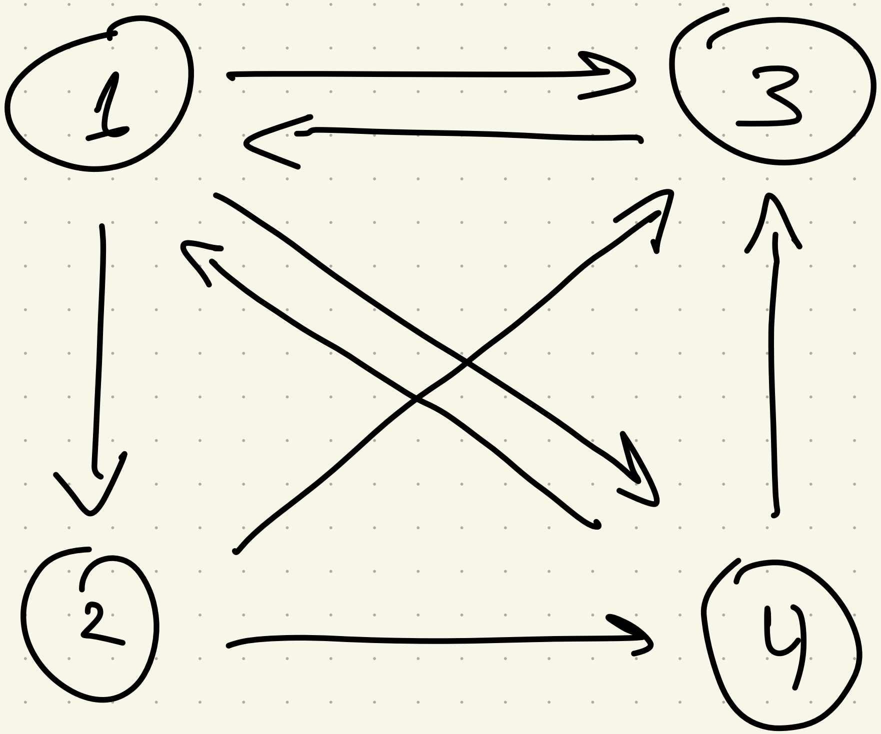
CS Lens

PageRank

Afonso Bandeira

6.12.2023

(see more in "Mathematics of Data Science" book draft on my page)



$x_k \in \mathbb{R}$ ← "importance of page k".

Idea #1: x_k = # back links to (k).

some backlinks matter more than others.

Idea #2:

$$X_k = \sum_{j \in E_k} x_j ,$$

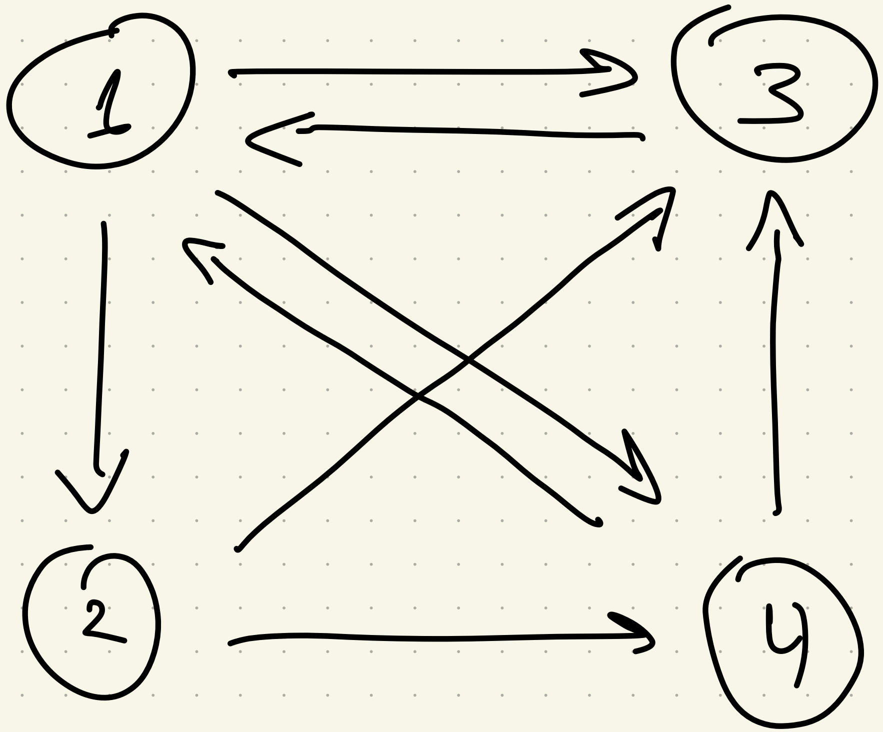
E_k set of pages that link to k

Idea #3:

normalize by # of out-links

$$X_k = \sum_{j \in E_k} \frac{X_j}{n_j}$$

n_j = # of outlink in page j.



$$x_1 = 1\,x_3 + \frac{1}{2}x_4$$

$$x_2 = \frac{1}{3}x_1$$

$$x_3 = \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_4$$

$$x_4 = \frac{1}{3}x_1 + \frac{1}{2}x_2$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$x = Ax.$$

X is an eigenvector of A corresponding to evalue 1.

Prop: if λ is an e-value of M then it is also an e-value of M^T.

$$A^T = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

$$A^T \underline{1} = \underline{1}.$$

$$\underline{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

see: Perron-Frobenius.