 1} \lambda_i \varphi_i(x)\varphi_i(y)$ where $\tilde{P}\varphi_i = \lambda_i \varphi_i$, $\{\varphi_i\}$ is ONB.

Set $g \leftarrow \frac{f}{\|f\|} \perp \varphi_1$. $1 - \Phi(S) = g^T \tilde{P} g = \sum_{i\geq 2} \lambda_i (g\varphi_i^T)^2 \leq \lambda_2 \sum_{i\geq 2} (g\varphi_i^T)^2$

$= \lambda_2 \underbrace{\sum_{i=1}^{n} (g\varphi_i^T)^2 = \lambda_2}_{\text{ONB}}. \Rightarrow \underbrace{\Phi = \min_S \Phi(S) \geq 1 - \lambda_2}$. //

Thm 2) Mixing Time: $\mu_{x,t}(v) = M^t_{x,v}$. $\tilde{d}(t) = \max_{x,y} \|\mu_{t,x} - \mu_{t,y}\|_1$, $T = \min\{t \mid \tilde{d}(t) \leq \frac{1}{e}\}$.

$d(t) = \min_x \frac{1}{2}\|\mu_{t,x} - \pi\|_1 \leq \tilde{d}(t) \leq e^{-\lfloor t/T \rfloor}$. For some lazy chain $M = \frac{1}{2}(I + P)$,

$\tilde{\lambda}_i := \frac{1+\lambda_i}{2}$ are evs of M. Then $\tilde{\lambda}_2^t \leq \tilde{d}(t) \leq \frac{\tilde{\lambda}_2^t}{\sqrt{\pi_{min}}}$. $[\tilde{\lambda}_2^t \leq e^{-(1-\tilde{\lambda}_2)t}]$

Proof Sketch of UB: $\mu_{x,t}(v) = M^t_{x,v}$. $\tilde{M} := D^{1/2} M D^{-1/2} = \sum_i \tilde{\lambda}_i \varphi_i \varphi_i^T$. [LB in HW]

Then $M^t = (D^{-1/2} \tilde{M} D^{1/2})^t = D^{-1/2} \tilde{M}^t D^{1/2}$. $M^t_{xv} = \sum_{i\geq 1} \tilde{\lambda}_i^t \varphi_i(x)\varphi_i(v) \sqrt{\frac{d_v}{d_x}}$

$= \pi(v) + \sum_{i\geq 2} \tilde{\lambda}_i^t \varphi_i(x)\varphi_i(v)\sqrt{\frac{d_v}{d_x}}$. $[\tilde{\lambda}_1^t \varphi_1(x)\varphi_1(v)\sqrt{\frac{d_v}{d_x}} = \frac{1}{2|E|} d_v = \pi(v)]$

$\rightarrow \|\mu_{x,t} - \pi\|_1 \leq \tilde{\lambda}_2^t \frac{1}{\min\sqrt{d_x}} \sum_i |\varphi_i(x)| \sum_v |\varphi_i(v)| \sqrt{d_v} \leq \tilde{\lambda}_2^t \frac{1}{\min\sqrt{d_x}} \underbrace{\left[\sum_i \varphi_i(x)^2\right]^{1/2}}_{\substack{1 \\ \text{(ONB)}}} \underbrace{\left[\sum_i \left(\sum_v \sim\right)^2\right]^{1/2}}_{\substack{\sqrt{2|E|} \\ \text{(omitted)}}}$

$= \tilde{\lambda}_2^t \sqrt{\frac{2|E|}{\min d_x}} = \frac{\tilde{\lambda}_2^t}{\sqrt{\pi(x)}}$. //

Graph Partitioning

Goal: Given a (weighted) graph on V, partition $V \rightarrow (S, S^c)$ "well-seperated".

Def) Volume: $vol(S) := \sum_{x\in S} d_x$, then $\pi(S) = \frac{\sum_{x\in S} d_x}{2|E|} = \frac{vol(S)}{vol(V)}$.

