9/14: Lecture 3, PageRank 2.

Tuesday, September 12, 2006
2:19 PM

Since the beginning of lecture was spent moving between classrooms, I didn't get to point out the first place that a measure like PageRank was used. It was in the paper:

"Power and Centrality: A Family of Measures", by Phillip Bonachich,
The American Journal of Sociology, Vol. 92, No. 5 (March 1987), pp. 1170-1182.

The family of measures introduced in this paper includes a term that is equivalent to PageRank's term that causes every page to give a weak reference to every other page.

The main purpose of this lecture is to introduce two papers that provide an axiomatic approach to the ranking problem. In both cases, they introduce a set of axioms that they would like a ranking system to have, and then show that the only ranking system that satisfies these is something like PageRank.

The first paper is "The Measurement of Intellectual Influence",
by Ignacio Palacios-Huerta and Oscal Volij,
Econometrica, Vol. 72, No. 3 (May 2004) 963-977.

The second is "The PageRank Axioms", by Alon Altman, Moshe Tennenholtz,
which appeared in the Proceedings of the Sixth ACM Conference on
Electronic Commerce.

The first paper approaches a more general problem: that of ranking journals in order of importance. It considers ranking functions that map journals into real numbers, with higher numbers indicating higher rankings. It states its axioms in an arithmetic fashion, and shows that the only ranking functions that satisfies its axioms corresponds to an eigenvector.

The second paper just focuses on unweighted directed graphs. It considers ranking functions just to be orders, and it states its axioms in a graph theoretic fashion. Each says that if the graph is transformed in some way, then the ordering should change (or not change) in some corresponding way.

In the middle of the lecture, I will observe that one could have taken an approach like this with the first paper, which would have made it far more compelling.

Now, let's examine the first paper.

I now introduce the notation that I will use. This notation is not the same as was used in the paper, but is consistent with the notation I will use in the rest of the course.

We assume that there are journals numbered 1, …, n.
We let l_i be the number of articles published by journal i.
Let $A_{i,j}$ be the number of times that journal i cites journal j.
So, the number of times that each journal cites another provides a weighted graph. We will only consider the case in which this graph is strongly connected, under the assumption that this corresponds to considering one intellectual community.

For now, we will only consider the case in which each journal publishes the same number of articles. In this case, we will need one less axiom, and our fundamental objects become journals instead of articles.

I'll tell you now that the ranking that the authors derive will be given by r, where r satisfies

$$r = r D^{-1} A,$$

Where $d_i = \sum_j A_{i,j}$ and $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ $D^{-1} = \begin{pmatrix} 1/d_1 & & 0 \\ & \ddots & \\ 0 & & 1/d_n \end{pmatrix}$

Note: from last lecture we know that this equation has a solution.

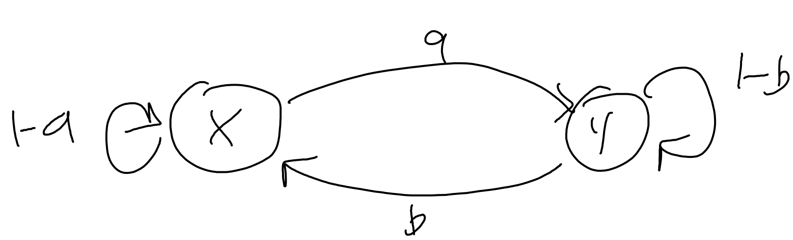
Axiom 1: if C is diagonal matrix with positive diagonals, than the ranking we obtain from A is the same as the ranking we obtain from C A. That is, if any journal multiplies all its citations through by some factor, then the ranking should not change.

Now, let's observe that this axiom is satisfied by the vector r.

Let $A' = CA$. Then, $D' = CD$, $(D')^{-1} = D^{-1}C^{-1}$.
So, $(D')^{-1}A' = D^{-1}C^{-1}CA = D^{-1}A$,
the same matrix!

