ECS 230: Homework 4

Bingxi Li

March 2016

1 Proof

The transpose matrix of matrix $\boldsymbol{A}$ is $\boldsymbol{A}^T$. $\boldsymbol{A}$ has the eigenvalue $\{\lambda, p(\lambda) = det(A - \lambda I) = 0\}$.

$$\begin{aligned} 0 &= det(A - \lambda I) \\ &= det((A - \lambda I)^T) \\ &= det(A^T - (\lambda I)^T) \\ &= det(A^T - \lambda I) \end{aligned} \tag{1}$$

From Equation 1, λ is therefore also the eigenvalue of A^T.

2 Implementation of Power Method

Please check the attached code named "power_2.c". It should be compiled with command "icc -O0 -o power_2 power_2.c -mkl=sequential". Then you can perform the "power_2 web Xinit", where web file contains the nodes information in the network and Xinit consists of the initial guess of the rank of different n nodes numbering from 1 to n.

3 Reproduction of Rank of Nodes in Fig 1

```
login4.stampede(11)$ ./power_2 web Xinit
The Web Matrix A
0.000000 0.000000 1.000000 0.500000
0.333333 0.000000 0.000000 0.000000
0.333333 0.500000 0.000000 0.500000
0.333333 0.500000 0.000000 0.000000
The Initial Scores of X
0.250000
```

```
0.250000
0.250000
0.250000
Final Scores of  X
0.387097
0.129032
0.290322
0.193548
Node      Score     Rank
 1      0.387097     1
 3      0.290322     2
 4      0.193548     3
 2      0.129032     4
Eigen Value of X
0.999999
```

The normalized scores of each nodes are consistent with the results on Page 572 of Reference 1. The vector of final scores has been confirmed as the eigenvector of the column stochastic matrix **A** with eigenvalue being 1.

4 Score of Node 3 after Introducing Node 5

```
login3.stampede(25)$ ./power_2 web-4 Xinit-4
The Web Matrix A
0.000000 0.000000 0.500000 0.500000 0.000000
0.333333 0.000000 0.000000 0.000000 0.000000
0.333333 0.500000 0.000000 0.500000 1.000000
0.333333 0.500000 0.000000 0.000000 0.000000
0.000000 0.000000 0.500000 0.000000 0.000000
The Initial Scores of X
0.200000
0.200000
0.200000
0.200000
0.200000
Final Scores of  X
0.244898
0.081632
0.367346
0.122449
0.183674
Node      Score     Rank
 3      0.367346     1
 1      0.244898     2
```

```
 5      0.183674     3
 4      0.122449     4
 2      0.081632     5
Eigen Value of X
1.000000
```

By introducing the node 5, the rank of node 3 is increased from 2nd to 1st. The vector of final scores has been confirmed as the eigenvector of the new column stochastic matrix **A** given eigenvalue being 1. It means that the introduction of node 5 only affect the rank of the node but does not change the dominant eigenvalue of the column stochastic matrix.

5 Chained Network

The link-matrix with size being 4 is generated based on the relation of nodes in the chained network.

Using power method with two different initial vectors a=[1, 1, 1, 1] and b=[1, 2, 3, 4], the eigenvector of this link-matrix is estimated.

For the implementation of power method with initial vector a=[1, 1, 1, 1], the estimated eigenvector is as follows.

```
The Web Matrix A
0.000000 0.500000 0.000000 0.000000
1.000000 0.000000 0.500000 0.000000
0.000000 0.500000 0.000000 1.000000
0.000000 0.000000 0.500000 0.000000
The Initial Scores of X
1.000000
1.000000
1.000000
1.000000
Final Scores of  X
0.166666
0.333334
0.333334
0.166666
Node      Score     Rank
 2      0.333334     1
 3      0.333334     2
 1      0.166666     3
 4      0.166666     4
Eigen Value of X
0.999999
```

Therefore the power method applied to link-matrix of chained network can successfully converge to an eigenvector associated with eigenvalue 1.