Def) Cut: $cut(S, S^c) := \sum_{x\in S, y\in S^c} A_{xy}$. $\rightarrow \Phi(S) = \frac{cut(S,S^c)\, vol(V)}{vol(S)\, vol(S^c)}$.

Let $\bar{d}(S) = \bar{d}_{int}(S) + \bar{d}_{ext}(S)$ where $\bar{d}_{int}(S) := \frac{1}{vol(V)} \sum_{\substack{x\in S \\ y \in S}} A_{xy}$, $\bar{d}_{ext}(S) := \frac{1}{vol(V)} \sum_{\substack{x\in S \\ y \in S^c}} A_{xy}$.

$\to$ $\bar{d}_{ext}(S)$ is also expansion$(S) = \alpha(S)$ (if $|S| \le |S^c|$).

What is a good partition?

① minimize $cut(S, S^c)$ s.t. $\varepsilon \le \frac{|S|}{|V|} \le \frac{1}{2}$.

② minimize $\bar{d}_{ext}(S) = \alpha(S)$ s.t. $|S| \le |S^c|$.

③ minimize $\frac{\bar{d}_{ext}(S)}{\bar{d}_{int}(S)} \Leftrightarrow \min \frac{\bar{d}_{ext}(S)}{\bar{d}_{ext}(S) + \bar{d}_{int}(S)} = \frac{cut(S, S^c)}{vol(S)}$ s.t. $vol(S) \le vol(S^c)$

④ minimize $\Phi(S)$

ex) $V = V_1 \sqcup V_2$, $|V_1| = n_1 \le |V_2| = n_2$.

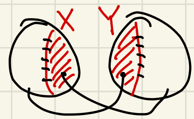
i) $K_{n_1} \sqcup K_{n_2}$. K_{n_1} K_{n_2} Set $S \leftarrow V_1$, then $cut(S, S^c) = 0 \to \Phi(S) = \emptyset$.

ii) K_n^{loop} ($A_{xy} = 1 \ \forall x, y$). $d_x = n \ \forall x$. $vol(S) = n|S|$, $cut(S, S^c) = |S||S^c|$.

$\to \Phi(S) = \frac{|S||S^c| n^2}{(n|S|)(n|S^c|)} = 1 \ \forall S \Rightarrow$ all partitions are equally bad.

Planted Clique in $G(V, p)$: $\underset{V_1}{K_{n_1}^{loop}} \overset{w.p.\ p}{\equiv} \underset{V_2}{K_{n_2}^{loop}}$, i.e. $P_{ij} = \begin{cases} 1 & \text{if } i,j \in V_1,\ i,j \in V_2 \\ 0 & \text{if } i \in V_1, j \in V_2, \text{ vice versa} \\ p & \text{o.w.} \end{cases}$

$\to E[d_x] = \begin{cases} d_1 = n_1 + n_2 p & \text{if } x \in V_1 \\ d_2 = n_2 + n_1 p & \text{if } x \in V_2 \end{cases}$. $E[cut(S, S^c)] = \sum_{\substack{x \in S \\ y \in S^c}} P_{xy}$.



Let $X := \frac{1}{n_1}|S \cap V_1|$, $Y := \frac{1}{n_2}|S \cap V_2|$. Assume concentration.

Replace d_x by $E[d_x]$, cut by $E[cut]$. Then $vol(S) = d_1 X n_1 + d_2 Y n_2$,

$vol(S^c) = d_1 \bar{X} n_1 + d_2 \bar{Y} n_2$ where $\bar{X}, \bar{Y} = (1-X), (1-Y)$.

$\mathrm{cut}(S,S^c) = n_1^2 X\bar{X} + n_2^2 Y\bar{Y} + p n_1 n_2 (X\bar{Y} + Y\bar{X})$.

$\rightarrow$ minimize $f(X,Y) = \dfrac{\mathrm{cut}(S,S^c)}{\mathrm{vol}(S)\,\mathrm{vol}(S^c)}$. $\rightarrow$ gives $S = V_1$ or V_2 as p is bounded $[0,1]$.