The second axiom gives what the solution should be when there are just two nodes. By the first axiom, we can assume without loss of generality that $d_i = 1$, for all i.

Axiom 2: If there are two nodes, x and y, and $A_{x,y} = a$ while $A_{y,x} = b$, then $r(x) = b / (a+b)$, while $r(y) = a / (a+b)$.

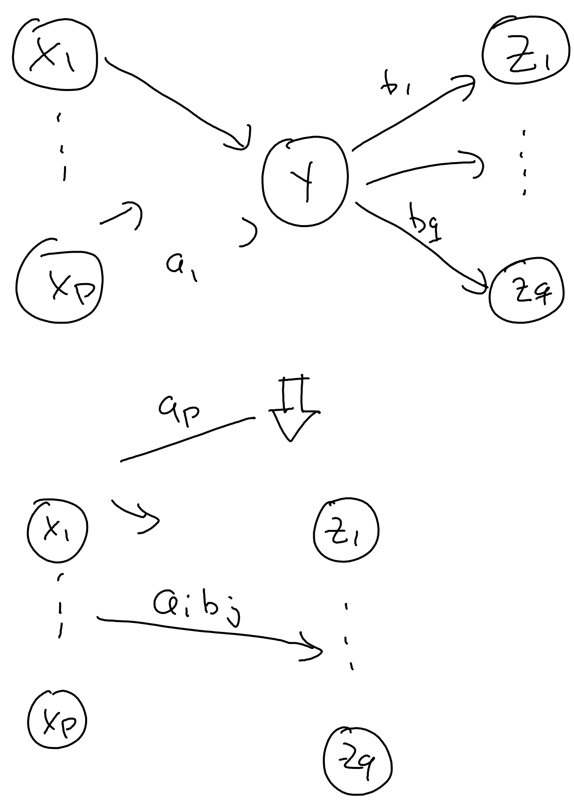


It is routine to verify that r satisfies this axiom.

The last axiom says how we should modify the graph to obtain the same rankings if we removed a vertex. That is, say there is a vertex y that we remove, and we want to add edges so that we still have the same rankings on the remaining vertices. The idea is that if x_i is a predecessor of y and z_j is a successor of y, then we should put a link from x_i to z_j whose strength is proportional to the strength of the link from x_i to y times the strength of the link from y to z_j.

Formally, assume that the sum of the weights of the edges leaving each vertex is 1, that the predecessors of y are $x_1, \ldots, x_p$, that these connect to y with edges of weights $a_1, \ldots, a_p$, and that the successors of y are $z_1, \ldots, z_q$ and that the weights of the edges from y to these are $b_1, \ldots, b_q$. Let G be the graph containing y.
We then form the graph G' that does not have y by removing y and all edges attached to it, and adding edges of weight $a_i b_j$ between each x_i and z_j. If there already is an edge between two of these nodes, then this just increases the weight of that edge by this amount. If these two nodes happen to be the same, then this creates a self-loop.

Here's the picture:



The assertion of the axiom is that if r is the ranking for G, and r' is the ranking in the modified graph G', then for all vertices u and v in G',

$$r(u) / r(v) = r'(u) / r'(v).$$

It is not difficult to observe that the vector r satisfies this condition.
To prove it, first note that, as the sum of the b_js is 1, the sum of the weights of the edges coming from each vertex is still 1. So, D' is just the restriction of D. Then, let r' be the restriction of r to the vertices other than y. To show that

$$r' = r' (D')^{-1} A',$$

it suffices to show that the equations are satisfied for the successors of y.
Consider z_1. We know that the vector r such that $r = r D^{-1} A$ satisfies

$$r'(z_1) = r(z_1) = \sum_{x:(x,z_1)\in E} r(x) A_{x,z_1}$$

$$= r(y)\, b_1 + \sum_{x\neq y} r(x) A_{x,z_1}$$

$$= r(y)\, b_1 + \sum_{x\neq y} r'(x) A_{x,z_1}$$

$$= \left(\sum_{i=1}^{p} r(x_i) a_i\right) b_1 + \sum_{x\neq y} r'(x) A_{x,z_1}$$

