Page Rank and Random Walks

also Perron Frobenius Theorem

Early web search

- categorize pages manually into various classes (hierarchically) Yahoo! and others
- key word based + ad hoc methods Alta Vista etc.

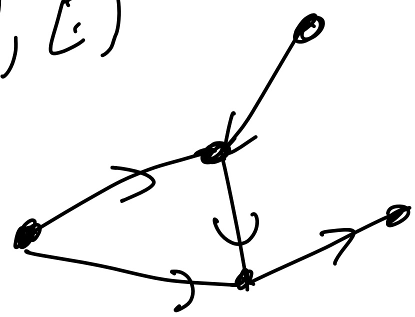
Issues: scalability and spamming and unreliability.

Google approach: use link information

Web graph.

each node is a webpage/url.
links create directed arc.

$G=(V,E)$



<u>Important</u>: graph is directed.

Links encode information that gives information on how <u>important</u> pages are. Create a <u>ranking</u> of webpages. "global". Use ranking + query words (later).

How to rank webpages in terms of importance?

① $\text{Score}_1(v) = \sum_{(u,v)\in E} 1 = \text{indeg}(v)$. Can spam easily.

weigh by out-degree.

② $\text{Score}_2(v) = \sum_{(u,v)\in E} \frac{1}{d^+(u)}$

(3) $$\text{score}_3(v) = \sum_{(u,v)\in E} \frac{\text{score}_3(u)}{d^+(u)}$$

"recursive" definition.

Main question: does $\text{score}_3(v)$ exist?

Can normalize score since scaling does not violate equation

$$\sum_v x(v) = 1 \qquad x(v) \geq 0$$

x is a probability distribution.

Theorem: For any strongly connected graph $\exists$ unique x s.t. s.t. $\bar{x}$ is a probability distribution and

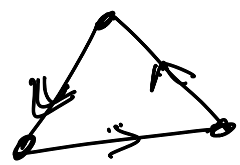
$$x(v) = \sum_{(u,v)\in E} \frac{x(u)}{d^+(u)}.$$

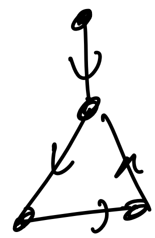
"Can be computed efficiently".

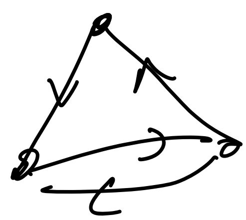
//

"stationary distribution" of a random walk on the web graph.

Ex:







Random Walks and Markov Chains

$G = (V, E)$

A stochastic process. $X_0, X_1, \ldots, X_n, \ldots$

$X_i \in V$

X_0 is initial "random" vertex

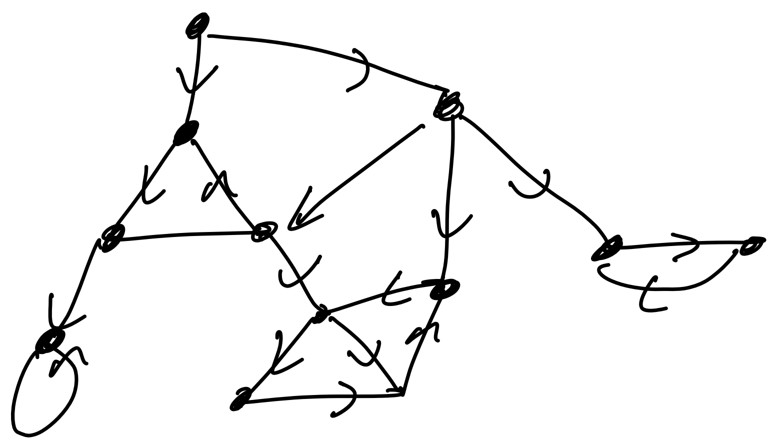
$\Pr[X_{t+1} = v \mid X_t = u, X_{t-1} \ldots X_0 =]$

$= \Pr[X_{t+1} = v \mid X_t = u] = \frac{1}{d^+(u)}$

process is Markovian

Q: - what happens in the long run?
- how does the process depend on the starting vertex/distribution?
- how does this depend on the graph?

