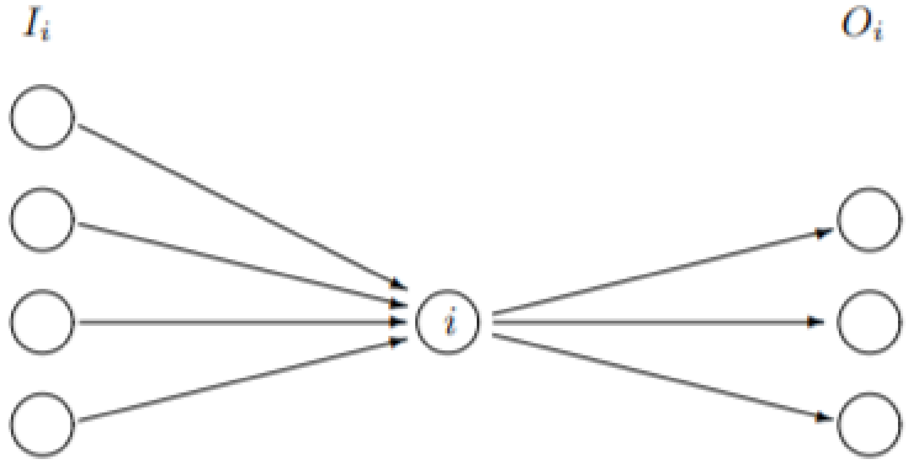
Page Ranking for a Web Search Engine

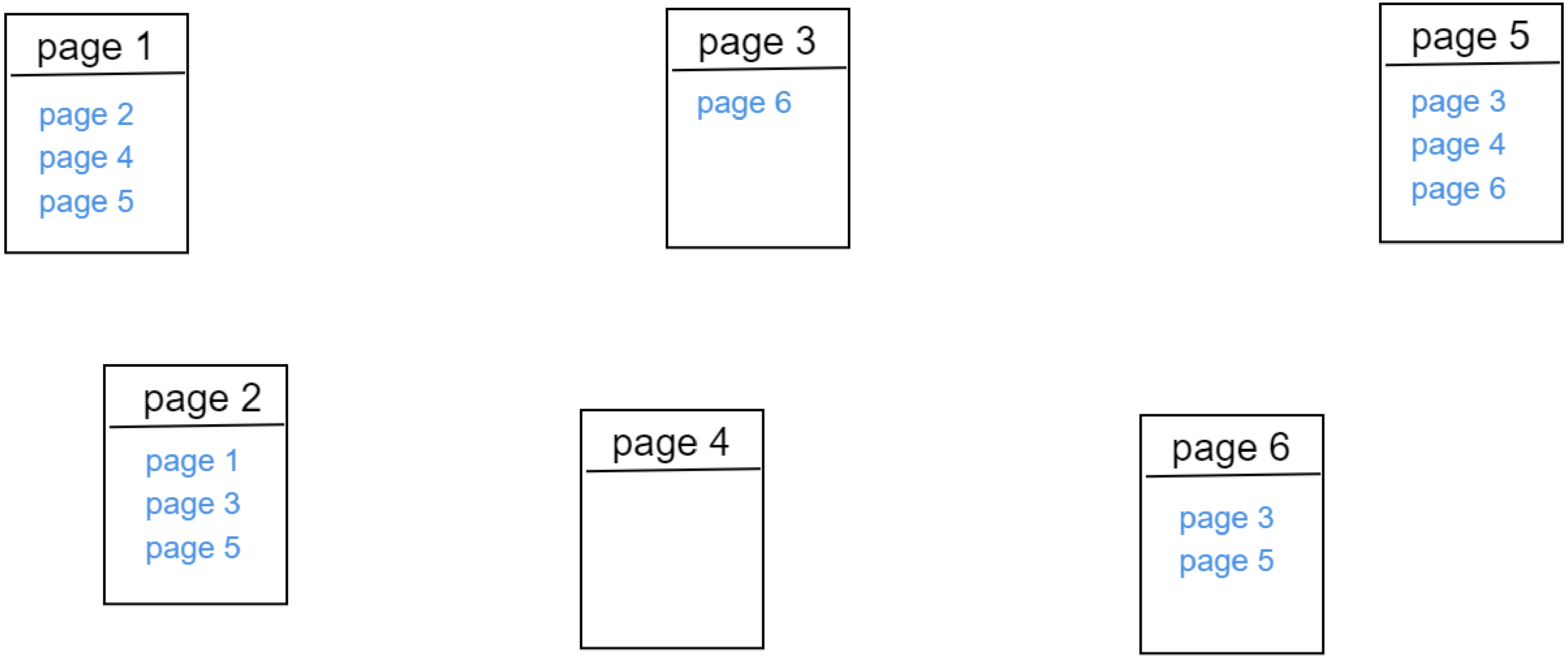
- the number of links to and from a page give information about the importance of a page.