$$= \sum_{i=1}^{p} r'(x_i) a_i b_1 + \sum_{x\neq y} r'(x) A_{x,z_1}$$

$$= \sum_{x\neq y} r(x) A'_{x,z_1}$$

Now, since the vector r satisfies all these three axioms, we know that they are consistent. Now, it is a simple matter to show that there can only be one ranking system that satisfies all three axioms. To prove this, we can apply induction on the number of vertices. Our base case will be the case of two vertices. In this case, we have specified the ranking. If there are more vertices, say n, then by applying axiom 3, we can reduce the system to one with n-1 vertices, which by induction can be assumed to be ranked according to r. By obtaining this ranking in two different ways--that is applying axiom 3 to remove one vertex and then another, we obtain the ranking on all the vertices.

For example, if there are just three vertices x, y, and z, then by removing z and applying axiom 2, we fix r(x) / r(y). Similarly, by removing x we fix r(y)/ r(z). Combining these fixes r(x)/ r(z), and the entire ranking.

Before we move on to the second paper, let me point out that if we do allow different journals to publish different numbers of articles, then we need one more axiom. In this case, the authors suggest allowing journals to split. When we split a journal, we divide its papers between the two journals, but keep the same citation patterns (appropriately scaled for the changed number of papers). Their axiom says that, in this case, the rankings should remain unchanged.

The solution all four axioms is given by:

$$r = r L D^{-1} A L^{-1}, \text{ where } L = \begin{pmatrix} l_1 & & 0 \\ & \ddots & \\ 0 & & l_n \end{pmatrix}$$

They also show that the axioms are independent: if you drop any one, than other solutions become possible.

Before moving on to the second paper, let me point out that it is possible to make the statements in the first paper much stronger by stating the axioms in terms of total orders, rather than specifying numerical constraints on the ranking function.

Recall that an order is a relation ≤ that satisfies

a. x ≤ x, for all x,
b. for all x and y at least one of x ≤ y or y ≤ x must hold
c. if x ≤ y and y ≤ z, then x ≤ z,
d. if x ≤ y and y ≤ x, then we say x ≈ y. This does not mean they are equal, just that they have the same rank.

Axiom 1 could then be weakened to say that A and CA produce the same order.
Axiom 2 could be weakened to say that if a < b, then x ≤ y.
Axiom 3 could be weakened to say that if ≤' is the ordering in G', then for all u, v in G', u ≤' v if and only if u ≤ v.

In this way, we eliminate the numerical prescriptions upon the ranking function, but still get a set of axioms that is consistent with the vector r, and that also uniquely determine the ranking! To see that they uniquely determine the ranking, consider any two vertices x and y in G. Now, apply axiom 3 to all the other vertices. When you are done, you wind up in the 2-vertex case, and can determine from axiom 2 which vertex ranks higher in the 2-vertex graph. Then, by axiom 3, this vertex also ranked higher in G.

The second paper states its axioms just for the case of unweighted graphs, but allowing weights on self-loops.

The one problem with this paper is that it claims to recover PageRank in this way. However, there is no way to obtain PageRank this way, since it includes an edge from each vertex to every other of low weight. Rather, this paper just characterizes the vector r solving $r = r D^{-1} A$. However, I still think it is interesting.