Directed graphs:

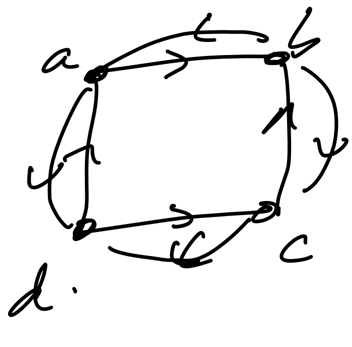


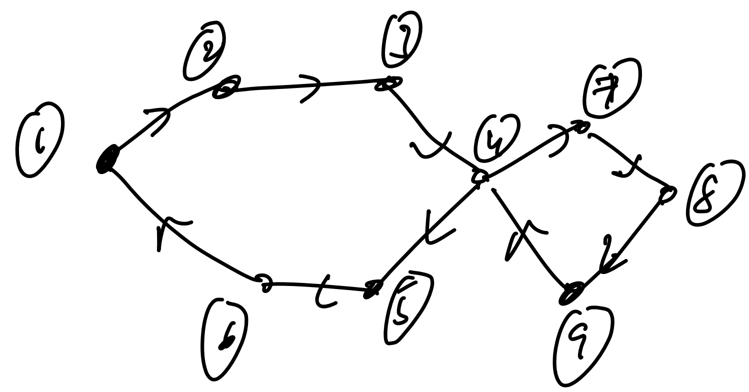
Strongly connected components

Walk will get stuck in a strongly connected sink component.

Thus it is useful to focus on strongly connected graphs.

Periodicity:





Defn: Period of state i or vertex v

$d(v) = \gcd(S(v))$ where

$S(v) = \{ n \mid$ there is a closed walk of length n from v to $v \}$.

v is aperiodic if $d(v) = 1$

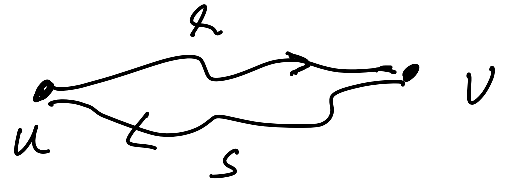
is periodic otherwise.

$d(v) > 1$.

Lemma: Let G be strongly connected

then $\forall\, v$ $d(v)$ is same.

Fix, u, v.



$r+s$ divisible by $d(u)$

Let t be any closed $v \to v$ walk length

$r+s+t$ is a closed $u \to u$ walk

$\Rightarrow$ $r+s+t$ is divisible by $d(u)$.

$\Rightarrow$ t is divisible by $d(u)$.

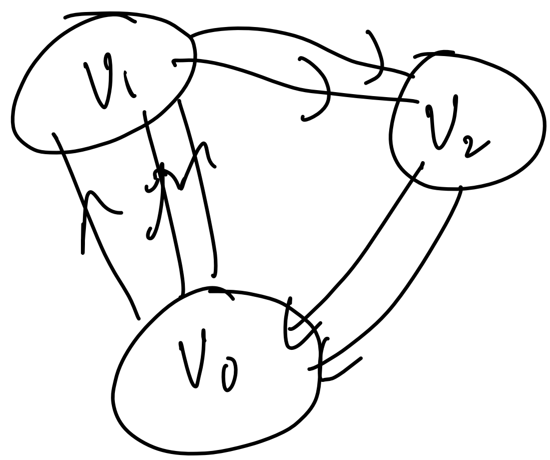
$\Rightarrow$ $d(u)$ divides $d(v)$ since

$d(v)$ is gcd of all t.

Similarly $d(v)$ divides $d(u)$.

$\Rightarrow$ $d(v) = d(u)$.

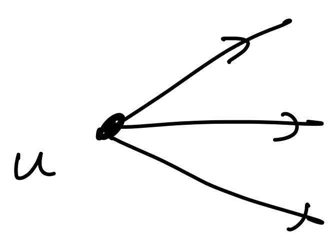
<u>Lemma:</u> Suppose G is strongly connected and $d(G)$ is period of G. Then V can be partitioned into $V_0, V_2, \ldots, V_{d-1}$ s.t. $\forall (u,v) \in E$ $u \in V_i \Rightarrow v \in V_{(i+1) \bmod d}$



$d = 3$

Edge-weights and Finite State Markov Chains and Matrices

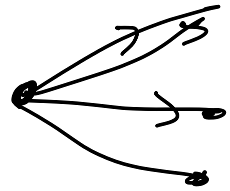
Random walk:

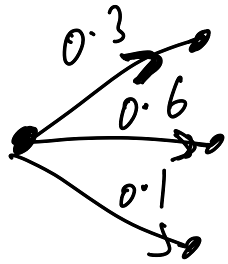


go to a uniform neighbor at random.

multigraph?

pick a uniform edge nbr at random





In general a probability distribution on out edges

Finite state Markov chain.