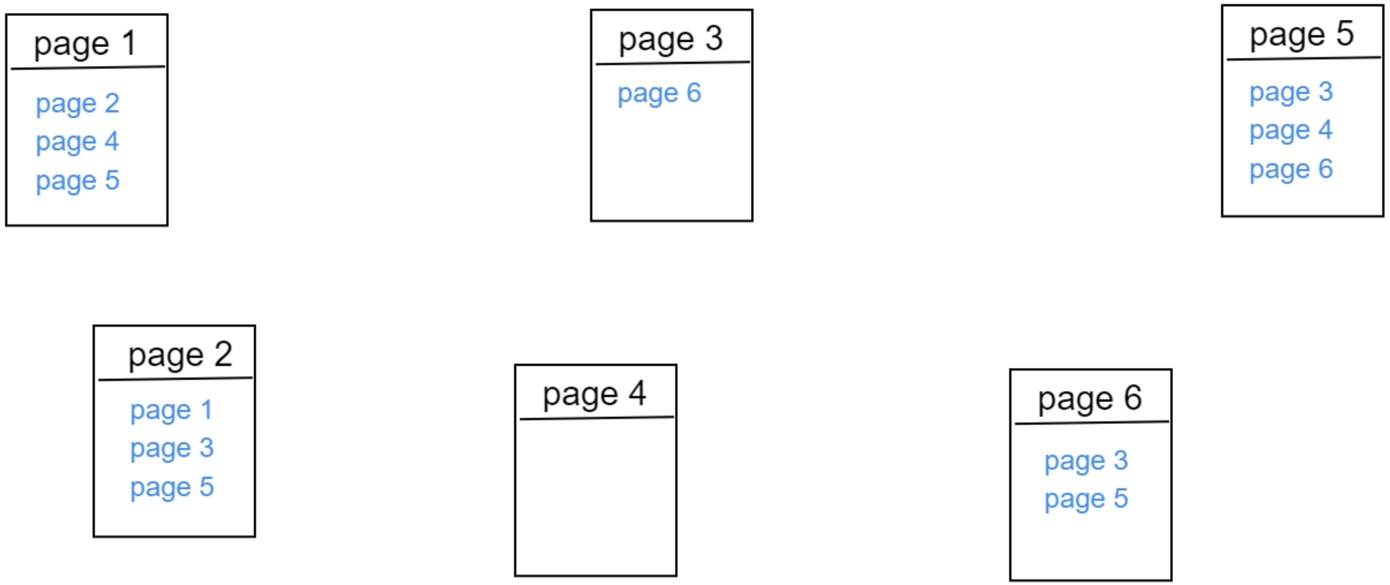


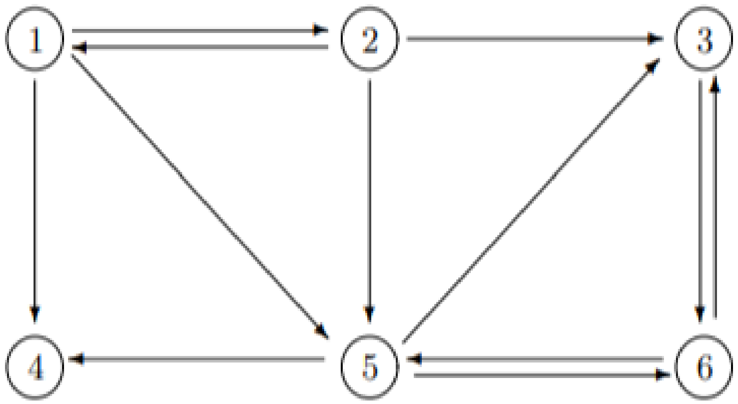
$$r_i = \sum_{j \in I_i} \frac{r_j}{N_j}.$$