Spectral Clustering

Recall $(1-\lambda_2) \leq \min_S \Phi(S) = \Phi$. $\rightarrow$ 2^{nd} ev approximates minimal conductance.

Q: What about eigenvectors?

First ev $\varphi_1(u) = \sqrt{\dfrac{d_u}{2|E|}} > 0$. All others are orthogonal. $\rightarrow$ must have signs!

Try $S := \{v \in V \mid \varphi_2(v) < 0\}$, $S^c := \{v \in V \mid \varphi_2(v) \geq 0\}$.

Consider the weighted adj. matrix of a "clustered" graph:

$$A = \begin{array}{c} \\ n_1 \\ n_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \left[\begin{array}{c|c} a & \varepsilon \\ \hline \varepsilon & b \end{array}\right] \end{array}$$

$d_i = a n_1 + \varepsilon n_2$ if $i \in V_1$, $b n_2 + \varepsilon n_1$ if $i \in V_2$.

$P = \frac{1}{D} A$ has $\lambda_1 = 1$ with $\varphi_1 = \vec{1}$.

We want to solve $D^{-1} A \psi = \lambda \psi$. $(A\psi)_i = a \sum_{i \in V_1} \psi_i + \varepsilon \sum_{i \in V_2} \psi_i = D_1 \psi_i \rightarrow \psi_i = \underline{c_1}$ if $i \in V_1$.

$(A\psi)_i = b \sum_{i \in V_2} \psi_i + \varepsilon \sum_{i \in V_1} \psi_i = D_2 \psi_i \rightarrow \psi_i = \underline{c_2}$ if $i \in V_2$.

Observe that $\lambda_1 = 1 \Rightarrow c_1 = c_2 = c$, $\psi = [c, c, \ldots, c]^T$.

If $\varepsilon = 0$, $\lambda_2 = 1$, and $\psi = \alpha[1,1,\ldots,0,0\ldots]^T + \beta[0,0,\ldots,1,1,\ldots]^T$.

If $\varepsilon^2 < ab$, $\lambda_2 < 1$, and we get a sign cut $S = \{i \mid \psi_i < 0\}$. [Proof in HW]

Algorithm: Let Λ_2 be the second eigenvalue* with eigenvector φ_2, $\varphi_2^T D \varphi_1 = 0$

where $\varphi_1 = [1 1 \ldots 1]^T$. ($*$ of $(D-A)\psi = \Lambda D \psi$)

- order vertices s.t. $\psi_2(v_1) \leq \psi_2(v_2) \leq \dots$.
- output $S_k = \{v_1, v_2, \dots, v_k\}$ where (choose 1):
  - $k = \min_k \{\psi_2(v_k) = 0\}$ [sign cut]
  - $k = \lceil n/2 \rceil$ [median cut]
  - $k = \arg\min_k \{\Phi(S_k)\}$ [sweep cut]

Q1: Why does this work? Why $\psi_2^T D \psi_1 = 0$?

Remark: we get a witness $\Lambda_2 = 1 - \lambda_2 \leq \Phi \leq \Phi(S_k) \Rightarrow \text{error} \leq \Phi(S_k) - \Lambda_2$

Q2: What if $\Lambda_2 = 0$?

Thm 1) $\Lambda_2 = \min_{\psi} \left\{ \frac{\psi^T(D-A)\psi}{\psi^T D \psi} \right\}$ where $\psi^T D \psi_1 = 0$. Any minimizer is a generalized eigenvector with $(D-A)\psi = \Lambda_2 D\psi$.

Proof: Recall $\tilde{P} = D^{-1/2} A D^{-1/2} = \sum_i \lambda_i \varphi_i \varphi_i^T$ for ONB $\{\varphi_i\}$. Then let $\tilde{L} = I - \tilde{P} = D^{-1/2}(D-A)D^{-1/2} = \sum_i \Lambda_i \varphi_i \varphi_i^T$, the normalized Laplacian.