Represent as a weighted graph or as a probability transition matrix

P

$$
\begin{array}{c|ccc}
 & 1 & 2 & n \\
\hline
1 & & & \\
2 & & & \\
\vdots & & & \\
n & & &
\end{array}
$$

self loops allowed.

P_{ii} can be > 0.

P_{ij} = Probability of going to j from i.

$$\sum_{j=1}^{n} P_{ij} = 1 \quad \forall i \qquad P \geq 0$$

stochastic matrix.

Suppose $p(t)$ is the probability distribution over states at time t

. - -

P^k is the k-step transition matrix.

ie P^k_{ij} = probability of i going to j after k steps.

$P^k_{ij} = (P^k)_{ij}$.

If $\Pi = (\Pi_1, \Pi_2, \ldots, \Pi_n)$ is a probability row vector.

ΠP is the distribution after one step $(\Pi P) P$ is after 2 steps

ΠP^k after k steps.

<u>Defn:</u> A distribution π is a stationary distribution if
$\pi P = \pi$.

<u>Q:</u>

① Does $\pi = \pi P$ always have a probability vector solution

② Does $\pi = \pi P$ have a unique prob vector solution.

③ Do rows of P^n converge to a prob vector soln $\pi = \pi P$.?

<u>Answers:</u>

① Always!

② Uniqueness iff ~~G~~ G is strongly connected

③ Yes if G is aperiodic.

$\Rightarrow$ $\forall$ any starting distribution π_0. $\pi_0 P^n \to \pi$. stationary distribution

Back to Page Rank

Web graph. want to find a solution

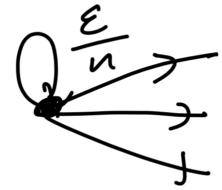
$$x(v) = \sum_{(u,v)\in E} \frac{x(u)}{d^+(u)}$$

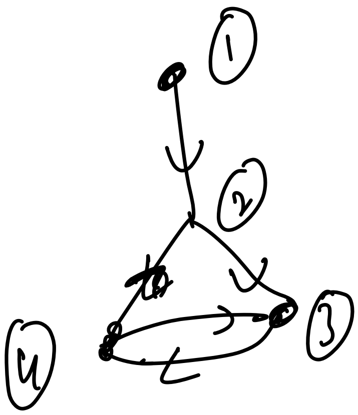
But web graph ~~may~~ is not be strongly connected ~~and~~.

Brin-Page "trick"

$(1-\varepsilon) G + \varepsilon K_n$.

↗ random surfer!





$$(1-\varepsilon)\begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \end{array} + \varepsilon \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Makes Markov chain ergodic and ~~complete~~ strongly connected

$\Rightarrow$ unique station distribution π for any $\varepsilon > 0$.

How to compute π?

Power method: $\pi_0 P^n \to \pi$ for any π_0

start with $\pi_0 = \begin{bmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{bmatrix}$.

$\left(\frac{1}{n}, \dots \frac{1}{n}\right)$.

$$\prod P = \prod \left((1-\varepsilon) P_{G} + \varepsilon P_{Kn} \right)$$

$$= \prod (1-\varepsilon) P_{G} + \varepsilon \prod P_{Kn}$$

can exploit sparsity of G.

is easy.

So computation is easy.

How to prove the Markov Chain Theorem?

Perron-Frobenius Theorem from linear algebra:

$\pi P = \pi \Rightarrow \pi$ is a left eigenvector of P with eigenvalue 1.

We are typically used to right eigen vectors.

Eigen vectors: A is a $n \times n$ matrix.

$Ax = \lambda x$ has a non-zero solution x

$\Leftrightarrow$ λ is an eigen value.

x is a corresponding eigen vector.

$\det(A - \lambda I)$ is the characteristic polynomial.

roots are eigen values. In general need not be real.

Spectral theory:

- If A is symmetric then all eigen values are real.
- If A is psd then all eigen values are ≥ 0.

But P is ~~neither~~ not symmetric.