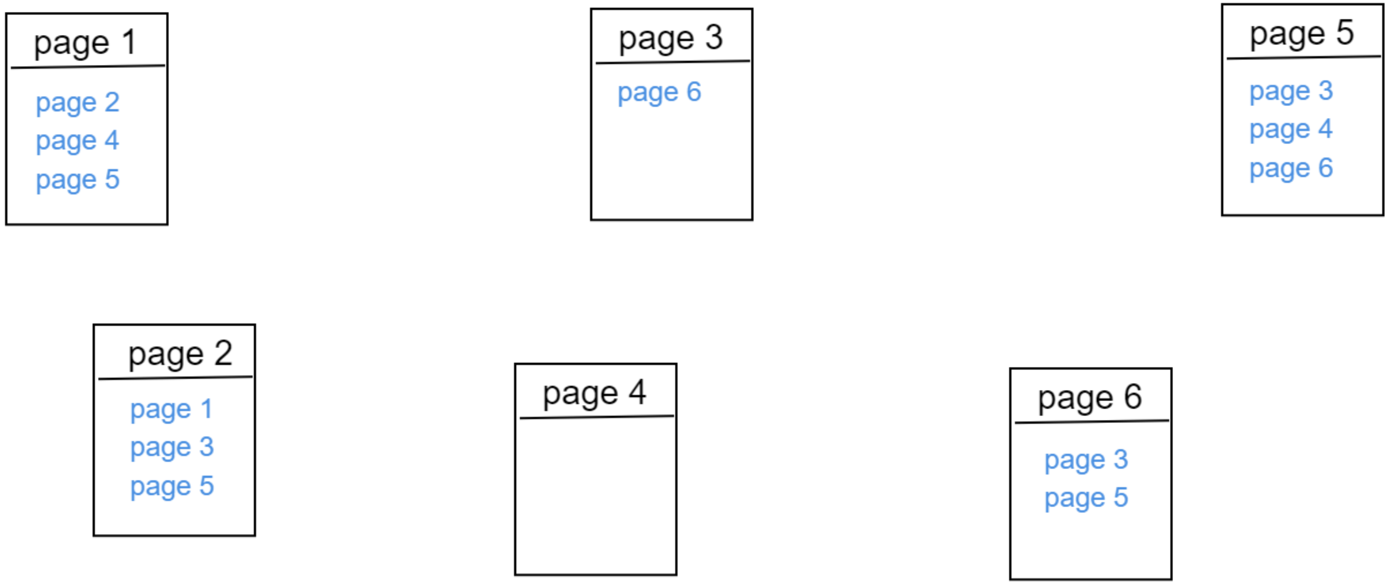
$$r_i = \sum_{j \in I_i} \frac{r_j}{N_j}.$$

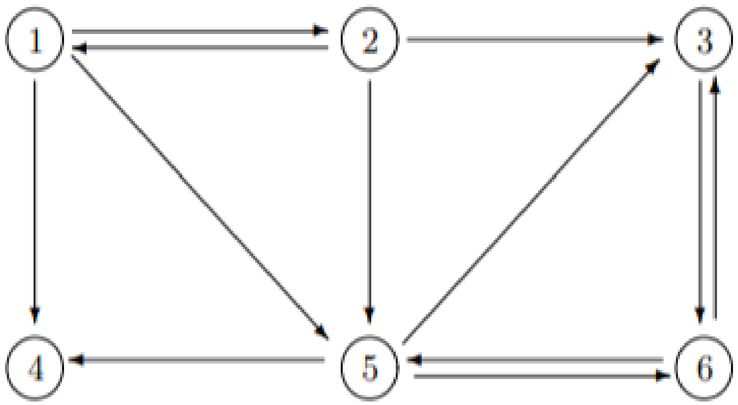
$$r_i^{(k+1)} = \sum_{j \in I_i} \frac{r_j^{(k)}}{N_j}, \quad k = 0, 1, \ldots.$$


page 1
page 2
page 4
page 5
page 3
page 6
page 5
page 3
page 4
page 6
page 2
page 1
page 3
page 5
page 4
page 6
page 3
page 5


page 1
page 2
page 4
page 5
page 3
page 6
page 5
page 3
page 4
page 6
page 2
page 1
page 3
page 5
page 4
page 6
page 3
page 5