$\Lambda_2 = \varphi_2^T \tilde{L} \varphi_2 = \min_{\varphi \perp \varphi_1} \left\{ \frac{\varphi^T \tilde{L} \varphi}{\varphi^T \varphi} \right\}$ and any minimizer is an eigenvector with $\tilde{L}\varphi = \Lambda_2 \varphi$. [Detail in HW]

$\varphi^T \tilde{L} \varphi = \varphi^T D^{-1/2}(D-A)D^{-1/2}\varphi$. $\underline{\text{Let } \psi := D^{-1/2}\varphi} \Rightarrow \psi^T(D-A)\psi$.

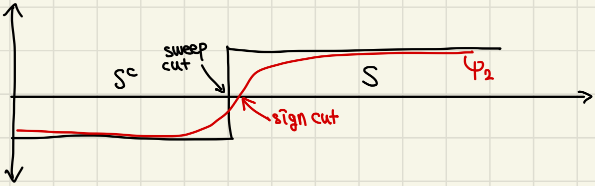
$\varphi^T \varphi = \psi^T D \psi$. $\varphi_1(v) = \sqrt{\frac{d_v}{2|E|}}$, so $\varphi \perp \varphi_1 \Rightarrow \sum_v d_v^{1/2} \varphi(v) = 0$.

$\psi^T D \psi_1 = \sum_v \psi(v) d_v = \sum_v \varphi(v) d_v^{1/2} = 0$. So two are equivalent. //

Thm 2) $\Phi = \min_{\varphi} \left\{ \frac{\varphi^T(D-A)\varphi}{\varphi^T D \varphi} \right\}$ where $\varphi^T D \varphi_1 = 0$ and φ has two levels. (values)

Then any minimizer is of the form $c\cdot(\pi(S^c)\mathbb{1}\{x\in S\} - \pi(S)\mathbb{1}\{x \in S^c\})$.

↳ Interpretation: Λ_2 is the continuous relaxation of the discrete conductance minimization problem. ← NP-Hard



$\varphi^T(D-A)\varphi = \frac{1}{2}\sum_{i,j}(\varphi_i - \varphi_j)^2 A_{ij}$

↳ What if $\Lambda_2 = 0$? [HW: If G has k components, $\dim\{\varphi \mid (D-A)\varphi = 0\} = k$]

→ $\Lambda_1 = \Lambda_2 = \cdots = \Lambda_k = 0$ and $\Lambda_{k+1} > 0$. In particular, $\Lambda_2 = 0 \Leftrightarrow$ $\Phi = 0 \Leftrightarrow$ perfect clustering, and φ is constant on all components.

Claim: $\Phi_S = \frac{\varphi_S^T(D-A)\varphi_S}{\varphi_S^T D \varphi_S}$ where $\varphi_S = \pi(S^c)\mathbb{1}_S - \pi(S)\mathbb{1}_{S^c}$.

Proof: Recall $\Phi_S = \frac{f_S^T(I - D^{-1/2} A D^{-1/2}) f_S}{\|f_S\|^2}$ where $f_S(v) = \sqrt{d_v}\,\varphi_S$. Expand f_S to get $\varphi_S = D^{-1/2}\varphi_2$ and every thing works out. //

Proof of Thm 2: $0 = \varphi_S^T D \varphi_1$. $\varphi = a\mathbb{1}_S + b\mathbb{1}_{S^c} \Rightarrow 0 = a\sum_{x\in S} d_x + b\sum_{x\in S^c} d_x$ $= 2|E|\left(a\sum_{x\in S}\pi(S) + b\sum_{x\in S^c}\pi(S^c)\right) = 2|E|(a\pi(S) + b\pi(S^c))$. Then it must be the case that $a = c\cdot\pi(S^c)$, $b = -c\cdot\pi(S)$. $\Rightarrow \varphi = c(\pi(S^c)\mathbb{1}_S - \pi(S)\mathbb{1}_{S^c})$.

