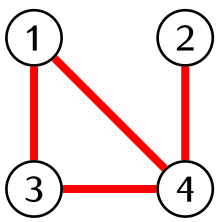
Steady state vector of a random walk on an undirected connected network

Example.



$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1/2 & 1/3 \\ 0 & 0 & 0 & 1/3 \\ 1/2 & 0 & 0 & 1/3 \\ 1/2 & 1 & 1/2 & 0 \end{bmatrix} \end{array}$$

$$A = \underbrace{\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}}_{\text{adjacency matrix}} \qquad D = \underbrace{\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \end{array}}_{\text{degree matrix}} \qquad D^{-1} = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}$$

Note: $P = A \cdot D^{-1}$. It follows that if Y is a steady state vector of P then:

$$Y = PY$$
$$IY = AD^{-1}Y$$
$$(I - AD^{-1})Y = 0$$
$$(DD^{-1} - AD^{-1})Y = 0$$
$$(\underbrace{D - A}_{\text{Laplacian}})D^{-1}Y = 0$$

Recall: If L is the Laplacian of a connected undirected graph then $Lv = 0$ if and only if $v = c \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

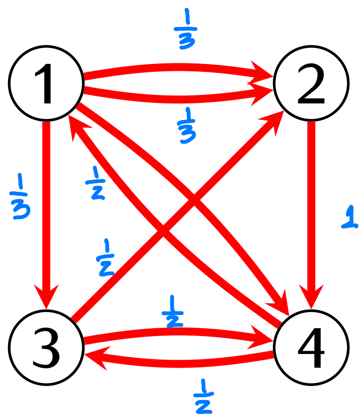
We obtain: $D^{-1}Y = c\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ so: $Y = cD\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

$$= c\begin{bmatrix} k_1 \\ \vdots \\ k_N \end{bmatrix} \Big\} \text{ degrees of vertices of the graph}$$

Y - probability vector $\Rightarrow c = \frac{1}{\sum k_i}$. Thus $Y = \frac{1}{\sum k_i}\begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix}$

Random walks on directed networks and the Google PageRank

Example. Network of web pages:



Question. How to rank web pages?

PageRank:

- Consider the steady-state vector of a random walk on the network of pages. It gives probabilities that a walker will visits each page in a long run.
- Higher probability of a page means that the page is more popular.
- Rank pages according to these probabilities.

Recall:

Definition

A stochastic matrix P is *regular* if there is $N \geq 0$ such that all entries of P^N are positive.

Perron-Frobenius Theorem

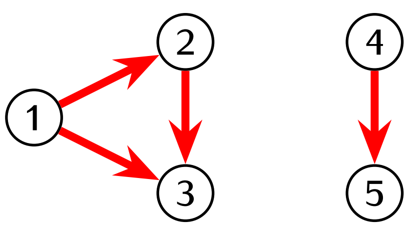
If P is a regular stochastic matrix then:

- There exists only one steady state vector Y of P
- For any probability vector X we have

$$\lim_{n} P^n X = Y$$

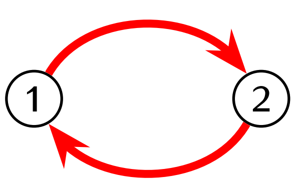
Note. The transition matrix for a random walk on a network of web pages need not be regular.

Problem: a disconnected network.



It is impossible to get e.g. from 1 to 5 in any number of steps.

Problem: cycles.



We can get from 1 to 2 only in an odd number of steps, but from 1 to 1 only in an even number of steps. Thus if $P^n = (a_{ij})$ then $a_{11} = 0$ if n - odd, $a_{21} = 0$ if n - even.
So P is not regular

Solution.

Modified transition matrix

$P =$ the transition matrix of the random walk

$q =$ some number $0 < q < 1$

$$R = \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & & \vdots \\ 1/n & \cdots & 1/n \end{bmatrix}$$

– a matrix of the same size as P with all entries equal to $\frac{1}{n}$ (where $n =$ the number of webpages in the network)

Take:

$$Q = \underbrace{q}_{\text{probability that a user will continue the random walk}} P + \underbrace{(1-q)}_{\text{probability that a user will abandon the random walk and pick up a new webpage at random}} R$$

Q is a regular matrix, so Perron-Frobenius theorem applies to it.