1
2
3
4
5
6





Corresponding Transition Probibility Matrix:

$$Q = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

$$Q = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

$$r_i = \sum_{j \in I_i} \frac{r_j}{N_j}.$$

$$Q = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

$$r_i = \sum_{j \in I_i} \frac{r_j}{N_j}.$$

is equivalent to the scalar product of row i and the vector r, which holds the ranks of all pages

$$Q = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

$$r_i = \sum_{j \in I_i} \frac{r_j}{N_j}.$$

is equivalent to the scalar product of row i and the vector r, which holds the ranks of all pages :

$$\lambda r = Qr, \qquad \lambda = 1,$$

i.e., r is an eigenvector of Q with eigenvalue $\lambda = 1$

$$r_i^{(k+1)} = \sum_{j \in I_i} \frac{r_j^{(k)}}{N_j}, \quad k = 0, 1, \ldots.$$

$$r^{(k+1)} = Qr^{(k)}, \qquad k = 0, 1, \ldots,$$

Equivalent definition

$$r^{(k+1)} = Qr^{(k)}, \qquad k = 0, 1, \ldots,$$

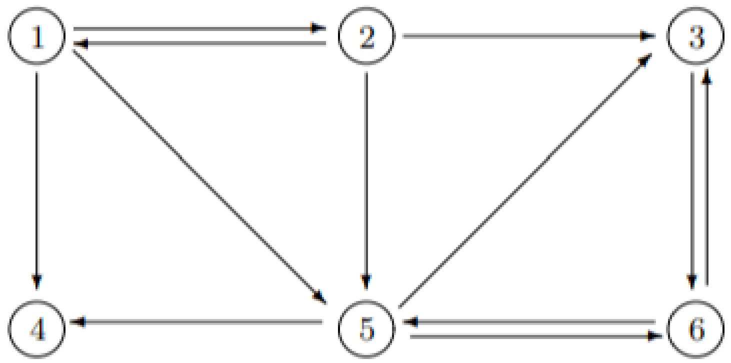
Equivalent definition

$$r^{(k+1)} = Qr^{(k)}, \qquad k = 0, 1, \ldots,$$

- **If i were to randomly click around on this graph where would i end up spending most of my time?**

Trap

If i were to randomly click around on this graph where would i end up spending most of my time?



$$Q = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

A square matrix A is called reducible if there is a permutation matrix P such that :

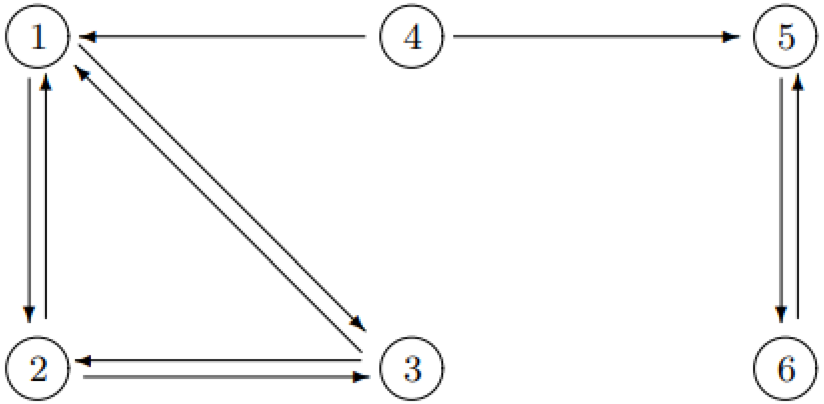
$$PAP^T = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$$

where X and Z are both square. Otherwise the matrix is called irreducible

A square matrix A is called reducible if there is a permutation matrix P such that :

$$PAP^T = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$$

where X and Z are both square. Otherwise the matrix is called irreducible



$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

Where is the trap of this graph?

Eliminate traps:

$$Q' = (1-\alpha)Q + \frac{\alpha}{N}\ 1_{n\times n}$$

$$\frac{\alpha}{N}: random\ jump\ probability$$

$$Q' \, is \; irreducible \; .why?$$

$$Q' = (1-\alpha)Q + \frac{\alpha}{N} \; 1_{n\times n}$$

$$\frac{\alpha}{N} : random \; jump \; probability$$

We know the vector r, which holds the ranks of all pages
r is an eigenvector of Q’ with eigenvalue $\lambda = 1$ but :

the existence of a unique eigenvalue with eigenvalue 1
is still not guaranteed

$\boldsymbol{Let\ Q\ \ be\ an\ irreduible(Notraps), column - stochastic\ matrix\ with\ \ non\ negetive\ entries}$