By claim, minimizer φ of such form yields Φ. //

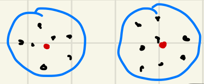
What do people do in practice?

k-means Spectral Clustering: calculate d lowest eigenvalues/vectors $\varphi_{1,\dots,d}$.

Embed the graph into $\mathbb{R}^d$ as follows: map $v \in V \mapsto \tilde{v} = (\varphi_1(v), \varphi_2(v), \dots, \varphi_d(v))$.

Now we have $\tilde{v}_1, \dots, \tilde{v}_n \in \mathbb{R}^d$ that we can geometrically cluster into k pieces.

Specifically, we can use k-means, $S = (S_1, \dots, S_k) = \underset{S}{\arg\min} \sum_{i=1}^{k} \sum_{j \in S_i} \|\vec{v}_j - \vec{m}_i\|_2^2$

where $\vec{m}_i$ is the center of S_i, $\frac{1}{|S_i|} \sum_{j \in S_i} \vec{v}_j$. 

Runtime: $O(n^{(kd+1)})$ in worst case. In practice, $O(\overset{\#\text{ iterations.}}{T} nkd)$ by local search.

Hopefully, we find some d where $\lambda_1 \simeq \lambda_2 \simeq \dots \simeq \lambda_d \ll \lambda_{d+1}$ and $d \simeq k$.

Information Cascades

Model the spread of information/innovation over social network

- Maximize the total # of people that get the information w/ seeding

History: Socio/economics [Shelling '78, Gronevetter '78], first math model [Domingos, Richardson '01] introduced influence maximization

[Kempe, Kleinberg, Tardos '03] TCS analysis

Generic Model: Digraph $G(V, E)$. Except for few seeds, everyone starts in an "inactive" state. In discrete timesteps $t \mapsto (t+1)$, some of inactive neighbors of active nodes get activated w.r.t. some rule. Active nodes stay active forever.

Linear Threshold Model: Digraph $G(V,E)$, weight matrix $B=(b_{uv})_{uv\in E}\geq 0$.

Spreading process: $t=-1$, all nodes iid uar choose threshold $\theta_v\in[0,1]$.

At $t=0$, $A_0\subseteq V$ is the initial seed.

When $(t-1)\mapsto t$, $A_t \leftarrow A_t \cup \{v\notin A_{t-1} \mid \sum_{w: wv\in E} b_{wv} \geq \theta_v\}$. Stop when $A_{t+1}=A_t$.

Independent Cascade Model [KKT '03]: Digraph $G(V,E)$, matrix $P=(\overset{\in[0,1]}{p_{uv}})_{uv\in E}$.

Spreading process: each newly active node v gets only one chance to spread its information to neighbors $u\in N(v)$ w.p. p_{uv}.

Def) Influence: For $A\subseteq V$, $I(A):=E[|A_\infty| \mid A_0=A]$.

Influence Maximization: For input $D=(V,E)$, either B or P, and integer k, find set $A^*\subseteq V$ of size $\leq k$ s.t. $I(A^*)=\max_{|A|\leq k}\{I(A)\}$.

↳ This is NP-Hard. Mappable to Vertex Cover, etc.

Greedy Algorithm: Given finite V ground set, function $f: 2^V \mapsto \mathbb{R}_+$ which is monotone ($f(S\cup\{v\})\geq f(S)$ $\forall S\subseteq V, v\in V$), construct a sequence A_t, $A_0=\emptyset$, $A_t=A_{t-1}\cup\{v_t\}$ s.t. $f(A_{t-1}\cup\{v_t\})$ is maximized.

Thm) If $f: 2^V\to\mathbb{R}_+$ is <u>monotone</u> and <u>submodular</u>, and A_t is greedily selected, then $f(A_k)\geq f(A^*)(1-(1-\frac{1}{k})^k)\geq f(A^*)(1-\frac{1}{e})$ where $f(A^*)$ is maximal.

