Chan and Golab
Page Rank

Introduction

PageRank is Google's way of determining the relevancy / importance of any given web page. It does this by looking at
a) the number of links going into a page
b) how important the sources of those incoming links are
c) how many other links the source of a given link has.

PageRank uses the "random surfer" model - a bored internet user mindlessly clicking on links available in a page. The actual PageRank value is the probability that this user will end up on some given page in a very large, but finite number of clicks.

Computational Varieties

Variant 1: Monte Carlo
a) Choose a page (vertex) at random
b) Choose a large integer k
c) Follow a link on the page (an outgoing edge) k times uniformly at random
d) Average the final results over a bunch of trials to determine probability

This is a totally okay variant, and isn't too difficult to understand. Given it's nature, though, it's hard to know how accurate your final result is.

That's why the linear algebra method exists

Variant 2: Linear Algebra

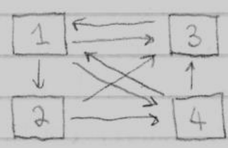
a) Consider all possible starting positions simultaneously using a stochastic vector

b) Consider all possible out-edges simultaneously using a stochastic matrix

c) Repeatedly multiply the vector by the matrix until it converges or an iteration limit is reached

Okay, so what the fuck does that mean? Let's go through an example.

ex.



This graph represents our network of connected web pages.

We'll solve this using steps.

1) Assign uniformly random probabilities to edge transitions.

1 has 3 transitions: $1 \to 2 : 1/3$
$1 \to 3 : 1/3$
$1 \to 4 : 1/3$

2 has 2 transitions: $2 \to 4 : 1/2$
$2 \to 3 : 1/2$

3 has 1 transition: $3 \to 1 : 1$

4 has 2 transitions: $4 \to 1 : 1/2$
$4 \to 3 : 1/2$

2) Construct a transition matrix from the assigned probabilities.

Each column is one web page, and each row is the transition to a corresponding page's probability.

$$\begin{array}{c c} 4 & \\ 3 & \\ 2 & \\ 1 & \\ & \end{array} \begin{array}{c} \begin{bmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \\ \begin{matrix} 1 & \;\; 2 & \;\; 3 & \;\; 4 \end{matrix} \end{array}$$

So A_{ij} (i = row, j = column) is the probability of transitioning from $j \to i$.

$A_{41} = 1/2$, for example, as we've identified earlier. As these are representative of probabilities, each COLUMN must sum to 1 (not needed for rows).

3) Create an initial rank vector with even probability distribution.

We've got four nodes, so $v = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$

4) Calculate Av, then $A^2 v$, then $A^3 v$, etc, until the result converges.

Quick reminder on matrix multiplication

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix}$$

2 by 3   3 by 1 = 2 by 1

have to match

where $m_{11} = a_{11} \times b_{11} + a_{12} \times b_{21} + a_{13} b_{31}$

$m_{21} = a_{21} \times b_{11} + a_{22} \times b_{21} + a_{23} \times b_{31}$

As such,

$$\underset{4 \text{ by } 4}{\begin{bmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix}} \underset{4 \text{ by } 1}{\begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}} = \underset{4 \text{ by } 1}{\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix}}$$

$a_{11} = (0)(1/4) + (0)(1/4) + (1)(1/4) + (1/2)(1/4)$
$= 3/8$

$a_{21} = (1/3)(1/4) + (0)(1/4) + (0)(1/4) + (0)(1/4)$
$= 1/12$

$a_{31} = (1/3)(1/4) + (1/2)(1/4) + (0)(1/4) + (1/2)(1/4)$
$= 1/3$

$a_{41} = (1/3)(1/4) + (1/2)(1/4) + (0)(1/4) + (0)(1/4)$
$= 5/24$

$\therefore\ Av = \begin{bmatrix} 3/8 \\ 1/12 \\ 1/3 \\ 5/24 \end{bmatrix}$ To spare us the calculation trouble, $A^7 v = \begin{bmatrix} 0.38 \\ 0.12 \\ 0.29 \\ 0.19 \end{bmatrix}$, $A^8 v = \begin{bmatrix} 0.38 \\ 0.12 \\ 0.29 \\ 0.19 \end{bmatrix}$

And since it converges, the pageranks of 1, 2, 3, 4 are 0.38, 0.12, 0.29, and 0.19, respectively.

This result is EXACTLY the eigenvector of A with $\lambda = 1$. What does this mean, again?

Let's put this another way. The eigenvalue (λ) MUST be 1.
How do we get an eigenvector from A, knowing that?

$$\begin{bmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \overset{1}{\lambda} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

This is the expression for how to find the eigenvector. We can solve this by converting it to a system of equations.

