6. Page-Rank Centrality; Hubs and Authorities

Bookmark this page

Exercises due Oct 20, 2021 17:29 IST Completed

Page-Rank Centrality; Hubs and Authorities

Start of transcript. Skip to the end.

PROFESSOR: OK, and now, depending on the application,

you can make this even more complicated

if you want to, just to show you how you can go--

depending on the application, you

can extend it in the ways you like and you

0:00 / 0:00 2.0x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Video note: At 2:15, the hand-written derivation has a wrong index of x and should be:

$$x^{(k+1)} = \alpha\beta AA^{\intercal}x^{(k)},$$

(instead of $x^{(k+2)}$ and similarly for y.

Page-Rank Centrality

Katz centrality and eigenvector centrality assign a relatively high importance value to a node i that has an incoming edge from a node j that is of high importance and has no other incoming edges. If node j has a very high out-degree then node i is just one of the many neighbors that node j points to. In some applications, we may require that such a node i not have very high importance simply because it has an incoming edge from a node of very high importance.

Page-rank centrality modifies Katz centrality to obtain a centrality measure that addresses this requirement. In particular, page-rank centrality weighs the contributions of all neighbors of a node by their respective out-degree values:

$$\left(\mathbf{y}^{k+1}\right)^T = \alpha\left(\mathbf{y}^k\right)^T D^{-1}A + \beta\mathbf{1}^T, \quad \text{where } D = \text{diag}\left(k_1^{\text{out}}, \ldots, k_n^{\text{out}}\right).$$

With a choice of α in the interval $\left(0, 1/\lambda_{\max}\left(D^{-1}A\right)\right)$, we can show that the recursive updates converge to $\mathbf{v}^T$, where $\mathbf{v}^T = \beta\mathbf{1}^T\left(\mathbf{I} - \alpha D^{-1}A\right)^{-1}$.

The library **networkx** does not enforce a normalization condition on the page-rank centrality vector, and so we will follow that convention here.

Katz Centrality

7/7 points (graded)

Consider the following adjacency matrix A:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

In this problem, you will compute the Katz centrality vector, and then you will compute the Page-Rank centrality vector for this same graph in the next problem.

Raw matrix

Python:

```
    [[0, 1, 1, 1, 1, 1, 1],
[1, 0, 1, 0, 0, 0, 0],
[1, 1, 0, 0, 0, 0, 0],
[1, 0, 0, 0, 1, 0, 0],
[1, 0, 0, 1, 0, 0, 0],
[1, 0, 0, 0, 0, 0, 0],
[1, 0, 0, 0, 0, 1, 0]]
```

Mathematica:

```
    {{0, 1, 1, 1, 1, 1, 1},
{1, 0, 1, 0, 0, 0, 0},
{1, 1, 0, 0, 0, 0, 0},
{1, 0, 0, 0, 1, 0, 0},
{1, 0, 0, 1, 0, 0, 0},
{1, 0, 0, 0, 0, 0, 0},
{1, 0, 0, 0, 0, 1, 0}}
```

Hide

Compute the Katz centrality of the nodes using the `networkx.katz_centrality` function in **networkx**. Use the default values of parameters ($\alpha = 0.1, \beta = 1$), and the Katz centrality normalization of $\sqrt{\sum v_i^2} = 1$. Provide an answer accurate to at least three **significant figures** (graded to 1% tolerance).

Node 0: 0.49029248 ✔ **Answer:** 0.4902

Node 1: 0.362204 ✔ **Answer:** 0.3622

Node 2: 0.362204 ✔ **Answer:** 0.3622

Node 3: 0.362204 ✔ **Answer:** 0.3622

Node 4: 0.362204 ✔ **Answer:** 0.3622

Node 5: 0.35858196 ✔ **Answer:** 0.3586

Node 6: 0.32598361 ✔ **Answer:** 0.3260

Solution:

0.4902, 0.3622, 0.3622, 0.3622, 0.3622, 0.3586, 0.3260

Python:

```
graph = networkx.from_numpy_matrix(np.array(A), create_using=networkx.DiGraph)
networkx.katz_centrality(graph, alpha=0.1, beta=1)
```

Mathematica:

```
v = KatzCentrality[AdjacencyGraph[A], 0.1, 1]
v/Sqrt[Total[v^2]]
```

Note that Mathematica does not normalize the katz centrality vector, so make sure you do so after computation.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Page-Rank Centrality

7/7 points (graded)

Use the same adjacency matrix as in the previous problem:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Compute the page-rank centrality of the nodes using the `networkx.pagerank` function in **networkx**. Use the default values of parameters ($\alpha = 0.85$). Provide an answer accurate to at least three **significant figures** (graded to 1% tolerance).

Node 0: 0.34304517 **Answer:** 0.343

Node 1: 0.12178517 **Answer:** 0.1218

Node 2: 0.12178517 **Answer:** 0.1218

Node 3: 0.12178517 **Answer:** 0.1218

Node 4: 0.12178517 **Answer:** 0.1218

Node 5: 0.09978782 **Answer:** 0.0998

Node 6: 0.07002631 **Answer:** 0.0700

Solution:

0.3430, 1: 0.1218, 2: 0.1218, 3: 0.1218, 4: 0.1218, 5: 0.0998, 6: 0.0700

Python:

```
networkx.from_numpy_matrix(np.array(A), create_using=networkx.DiGraph)
networkx.pagerank(graph, alpha=0.85)
```

Mathematica:

```
PageRankCentrality[AdjacencyGraph[mat5], 0.85]
```

Note: The first node (row 1) is a node with high centrality as every other node points to it. The last node (row 7) benefits in the case of Katz centrality by being one of the nodes that the first node is pointing to. The effect of this is seen to be diminished when we compute the page-rank centrality as the first node is pointing to every node and the last node is just one of the many.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Hubs and Authorities

We now define two interdependent notions of centrality and combine them in a mutual recursion. An important **hub** is a node that points to many important **authorities** . An important authority is one that is pointed to by many important hubs.