Def) Submodularity: $f: 2^V \to \mathbb{R}_+$ s.t. $f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T)$ if $S \subseteq T \subseteq V$ and $v \in V$. [Diminishing returns]

Lemma: Both models are submodular.

Corollary: We can greedily select seeds for both models for $(1-\frac{1}{e})$-approximation.

Proof [for independent cascades]: When u gets activated at time t, it flips $X_{uv} \sim Ber(p_{uv})$ and activates v if $X_{uv} = 1$ $\forall v \in N(u)$. Preflip all coins.

$D(X) = (V, E(X)) \subseteq D$ where $xy \in E(X)$ iff $X_{xy} = 1$.

Then $I(A \mid X) = \bigcup_{x \in A} C_+(x)$. Conditioned on X, I is submodular since

$I(S \cup \{v\}) - I(S) = \left|\{w \in C_+(v), w \notin \bigcup_{y \in S} C_+(y)\}\right| \geq \left|\{ '' , w \notin \bigcup_{y \in T} C_+(y)\}\right|$

$= I(T \cup \{v\}) - I(T)$ conditioned on X for $T \supseteq S$.

Claim: If $f = \sum_i p_i f_i$ and f_i is monotone & submodular, so is f.

Corollary: $I(A) = \mathbb{E}_X[I(A) \mid X]$ is submodular.

Proof of Claim: $f(S \cup \{v\}) - f(S) = \sum_i p_i [f_i(S \cup \{v\}) - f_i(S)]$ ... use linearity. //

Claim: Let $|S^*| = k$ s.t. $f(S^*) = \max_{|S'| \leq k} \{f(S')\}$. If $|S| \leq k$, then

$\max_{x \notin S} \{f(S \cup \{x\}) - f(S)\} \geq \frac{1}{k}[f(S^*) - f(S)]$.

Proof: Let $v_1, \ldots, v_p \notin S$, $p \leq k$ s.t. $S \cup S^* = S \cup \{v_1, \ldots, v_p\}$. $f(S^*) \leq f(S^* \cup S)$

$= f(S) + [f(S \cup \{v_1\}) - f(S)] + [f(S \cup \{v_1, v_2\}) - f(S \cup \{v_1\})] + \cdots$

$$\le f(S) + \sum_{i=1}^{p} [f(S \cup \{v_i\}) - f(S)] \le f(S) + k \max_{x \in S} \{f(S \cup \{x\}) - f(S)\}. //$$

Proof of Thm: Let $A_0 \leftarrow \emptyset, A_1, \dots$ be the output of Greedy.

Then $A_k = A_{k-1} \cup \{v\}$ s.t. $f(A_{k-1} \cup \{v\}) - f(A_{k-1})$ is maximized.

$$\rightarrow f(A_k) \ge f(A_{k-1}) + \frac{1}{k}[f(A^*) - f(A_{k-1})]$$

$$= \frac{1}{k} f(A^*) + \left(1 - \frac{1}{k}\right) f(A_{k-1}) \quad \dots \text{ and unroll recurrence}$$

$$\ge \frac{1}{k} f(A^*) + \left(1 - \frac{1}{k}\right)\left[\frac{1}{k} f(A^*) + \left(1 - \frac{1}{k}\right) f(A_{k-2})\right] \quad \dots$$

$$\ge \frac{1}{k} f(A^*) \sum_{i=0}^{k-1} \left(1 - \frac{1}{k}\right)^k + \left(1 - \frac{1}{k}\right)^k f(A_0)$$

$$\ge \frac{1}{k} \sum_{i=0}^{k-1} \left(1 - \frac{1}{k}\right)^k f(A^*) = \frac{1}{k} \frac{1 - \left(1 - \frac{1}{k}\right)^k}{1 - \left(1 - \frac{1}{k}\right)} f(A^*) = \left(1 - \left(1 - \frac{1}{k}\right)^k\right) f(A^*). //$$