Spectral Ranking

Sebastiano Vigna
Dipartimento di Informatica
Università degli Studi di Milano
A Historical Talk

- This talk is about *spectral ranking*

- PageRank is just the currently trendy incarnation of spectral ranking

- The main ideas were developed at the end of the XIX century, and then in the late forties and in the early fifties of the XX century

- However, the connection between these ideas emerged during the study of PageRank
Perspective

- PageRank is probably the most talked-about ranking algorithm ever

- Nonetheless, we have no scientific, reproducible proof that it works (quite the opposite)...

- ...and it’s likely to be of minuscule importance in today’s ranking

- Nonetheless, the idea appear to be useful in several applications (Gleich 2015; 2018 SIAM Outstanding Paper)
Basic Setup

- $M$ is a 0/1 matrix representing relations between entities
- $M$ might contain inconsistent information, as in...
- $i$ likes $j$, $j$ likes $k$, but $i$ does not like $k$, or...
- $i$ is better than $j$, $j$ is better than $k$, but $i$ is not better than $k$
- Of course $M$ is also (the adjacency matrix of) a graph
Chess

• Imagine $M$ contains 0 or 1 in row $i$ and column $j$ depending on whether $i$ defeated $j$ (1/2 for ties)

• Row sums $M1^T$ provide an easy global score: you get 1 for each win, 1/2 for each tie

• The Austrian chess player Oscar Gelbfuhs proposed in 1873 to do it again: let’s add the global score of the players you defeated plus 1/2 of the global scores of the players you had a tie with

• This is just $M(M1^T) = M^21^T$

• In 1895 Edmund Landau (17!) in his first published paper suggest that since $M^k1^T$ oscillates, we should look for a global score $r$ satisfying $Mr^T = \lambda r^T$, as it will be essentially fixed by the procedure: that is, we should look for a right eigenvector
John R. Seeley (1949) wants to assign popularity scores to children of a kindergarten (he’s a psychologist)

Given $M$ containing 0 or 1 depending on whether $i$ likes $j$, Seeley argues that the score of a child should be the sum of the scores of the children that like him: $s_i = \sum_{j ightarrow i} s_j$

But you spread your popularity evenly: $s_i = \sum_{j ightarrow i} s_j/d_j$

And here we are! Seeley computes the dominant left eigenvector of $M$, $\ell_1$-normalized by row
# How It Works

\[
\begin{array}{c|ccccc}
X_0 & X_1 & X_2 & X_3 & X_4 \\
\hline
0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\
\frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\
\end{array}
\]

\[
\frac{1}{3}x_3 + \frac{1}{2}x_4 & \quad \frac{1}{3}x_0 + \frac{1}{3}x_3 & \quad \frac{1}{3}x_0 + \frac{1}{2}x_2 + \frac{1}{3}x_3 & \quad x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_4 & \quad \frac{1}{3}x_0 \\
\]
The Markovian View

- We normalise $M$ by row, getting $P$

- $P$ express the probability that we try to meet child $j$ after meeting child $i$...

- ...or, if you want, that we visit page $j$ after visiting page $i$.

- A dominant left eigenvector is a *stable state* or *stationary distribution*
Perron–Frobenius

- If $M$ is nonnegative, the spectral radius is a dominant eigenvalue and there is a nonnegative dominant eigenvector.

- $M$ is irreducible (iff the underlying graph if strongly connected) iff the dominant eigenvector is unique and strictly positive.

- $M$ is unichain iff the dominant eigenvector is unique.

- Otherwise, many possible solutions (Markovianly speaking, depending on the initial distribution).

- Note: at the time of Landau’s paper, these results were not known: Landau will go back on this topic in 1915, in fact dismissing the method because of a pathological counterexample built in collaboration with Perron (!).
• Teh-Hsing Wei (1952) wants to rank sport teams

• Given $M$ containing 0, 1/2 or 1 depending on whether $i$ defeated $j$, $i$ tied with $j$, or $i$ lost with $j$...

• ...Wei argues that the score of a team should be the sum of the scores of the teams it defeated, plus half the sum of the scores of the teams with which there was a tie...

• ...so he redisCOVERS Landau’s idea. Kendall will popularize it.

• Berge (1958) generalizes to arbitrary 0/1 matrices and remarks the connection with path counting
Spectral Ranking

• Given a matrix $M$ with a real, positive, strictly dominant eigenvalue

• The (left) spectral ranking of $M$ is its (left) dominant eigenvector

• Left eigenvectors are good for endorsement; right eigenvectors for “better than” relationships (or you can just transpose your matrix, of course)
Damping

• In 1953, Leo Katz introduces his famous index

• Given $M$ containing 0 or 1 depending on whether $i$ chooses/ endorses/votes for $j$...

• ...Katz claims that the importance of $i$ depends not only on the number of the voters, but on the number of the voters’ voters, etc., with suitable attenuation $\alpha < 1 / \rho(M)$

• He computes $1 + 1\alpha M + 1\alpha^2 M^2 + 1\alpha^3 M^3 + \cdots = 1\sum_{n \geq 0} \alpha^n M^n = 1(1 - \alpha M)^{-1}$
How it works

$$\begin{array}{cccc}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
\end{array}$$

$$\begin{array}{cccc}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
\end{array}$$

$$\begin{array}{cccc}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
\end{array}$$

= 

$$\begin{array}{cccc}
\end{array}$$

2

$$\begin{array}{cccc}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
\end{array}$$

1 2 3

5 4
Preference

• In 1965, Hubbell discusses clique identification (sociologists’s clustering) on a relationship matrix $M$ using $1 + M + M^2 + M^3 + \cdots$

• He comes up with the equation $r = rM + v$

• $v$ is called a “border condition”

• He proposes the status index $v\sum_{n\geq0}M^n = v(1 - M)^{-1}$
PageRank

- In 1998, Page, Brin, Motwani and Winograd propose a spectral ranking for the web

- The definition is in term of an equation \( r = \alpha(rP + v) \), where \( P \) is the web graph normalized by rows

- It will evolve into \( r = \alpha rP + (1 - \alpha)v \), and then into a perturbed Markov chain \( \alpha P + (1 - \alpha)1^Tv \)

- Steady state \( r = r(\alpha P + (1 - \alpha)1^Tv) \Rightarrow r = (1 - \alpha)v(1 - \alpha P)^{-1} \)

- But \( (1 - \alpha)v(1 - \alpha P)^{-1} = (1 - \alpha)v \sum_{n \geq 0} \alpha^n P^n \)