However, implementation of power method with initial vector b=[1, 2, 3, 4] **fails to converge**. In the long run, the normalized X vector will alternate between [0.200000, 0.266667, 0.400000, 0.133333] and [0.133333, 0.400000, 0.266667, 0.200000].

The reason for the failed convergence of power method estimation with initial vector b=[1, 2, 3, 4] is analyzed here. The column-stochastic link-matrix **A** for chained network with size 4 is

$$\begin{bmatrix} 0 & 0.5 & 0 & 0 \\ 1 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 1 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}$$

A has four eigenvalues,

$$\lambda_1\text{=-1} \quad \lambda_2\text{=1} \quad \lambda_3\text{=-0.5} \quad \lambda_4\text{=0.5}$$

associated with eigenvectors respectively being X_1, X_2, X_3 and X_4. The vector a=[1, 1, 1, 1] is the linear combination of X_2 and X_3. As a has no projection on the X_1, whose eigenvector is -1, the multiplications of a with **A** will finally converge to the vector associated with eigenvalue 1.

The vector b is a linear combination of X_1, X_2, X_3 and X_4, namely $b = c_1 \cdot X_1 + c_2 \cdot X_2 + c_3 \cdot X_3 + c_4 \cdot X_4$, where $c_1 \cdot c_2 \cdot c_3 \cdot c_4 \neq 0$. The power method will finally make the vector b converge to an vector from subspace of X_1 and X_2. For any vector t from this subspace, we have $t = m \cdot X_1 + n \cdot X_2$ with following properties.

$$\begin{aligned} A \cdot (m \cdot X_1 + n \cdot X_2) = m \cdot \lambda_1 \cdot X_1 + n \cdot \lambda_2 \cdot X_2) = -m \cdot X_1 + n \cdot X_2 \\ A \cdot (-m \cdot X_1 + n \cdot X_2) = -m \cdot \lambda_1 \cdot X_1 + n \cdot \lambda_2 \cdot X_2) = m \cdot X_1 + n \cdot X_2 \end{aligned} \tag{2}$$

This property explains the alternation between two final vectors, [0.200000, 0.266667, 0.400000, 0.133333] and [0.133333, 0.400000, 0.266667, 0.200000] when implementing power method with initial vector being [1, 2, 3, 4] in the long run.

6 Shifted Power Method

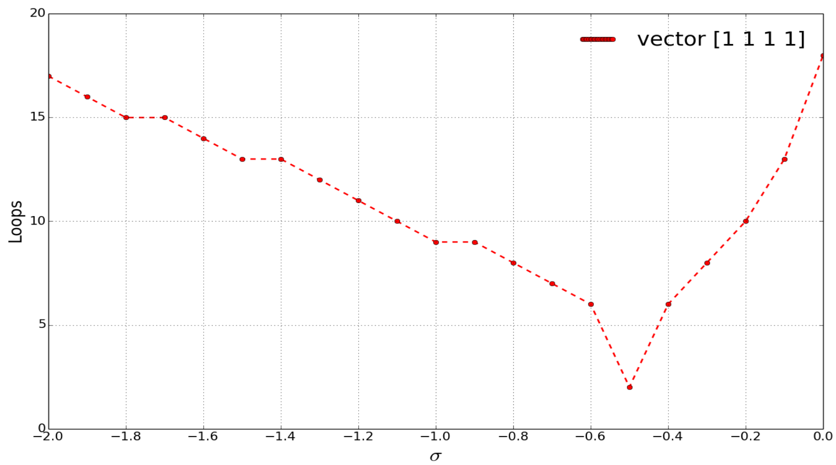
The vector [1, 2, 3, 4] has projection on the vectors associated with eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 1$. As $|\frac{\lambda_1}{\lambda_2}| = 1$, this fails the convergence of this vector.

When applying a shift σ to the linked matrix **A**, we have $A - \sigma \cdot I$. Correspondingly, the eigenvalues λ_1 and λ_2 become $\lambda_1 - \sigma$ and $\lambda_2 - \sigma$. Therefore $|\frac{\lambda_1 - \sigma}{\lambda_2 - \sigma}|$ is no longer 1 and the convergence is now possible for vector [1, 2, 3, 4]. In the case of shift $\sigma < 0$, $|\frac{\lambda_1 - \sigma}{\lambda_2 - \sigma}| < 1$, which allows the initial vector [1, 2, 3, 4] to converge to a vector associated with eigenvalue 1.