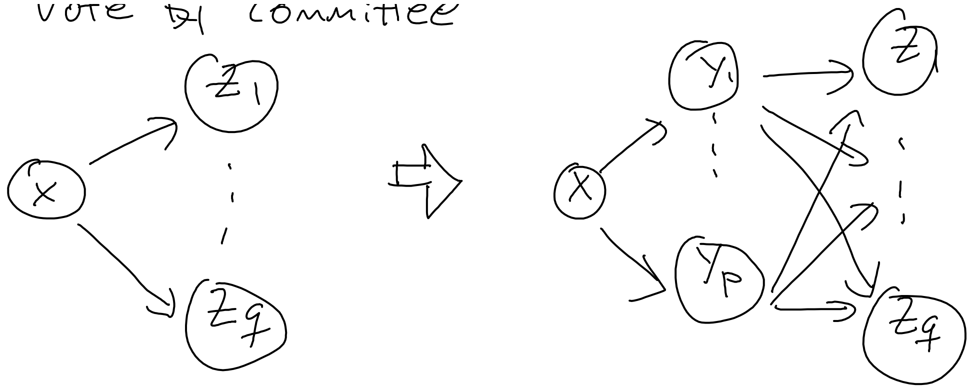
Here are their axioms.

1. The ordering is invariant under re-labeling of vertices. That is, the ordering does not depend on the names of the vertices. They mainly use this to show that if two vertices have the same set of predecessors and successors, and the same number of self-loops, then they must be equivalent in the ordering.

2. If G is a graph containing a vertex v, and G' is obtained by adding a self-loop to v, then
   a. for x, y different from v, $x \leq y$ if and only if $x \leq' y$,
   b. for x different from v, $x \leq v$ implies $v \not\leq' x$. That is, the rank of v can only increase, and if x was equivalent to v in G, then v is greater than x in G'.

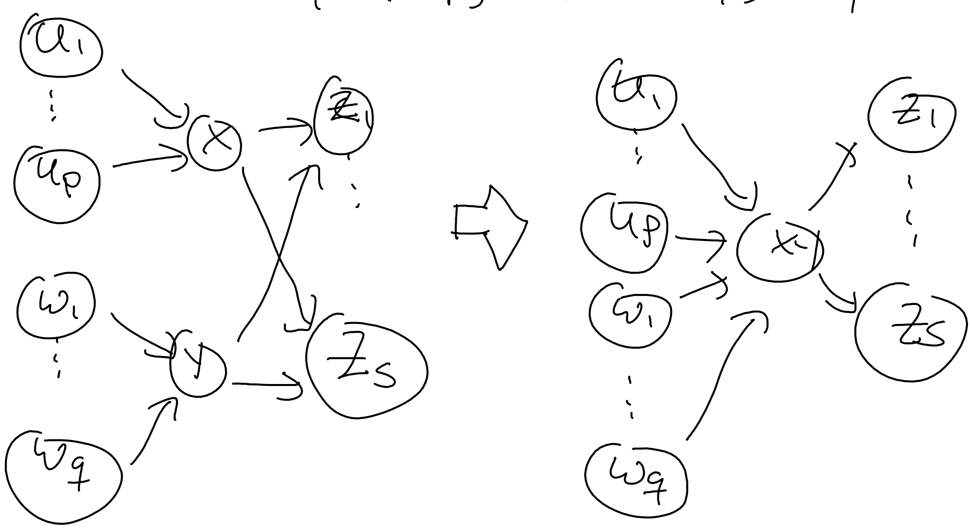
The rest of the axioms I will draw pictorially. In each case, the statement is that when you transform one graph to the next, the orderings on the vertices present in both graphs remains unchanged.