$\boldsymbol{\Rightarrow Perron - Frobenius}:$

$Largest\ eigenvalue = 1\ ,$

$eigenvector\ P = [p_1 \dots p_N]^T\ satisfies : p_i > 0\ , \sum p_i = 1,$

$u = \dfrac{1}{N}[1,..,1]^T\ then\ \lim_{k\to\infty} Q^k u = P$

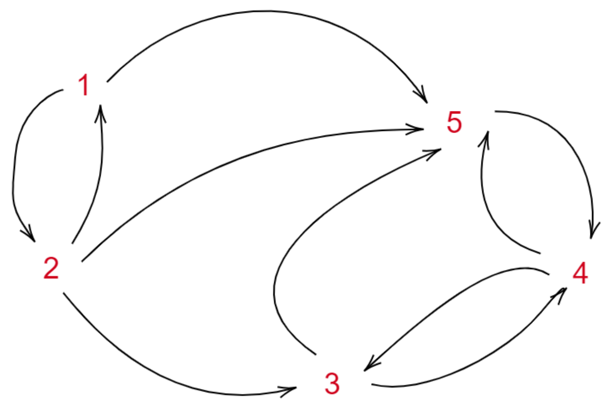
Summary

- Adjancency matrix
- Normalize columns
- Eliminate trapes
- Use power method for computing the eigenvector

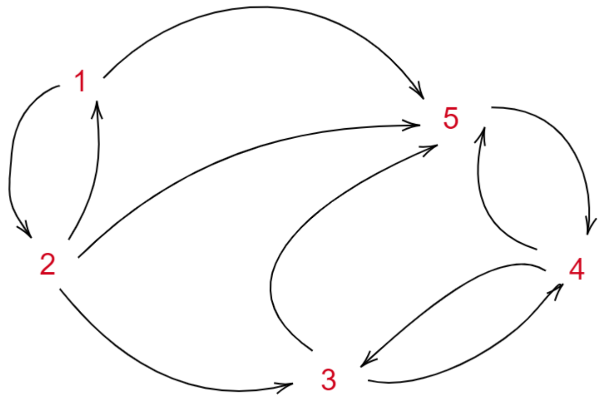
Summary

- Adjancency matrix
- Normalize columns
- Eliminate trapes
- Use power method for computing the eigenvector

Example:

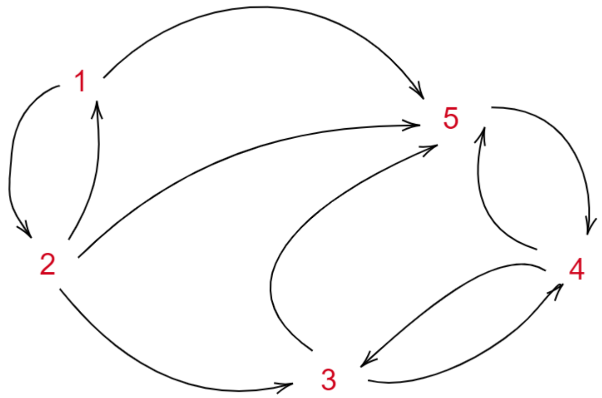


- Adjancency matrix :

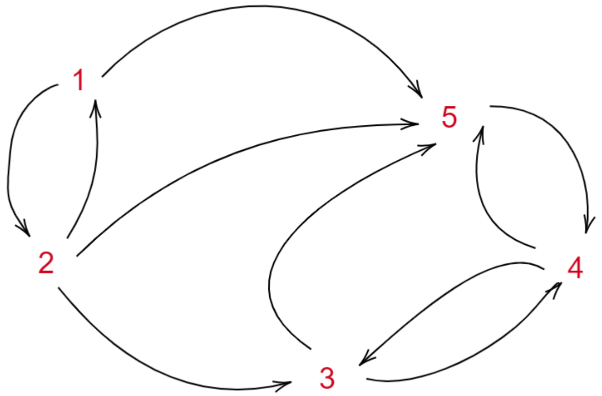


- Adjancency matrix :

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

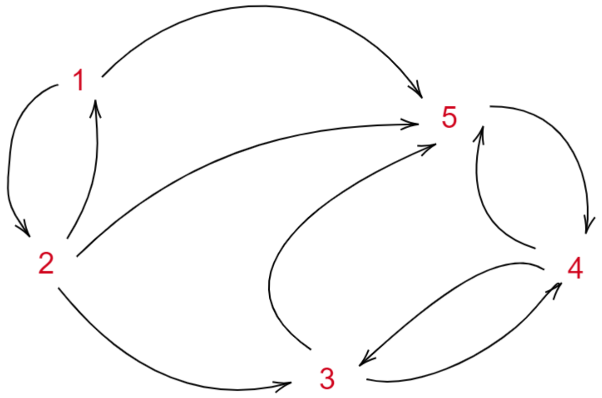


• Normalize columns:



- Normalize columns:

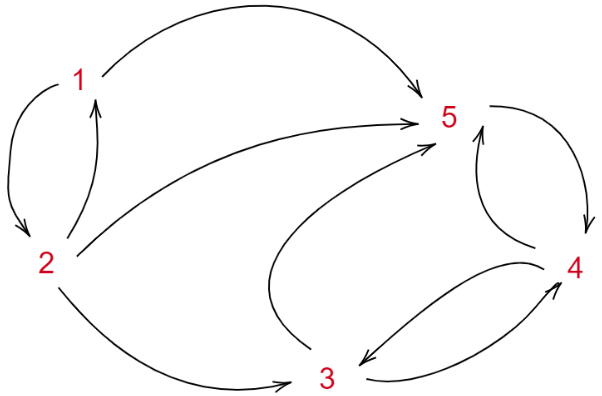
$$\begin{bmatrix} 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 1/2 & 1/3 & 1/2 & 1/2 & 0 \end{bmatrix}$$



- Eliminate trapes:

$$Q' = (1-\alpha)Q + \frac{\alpha}{N}\, 1_{n\times n}$$

$$\frac{\alpha}{N} : random\ jump\ probability$$

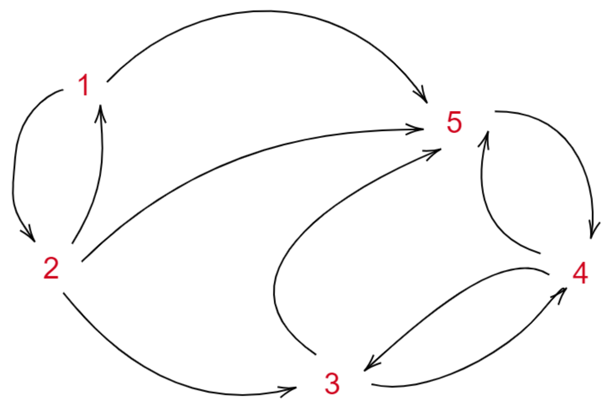


- Eliminate trapes:

$$Q' = (1 - \alpha)Q + \frac{\alpha}{N} 1_{n \times n}$$

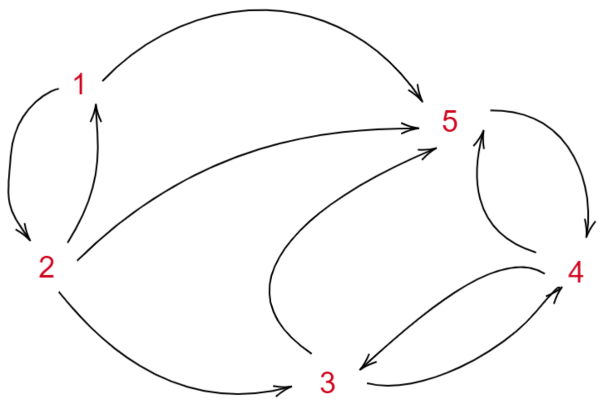
$$\frac{\alpha}{N}: random\ jump\ probability$$

$$\alpha = 0.01, Q' = \begin{bmatrix} 0.002 & 0.332 & 0.002 & 0.002 & 0.002 \\ 0.497 & 0.002 & 0.002 & 0.002 & 0.002 \\ 0.002 & 0.332 & 0.002 & 0.497 & 0.002 \\ 0.002 & 0.002 & 0.497 & 0.002 & 0.992 \\ 0.497 & 0.332 & 0.497 & 0.497 & 0.002 \end{bmatrix}$$



- Use power method for computing the eigenvector

$$u = \begin{bmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{bmatrix}, Q'u = \begin{bmatrix} 0.07 \\ 0.0 \\ 0.7 \\ 0.30 \\ 0.36 \end{bmatrix}, Q'^2u = \begin{bmatrix} 0.003 \\ 0.004 \\ 0.8 \\ 0.45 \\ 0.30 \end{bmatrix}, \ldots, Q'^{10}u = \begin{bmatrix} 0.003 \\ 0.004 \\ 0.22 \\ 0.44 \\ 0.33 \end{bmatrix}, P = \begin{bmatrix} 0.0032 & 0.0036 & 0.2211 & 0.4401 & 0.3320 \end{bmatrix}$$



Applications of PageRank Algorithm:

Search Engine Ranking
Internet Marketing
Social Network Analysis
Content Recommendation
Graph Analysis