Page Rank in R - STA 521

Abbas Zaidi

1 Agenda

1. Setting up a matrix in R.
2. Using the eigen() function in R.
3. Using the signif() function in R to capture the number of significant digits.

2 Lab Tasks

1. Read in the file titled pageRank.RData. Use the ls() command and ensure that the link matrix myData is present in the environment.
2. Count the number of links for each page (at each row) and place these into a diagonal matrix. Call this matrix M.
3. Invert the matrix M and store the inverse into the variable MInv.
4. Based on the link matrix and the matrix M compute the Broken Rank. What is the problem here? Hint: Before performing the calculation remember to transpose your link matrix since the you need the links to go from $j \to i$
5. Create a dampening parameter d and set this to 0.85. Next, count the number of webpages present in the link matrix and store this in the variable n. Next, initialize a matrix E with dimensions equal to the link matrix that contains only 1s. Now combine these elements to compute the Page Rank vector for the link matrix. How are this different from the Broken Rank. Hint: Before performing the calculation remember to transpose your link matrix since the you need the links to go from $j \to i$.
6. Read in the data from the following source **http://www.ats.ucla.edu/stat/data/mat25.txt** into the variable myNewData. Write a function myPageRank that takes as its inputs a link matrix and a dampening parameter and outputs the Page Rank vector. Now test this on the the link matrix myNewData and the dampening parameter d. Hint: Before performing the calculation remember to transpose your link matrix since the you need the links to go from $j \to i$

3 Directions

In general for Labs, at the top of any file you are asked to submit, please list the following:

1. First Name Last Name
2. Lab Date
3. Team Member(s)

With respect to any item for which you are asked to generate any output, please provide the actual R output as a part of your solution and any explanation needed as well. For any functions/ computations that you will write, please list the following as comments before the step in R:

1. Task number and descriptions.
2. Input(s) with descriptions.
3. Outputs(s) with descriptions.
4. Function/ output summary (along with intermediate step comments).

For Lab 4, please provide the following deliverable items:

1. Please provide your solutions using Markdown as a .pdf with the following naming convention: LastName_FirstName_Solutions_Lab4.pdf.
2. Provide your .Rmd file (this **MUST** compile) for the lab using the following naming convention: LastName_FirstName_Solutions_Lab4.Rmd