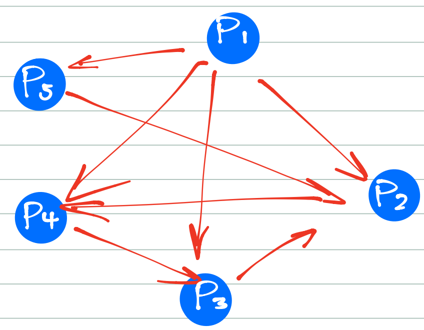
April 3rd

LAST LECTURE

PageRank

- web pages $\{P_1, \dots, P_n\}$
- links

Model:



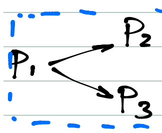
GOAL : rank pages

$I(P_i)$ = importance of P_i

Notation:
- P_i has l_i links
- B_i = pages links to P_i

e.g. $l_1 = 4$, $l_4 = 2$
$B_3 = \{P_1, P_4\}$

Axiom:
P_j will pass $\frac{1}{l_j}$ of its importance to P_i for every P_i that it links to.



$I(P_2) = \frac{1}{2}$

$$I(P_i) = \sum_{P_j \in B_i} \frac{I(P_j)}{l_j}$$

Def: Hyperlink matrix $H = \{H_{ij}\}$

$$H_{ij} = \begin{cases} \frac{1}{l_j} & P_j \in B_i \\ 0 & \text{otherwise} \end{cases}$$

In the example

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1 & 0 & 1 \\ 1/4 & 0 & 0 & 1 & 0 \\ 1/4 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 0 & 0 & 0 \end{bmatrix}$$

properties of H
1) all the entries are non-negative
2) all columns sum to 1
"stochastic matrices"

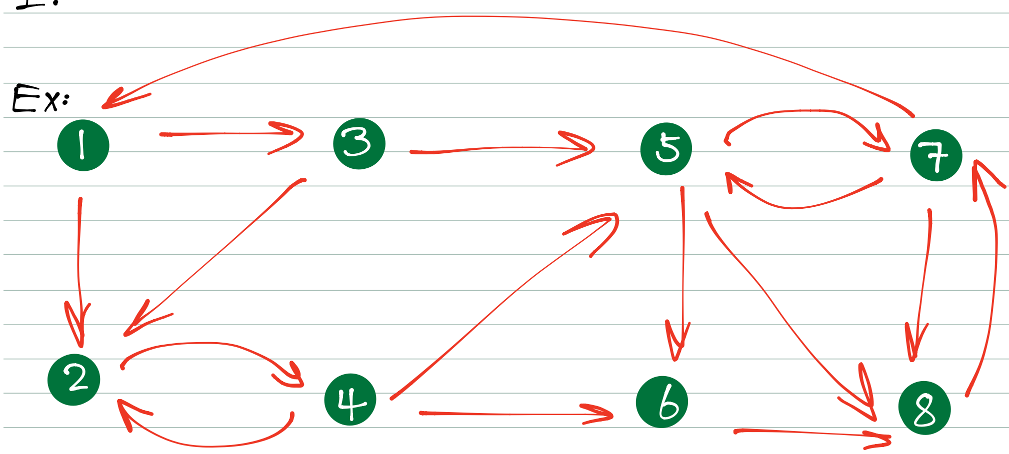
$$I = \begin{bmatrix} I(P_1) \\ \vdots \\ I(P_n) \end{bmatrix}$$

$$(HI)_i = \sum_{P_j \in B_i} \frac{I(P_j)}{l_j}$$

ith coord of this vector

* Same as requiring that $HI = I$, i.e. I is an eigenvector of H with evalue 1.

Ex:



$\rightsquigarrow H \rightsquigarrow$ find that H has an eigenvalue.

$$I = \begin{bmatrix} .06 \\ .0675 \\ .03 \\ .0675 \\ .0975 \\ .2025 \\ .180 \\ .29 \end{bmatrix}$$ of eigenvalue 1

ranking: 8
6
7
5
4, 2
1
3

Q: ① How to compute I? $H \sim 10^{10} \times 10^{10}$ matrix
② Is I unique?

The way Page Rank computes I:

1. Guess I_0
2. $I_1 = HI_0$
$I_2 = HI_1$
...
$I_k = H^k I_0$

H Stochastic $\Rightarrow I_k \to I$ as $k \to \infty$

$\begin{cases} \text{Under good circumstances, } H \text{ has eigenvalues } 1 = \lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \\ H \text{ diagonalizable} \end{cases}$

$\mathbb{R}^n$ has basis $\{v_1, \dots, v_n\}$ of eigenvectors

$I_0 = c_1 v_1 + \cdots + c_n v_n$ $\qquad$ $Hv_2 = \lambda_2 v_2$

$HI_0 = c_1 v_1 + c_2 \lambda_2 v_2 + c_3 \lambda_3 v_3 + \cdots + c_n \lambda_n v_n$

$$H^k I_0 = c_1 v_1 + c_2 \lambda_2^k v_2 + c_3 \lambda_3^k v_3 + \cdots + c_n \lambda_n^k v_n \longrightarrow c_1 v_1 \text{ as } k \longrightarrow \infty$$