With different choices of shift, we have different convergence rates for the initial vector [1, 1, 1, 1] and [1, 2, 3, 4]. The convergence rates are quantified with the number of iterations required for convergence. In Figure 1, I plot the variations of convergence rates of the two initial vectors when changing the shifts of the linked column vector from -2 to 0.

To implements the above, please compile the attached power_6.c with **"icc -O0 -o power_6 power_6.c -mkl=sequential"**. Then input **"./power_6 web-5-1 Xinit-5-1 -0.2"**, where **web-5-1** stores the linked web for initial vector [1, 1, 1, 1] and **Xinit-5-1** is the vector input. **-0.2** is the shift σ. The number of loops used to converge the initial vector is output as **count** in the last line. Similarly, input **"./power_6 web-5-2 Xinit-5-2 -0.2"** for initial vector [1, 2, 3, 4].

```
login3.stampede(8)$ icc -O0 -o power_6 power_6.c -mkl=sequential
login3.stampede(9)$ ./power_6 web-5-1 Xinit-5-1 -0.2
The Web Matrix A
0.000000 0.500000 0.000000 0.000000
1.000000 0.000000 0.500000 0.000000
0.000000 0.500000 0.000000 1.000000
0.000000 0.000000 0.500000 0.000000
The Initial Scores of X
1.000000
1.000000
1.000000
1.000000
shifted matrix
0.200000 0.500000 0.000000 0.000000
1.000000 0.200000 0.500000 0.000000
0.000000 0.500000 0.200000 1.000000
0.000000 0.000000 0.500000 0.200000
Final Scores of  X
0.166666
0.333334
0.333334
0.166666
Node      Score     Rank
 2      0.333334     1
 3      0.333334     2
 1      0.166666     3
 4      0.166666     4
X
0.166666
0.333334
0.333334
0.166666
  sigma     eigenvalue    count
-0.200000    1.199999      10
```



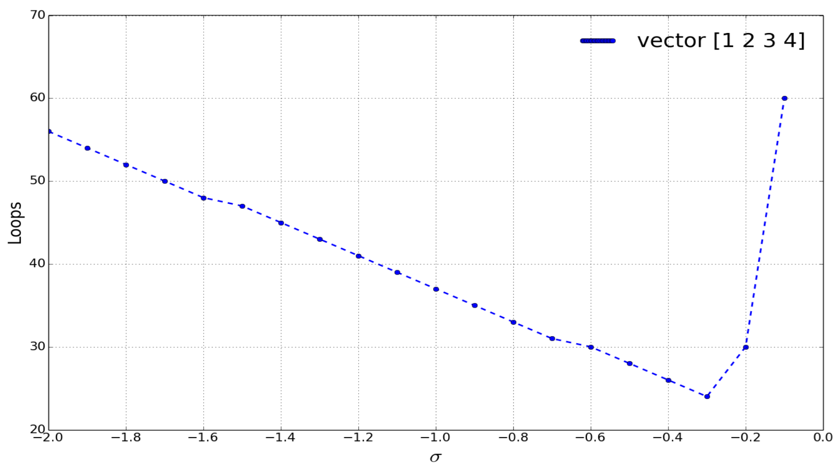


Figure 1: Convergence rates of Two Initial Vectors with Different Shifts to the Matrix.

Using the above command instructions, the loops are plotted in Figure 1 for two different initial vectors. As shown in Figure 1, the vector a=[1, 1, 1, 1] has fastest convergence when $\sigma = 0.5$. The shift is around 0.3 for the fastest convergence of b=[1, 2, 3, 4].