Theorem [Perron]: Let A be ~~x~~ $n \times n$ positive matrix: ie $A_{ij} > 0 \ \forall i, j$.
Then A has a real positive eigen value λ and corr eigen vector $v > 0$

~~o~~ ~~and~~

(i) $\forall$ other eigen values λ'

$|\lambda'| < \lambda$ hence λ is the unique largest eigen value.

(ii) If $\lambda x \leq Ax$ for $x \geq 0$ then

$\lambda x = Ax$

(iii) Unique eigen vector for λ

ie $Ax = \lambda x \Rightarrow x = \alpha v$ for some scalar α.

Above requires $A > 0$.

P for a general graph has $A \geq 0$.

<u>Defn</u>: ~~A~~ A is $n \times n$ matrix $A \geq 0$ ie $A_{ij} \geq 0$ $\forall$ ij. We say A is irreducible if corresponding weighted directed graph is strongly connected.

<u>Theorem</u> [Frobenius] Let $A \geq 0$ and irreducible. Then A has a positive eigenvalue $\lambda \geq |\lambda'|$ for all other eigenvalues λ'. There is a positive eigenvector $v > 0$ corr to λ and following hold for λ and v.

① For any non-zero $x \geq 0$ if $\lambda x \leq Ax$ then $\lambda x = Ax$.

② If $\lambda x = Ax$ ~~and~~ then $x = \alpha v$ for some scalar α. ie unique eigenvector up to scaling.

<u>Corollary:</u> The largest real eigenvalue λ of an irreducible matrix $A \geq 0$ has a positive left eigenvector π. π is unique (up to scaling) and is the only non-zero vector that sati

$$\lambda \pi \leq \pi A.$$

Proof: Consider A^T. $A^T \geq 0$ and irreducible. A^T has same eigenvalues as A. π is ~~the~~ right eigen vector corr to . . .

Corollary: Let $A \geq 0$ λ largest real eigen value and $v > 0$ ~~be~~ $\pi > 0$ be right and left eigen vector of λ. Then v is only non-neg eigen vector of A and π is only non-neg ev of A.

Proof: From theorem $v > 0$ is unique right ~~eigenvalue~~ eigen vector of A for λ. Suppose u is the right eigen vector of $\lambda' \neq \lambda$. Claim u is not > 0. Suppose it is $\pi A u = \lambda \pi u$ and and $\pi A u = \lambda' \pi u$

$\Rightarrow \lambda \pi u = \lambda' \pi u$ since $\lambda \neq \lambda'$

$\pi u = 0 \Rightarrow u$ has to have 0 or -ve ..

Similarly $\pi \Phi$ = left eigen vector.

Now consider stochastic matrix P for an irreducible Markov Chain

$$P \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \Rightarrow \lambda = 1 \text{ is an eigen value.}$$

λ' is a real eigen value

$\Rightarrow \lambda' \qquad \lambda' \leq 1$

$$P x = \lambda' x$$

$$\lambda' \leq 1$$

$\Rightarrow \lambda = 1$ is the largest real eigen value

$\Rightarrow \exists$ left eigen vector $\pi > 0$.

unique. !

Now consider $A > 0$

Corollary: Let λ be largest eigenvalue of $A > 0$. Then π, v be left, right eigenvectors. $\pi v = 1$.

Then $\lim_{n\to\infty} \frac{A^n}{\lambda^n} = v\pi$.

↑ outer product

$$\begin{bmatrix} \pi_0 \end{bmatrix} \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix} = [\pi_0] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} [\pi] = [1][\ ] = \pi.$$

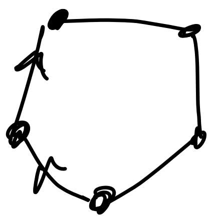
$$[a_1\ a_2\ a_n] \begin{bmatrix} \pi_1 & \pi_2 & \pi_n \\ \pi_1 & \pi_2 & \pi_n \\ \vdots\ \pi_1 & \pi_2 & \pi_n \end{bmatrix} = \begin{matrix} [\pi_1 .\ \pi_2 .\ \pi_n] \\ \pi \end{matrix}$$

---

what is π if P is periodic.

$\pi_v = \frac{1}{d}$ $\quad V_0 \cdots V_{d-1}$

as $\frac{1}{d|V_i|} = \pi_v \quad v \in V_i$



Personal / Contextual Page Rank

$$(1-\varepsilon) P_u + \varepsilon K_S$$

where S is a subset of "interesting" pages.

S is all pages with keyword. etc.

Exam :