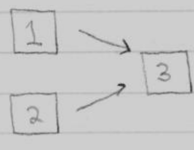
$$\begin{cases} 0a + 0b + c + \frac{1}{2}d = a \\ \frac{1}{3}a + 0b + 0c + 0d = b \\ \frac{1}{3}a + \frac{1}{2}b + 0c + \frac{1}{2}d = c \\ \frac{1}{3}a + \frac{1}{2}b + 0c + 0d = d \end{cases} \Rightarrow \begin{matrix} a = 0.38 \\ b = 0.12 \\ c = 0.29 \\ d = 0.19 \end{matrix}$$

Some questions:

1) Does an eigenvector for $\lambda = 1$ exist for A?
2) Is there a unique eigenvector, with unique values?
3) Will the power method converge on the eigenvector?
4) Will the power method converge at all?

Ideally, all of these should be yes. But when are they no?
We can answer these questions with another example.

ex.



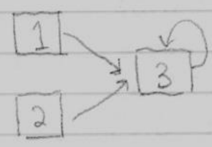
Here, we have a dangling node - one with no outbound transitions

This creates $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & \underbrace{0}_{\text{doesn't sum to 1!}} \end{bmatrix}$ $v = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$

Using the power method:

$$A v = \begin{bmatrix} 0 \\ 0 \\ 2/3 \end{bmatrix} \qquad A^2 v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Well. This is obviously an undesirable result. PageRanks of all zero doesn't make any sense at all, and $\lambda = 1$ is not an eigenvalue as all-zero vectors don't count as valid. Okay, so why don't we try something else, then. We'll add in a transition to itself on node 3:



$$\Rightarrow A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

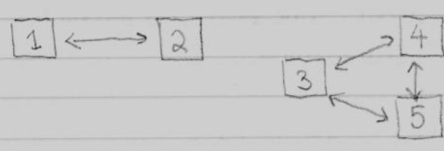
now sums to 1!

Let's try the power method once more.

$$A v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad A^2 v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

So this makes slightly more sense now, but we still shouldn't have pageranks of zero - it's not like these don't exist or that you can never get to them.

Let's see how else things can get messed up.



What if the sets of nodes are partitioned?

This creates

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{bmatrix}$$

which converges at $Av = \begin{bmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{bmatrix}$

We've converged, but the eigenvector doesn't have unique values. We want a distinct ordering of pages, so this won't do, especially when 3/4/5 are arguably more important than 1/2 as they have more incoming links.

One more.

$$\boxed{1} \rightleftarrows \boxed{2} \rightleftarrows \boxed{3}$$

This one has no dangling nodes and its nodes aren't disjoint, so we should expect it to work, right?

$$A = \begin{bmatrix} 0 & 1/2 & 0 \\ 1 & 0 & 1 \\ 0 & 1/2 & 0 \end{bmatrix} \quad v = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$$Av = \begin{bmatrix} 1/6 \\ 2/3 \\ 1/6 \end{bmatrix} \quad A^2 v = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \quad A^3 v = \begin{bmatrix} 1/6 \\ 2/3 \\ 1/6 \end{bmatrix}$$

You thought wrong. The value never converges

Convergence Theory

Theorem 1 (Perron-Frobenius):

A real square matrix with POSITIVE entries has a dominant real eigenvalue. The corresponding eigenspace has dimension one, and contains a vector with strictly positive components.

Definition 1:

A matrix is column stochastic if it is square, all entries are REAL and NON-NEGATIVE, and all columns sum to one.

Theorem 2:

The largest real eigenvalue of a column stochastic matrix is $\lambda = 1$.

Corollary:

A column stochastic matrix with positive entries has a dominant real valued eigenvalue $\lambda = 1$ with multiplicity one, and its eigenspace contains a vector with positive components

So in summary, it means if we can make

a) a square matrix

b) whose columns sum to 1

c) and whose values are strictly positive

then that means we'll converge using the power method

But how do we avoid 0s in our probability matrices? We give the user a chance to get bored with a page and randomly browse to any other page.

We create a new matrix M:

$$M = \underbrace{(1-\delta)}_{\text{probability user gets bored}} \begin{bmatrix} 1/n & 1/n \\ 1/n & 1/n \end{bmatrix} + \underbrace{\delta}_{\text{probability user is still interested and keeps following links}} \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

We add in a damping factor δ as a probability that the user is still interested. This allows us to create a column stochastic matrix with only positive values, which by the Perron-Frobenius theorem gives us an eigenvector of multiplicity one (unique values) with the eigenvalue $\lambda = 1$.

We then perform Mv, M^2v, etc, until we converge on the guaranteed eigenvector