We begin with an initial assignment of hub and authority scores for every node $\mathbf{x}^0$ and $\left(\mathbf{y}^0\right)^T$, respectively. The updates are as follows:

$$\mathbf{x}^{k+1} = \alpha A \mathbf{y}^k, \quad \left(\mathbf{y}^{k+1}\right)^T = \beta \left(\mathbf{x}^{k+1}\right)^T A.$$

Choosing $\alpha\beta = 1/\lambda_{\max}\left(AA^T\right)$, we can show that $\mathbf{x}^k \to \mathbf{v}$ and $\left(\mathbf{y}^k\right)^T \to \mathbf{w}^T$, where $AA^T\mathbf{v} = \lambda_{\max}\left(AA^T\right)\mathbf{v}$ and $A^TA\mathbf{w} = \lambda_{\max}\left(A^TA\right)\mathbf{w}$. In fact, the non-zero eigenvalues of AA^T and A^TA are the same and $\mathbf{w} = A^T\mathbf{v}$.

Hubs and Authorities

14/14 points (graded)

networkx has an implementation of the hubs and authorities concept in the algorithm called *HITS*. This implementation enforces the normalization conditions of

$$\sum_i v_i = 1 \quad \sum_i w_i = 1$$

which we will use here.

Compute the HITS hub and authority scores of the nodes for the adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Raw matrix

Python:

```
    [[0, 1, 1, 1, 1, 1, 1],
 [1, 0, 1, 0, 0, 0, 0],
 [1, 1, 0, 0, 0, 0, 0],
 [1, 0, 0, 0, 1, 0, 0],
 [1, 0, 0, 1, 0, 0, 0],
 [1, 0, 0, 0, 0, 0, 0],
 [1, 0, 0, 0, 0, 1, 0]]
```

Mathematica:

```
    {{0, 1, 1, 1, 1, 1, 1},
 {1, 0, 1, 0, 0, 0, 0},
 {1, 1, 0, 0, 0, 0, 0},
 {1, 0, 0, 0, 1, 0, 0},
 {1, 0, 0, 1, 0, 0, 0},
 {1, 0, 0, 0, 0, 0, 0},
 {1, 0, 0, 0, 0, 1, 0}}
```

Hide

1. Hub scores. Provide an answer accurate to at least three **significant figures** (graded to 1% tolerance).

Node 0: 0.2544010357250008 **Answer:** 0.2544

Node 1: 0.13175928394061134 **Answer:** 0.1318

Node 2: 0.13175928394061134 **Answer:** 0.1318

Node 3: 0.13175928394061134 **Answer:** 0.1318

Node 4: 0.13175928394061134 **Answer:** 0.1318

Node 5: 0.08680254457194272 **Answer:** 0.0868

Node 6: 0.13175928394061134 **Answer:** 0.1318

2. Authority scores. Provide an answer accurate to at least three **significant figures** (graded to 1% tolerance).

Node 0: 0.2544010367546554 **Answer:** 0.2544

Node 1: 0.13175928394061134 **Answer:** 0.1318

Node 2: 0.13175928394061134 **Answer:** 0.1318

Node 3: 0.13175928394061134 **Answer:** 0.1318

Node 4: 0.13175928394061134 **Answer:** 0.1318

Node 5: 0.13175928394061134 **Answer:** 0.1318

Node 6: 0.08680254457194272 **Answer:** 0.0868

Solution:

- Hub scores: **Node 0: 0.2544, 1: 0.1318, 2: 0.1318, 3: 0.1318, 4: 0.1318, 5: 0.0868, 6: 0.1318.**
- Authority scores: **Node 0: 0.2544, 1: 0.1318, 2: 0.1318, 3: 0.1318, 4: 0.1318, 5: 0.1318, 6: 0.0868.**

Python:

```
networkx.from_numpy_matrix(np.array(A), create_using=networkx.DiGraph)
v,w=networkx.hits(graph)
```

Mathematica:

```
{w, v} = HITSCentrality[AdjacencyGraph[A]]
v/Total[v]
w/Total[w]
```

One could also compute the eigensystem for $AA^\intercal$ and $A^\intercal A$ and use the largest eigenvectors of each for v and w respectively (after normalization).

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Module 3: Network Analysis:Graph Centrality Measures / 6. Page-Rank Centrality; Hubs and Authorities

Add a Post

All Posts

what if the out-degree is zero

question posted 2 months ago by **zcypaul**

"page-rank centrality weighs the contributions of all neighbors of a node by their respective out-degree values" what if the out-degree is zero? then is D-1 reasonable if there exists 0 entries?

This post is visible to everyone.

Add a Response 1 response

vilgalys (Staff)
2 months ago

Note that page-rank centrality still includes the Katz centrality correction, so we avoid this 0-out-degree problem.

Add a comment

Showing all responses

Add a response:

Preview

< Previous Next >

edX

About
Affiliates
edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate

GET IT ON Google play
Download on the App Store