6.1 Theoretical Analysis

For any initial vector V being the linear combination of eigenvectors $\{X_1 \dots X_4\}$ of linked matrix, $V = c_1 \cdot X_1 + c_2 \cdot X_2 + c_3 \cdot X_3 + c_4 \cdot X_4$. The eigenvalues associated with $\{X_1 \dots X_4\}$ are $\{\lambda_1 = -1, \lambda_2 = 1, \lambda_3 = -0.5, \lambda_4 = 0.5, \}$.

The shift σ for the fastest convergence of V to a multiple of λ_2 should satisfies:

$$\sigma : \min\{\frac{\max |\lambda_i - \sigma|}{|\lambda_2 - \sigma|}, i = 1, 3, 4\} \tag{3}$$

The initial vector a only has projection on the X_2 and X_3 while b has projection on all the eigenvectors of linked matrix. The difference in initial vectors affects the value of shift σ corresponding to the minimum value of $\frac{|\lambda_i - \sigma|}{|\lambda_2 - \sigma|}$. Therefore, the shift is **-0.5** for the fastest convergence of a and **-0.25** for that of b.

7 Comparisons between Shifted Power Method and Page Rank

The following is the comparison between shifted power method and page rank in terms of convergence rate. The performance of the two algorithms is analyzed in the limit of large n for linear chain.

7.1 Convergence Rate of Shifted Power Method

The linear chain matrix of size n is an irreducible matrix with n eigenvectors. For irreducible stochastic matrix, the largest absolute value of its eigenvalues is 1. Therefore the n eigenvalues distributed between -1 and 1. Suppose the first and second largest absolute eigenvalues are $|\lambda_1|$ and $|\lambda_2|$. After shift σ, they change to $|\lambda_1 - \sigma|$ and $|\lambda_2 - \sigma|$.The convergence rate is therefore $|\frac{\lambda_2 - \sigma}{\lambda_1 - \sigma}|$. The convergence to tolerance τ will take $O(log_{|\frac{\lambda_2 - \sigma}{\lambda_1 - \sigma}|}\tau)$ time.

When the n increases to very large, the limiting case of $|\frac{\lambda_2 - \sigma}{\lambda_1 - \sigma}|$ is as follows.

$$\lim_{n \to \infty} |\frac{\lambda_2 - \sigma}{\lambda_1 - \sigma}| = 1 \tag{4}$$

Correspondingly, the time will goes to infinity.

7.2 Convergence Rate of Page Rank

The matrix M used in the Page Rank Algorithm is shown as follows,

$$M = (1 - m) \cdot A + m \cdot S \tag{5}$$

According to the paper by Kurt Bryan and Tanya Leise, the **M** has the eigenvalue λ_1 equal to 1 and second largest value λ_2 is 1-m. Therefore the convergence rate of Page Rank is $|\frac{\lambda_2}{\lambda_1}| = \frac{1-m}{1} = 1 - m$. When choosing 0.15 for m, the convergence rate is then 0.85. This is also proven to be correct in reference [1].

Based on the above conclusions, the time for Page Rank to converge within tolerance τ is then $\log_{1-m} \tau$. And it is $\log_{0.85} \tau$ for m=0.15.

7.3 Merits of the Two Methods

Through the above analysis, the merits of shifted power method and Page Rank are generalized as follows.

- The shifted power method is robust to determine the ranks of nodes in the small linear chain web. And its convergence rate can be greatly enhanced through the selection of σ, which can minimize the $|\frac{\lambda_2-\sigma}{\lambda_1-\sigma}|$. With shifted power method, we can accurately have the eigenvector of column-stochastic matrix $\mathbf{A}$ associated with eigenvalue 1.

- For a large linear chain web, the convergence of shifted power method will be super long as the convergence rate goes to 1. However the loops for Page Rank convergence only relies on the choice of m and tolerance, namely, $\log_{1-m}\tau$.

- Obviously, the larger m, the faster convergence of Page Rank. However, the final vector deviates more from the eigenvector of column-stochastic matrix $\mathbf{A}$ associated with eigenvalue 1.

References

[1] Langville, Amy N., and Carl D. Meyer. *"Deeper inside pagerank."*. Internet Mathematics 1.3 (2004): 335-380.