$\boxed{\text{Lecture \#8}}$

April 2 2013

Page Rank Formula

$$\omega^T = \omega^T \left(sP + (1-s)E \right)$$

Page Rank Thesis.

"A page is important if it is pointed to by other impotant pages."

⇒ STATIONARY VALUES OF AN ENORMOUS MARKOV CHAIN. "MARKOV THEORY."

Problems with the Iterative Process

a) Convergence: Will this ~~process~~ iterative process continue indefinitely? Or will it converge?

b) Conditioning: Under what properties of P is it guaranteed to converge?

c) RELEVANCE: Will it converge to some thing "relevant"?

d) Will it converge to one solution or many?

e) Will the convergence depend on the initial vector ω_0^T?

f) $\boxed{\text{COMPLEXITY:}}$ If it will converge eventually, how long is this "eventuality?"

Recall the formulation:

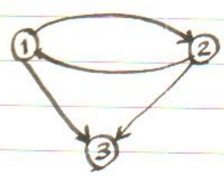
$$\omega_{k+1}^T = \omega_k^T P \Rightarrow \omega(i) = \sum \frac{A_{ij}\,\omega(j)}{d_{out}(j)}$$

1) Start with $\omega_0^T = \frac{1}{n} e^T$

Problem: Rank Sinks.

→ Few pages accumulate more and more PageRanks at each iteration (Dangling Nodes).

pp 46



$$A_{ij} \rightsquigarrow \quad A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$P_{ij} \sim \quad P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\omega_{k+1}^T = \omega_k^T P$$

Step 1. $$[1/3 \quad 1/3 \quad 1/3] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= [1/6 \quad 1/6 \quad 1/3]$$

Step 2. $$[1/6 \quad 1/6 \quad 1/3] \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= [1/12 \quad 1/12 \quad 1/6]$$

Rank $\quad \omega(1) = 2; \ \omega(2) = 3; \ \omega(3) = 1.$

CYCLE

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Suppose the iterative process starts with

$$\omega_0^T = (1 \quad 0)$$

$$\omega_1^T = (0 \quad 1)$$

$$\omega_2^T = (1 \quad 0) = \omega_0^T$$

Indefinite flip-flops.

Such iterations will not converge no matter how long the process runs.

$$\omega_0^T = (1/2 \quad 1/2)$$

$$\omega_1^T = (1/2 \quad 1/2) = \omega_0^T$$

THEORY OF MARKOV CHAIN.

P = Transition Probability Matrix

Power method, for any starting vector, applied to a Markov matrix P converges to a unique positive vector:

STATIONARY VECTOR.

P = Stochastic
&
Primitive $\Rightarrow$ ≡ Aperiodic
+
Irreducible.

Page Rank convergence problems (due to sinks and cycles) can be overcome if P is modified slightly so that

$$P' = \varepsilon P + (1-\varepsilon)\, ee^T$$

is a Markov matrix with desired property.

(I) A unique positive PageRank vector exists when the Google matrix is stochastic and irreducible.

(II) The power method converges to this PageRank vector (regardless of the initial starting vector), if the Google matrix is also aperiodic.

$$P = \qquad P_{ij} = \frac{A_{ij}}{d_{out}(i)}$$

$$P' = P + a\left(\frac{1}{n} e^T\right)$$

$$a_i = \begin{cases} 1 & \text{if } i \text{ is a dangling node} \\ 0 & \text{otherwise.} \end{cases}$$

$\frac{1}{n} a\, e^T$ = Rank-1 matrix.

a = dangling node vector.

$$P'' = sP' + (1-s)\frac{ee^T}{n}$$

$$= sP + \frac{s}{n} a e^T + \frac{(1-s)}{n} ee^T$$

$\equiv G$ = Google Matrix.

(I) G = Stochastic.
∵ A convex combination of three stochastic matrices:

$$P, \ \frac{ae^T}{n}, \ \frac{ee^T}{n}$$

(II) G = Irreducible
∵ Every page is connected (by teleportation) to every other page.

(III) G = Aperiodic
∵ Self-loops ($G_{ii} > 0$) creates a periodicity. Note that one can teleport to a node from itself.

$\exists_k \ G^k > 0$ (e.g. $G^1 > 0$).

G = Primitive.

Problem → G = DENSE
(although, a rank-1 update to a sparse matrix.)