3. Vote by committee

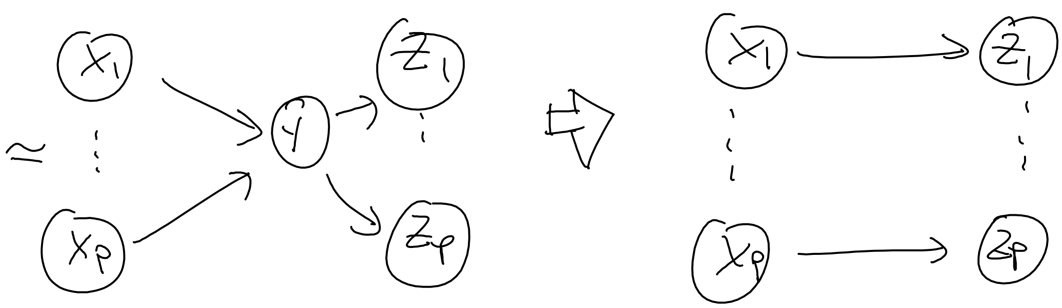


4. Collapsing:

$\{u_1, \ldots, u_p\} \cap \{w_1 \ldots w_q\} = \emptyset$

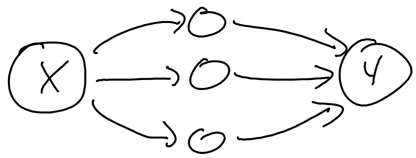


5. Vote by proxy: Here $x_1 \simeq x_2 \simeq \cdots \simeq x_p$



It is not too difficult to show that the vector r satisfies all these axioms.
To show that the vector r is the only one that satisfies all these axioms, we can try to use these axioms to derive the axioms of the previous paper. The one trick is that the only way to state those axioms is with weighted edges, while this model does not allow for weighted edges.

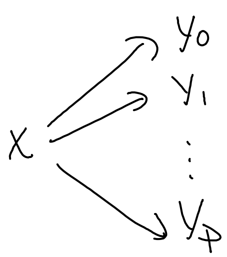
To fix this, we model weighted edges by inserting a vertex with one in-edge and one out-edge in the middle of an edge. You can have as many of these between two vertices as you like. For example, an edge of weight 3 between x and y is modeled as:



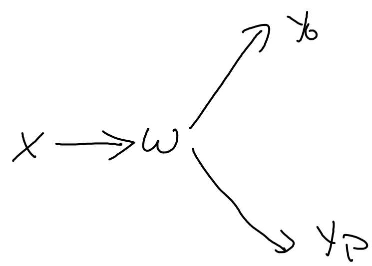
So, the first axiom that we need to derive is one that shows that ranks are unchanged by inserting one vertex inside an edge. Here is a picture of how such a proof goes.

The neighbors of x are $y_0, .., y_p$.
We will put a vertex between x and y_0

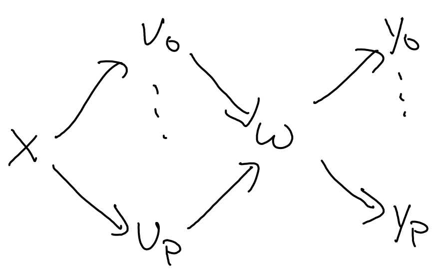
Start



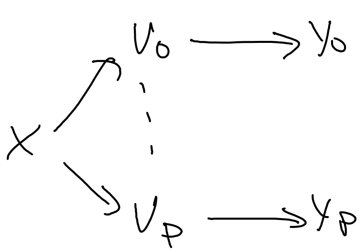
Vote by committee at x



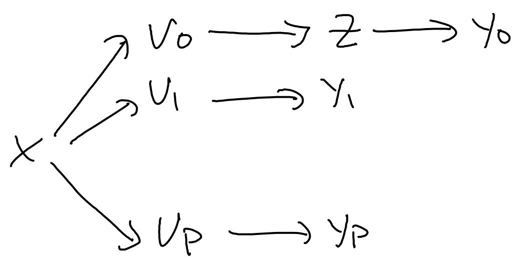
Vote by committee at x



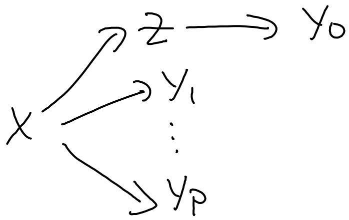
Proxy to elim w



Vote by committe at v_0



Vote by committe, backwards,
at $V_0, .., V_p$



Using similar techniques, one can model axiom 2 of the first paper, and eventually axiom 3, which is all one needs to show that ordering given by the vector r is the only solution.