Matrix Theory Report

Berepalli Rajesh,

SR No:21343,

Signal Processing.

1. Let $A \in M_n$ and non-zero vectors $x, v \in \mathbb{C}^n$ be given. Suppose $c \in \mathbb{C}$, $v^H x = 1$, $Ax = \lambda x$. eigen values of A are $\lambda, \lambda_2, \ldots, \lambda_n$.

We need to show eig. values of $A(c) = cA + (1-c)\lambda x v^H$ are $\lambda, c\lambda_2, \ldots, c\lambda_n$.

<u>proof</u>:

$$(tI - cA)x = (t - c\lambda)x$$

$$\Rightarrow \det((tI-cA)x) = (\mathrm{adj}(tI-cA))((tI-cA)x)$$

$$= (t-c\lambda)\,\mathrm{adj}(tI-cA)\,x \longrightarrow (1)$$

We then have,

$$(t-c\lambda)P_{A(c)}(t) \triangleq (t-c\lambda)\det(tI-cA-(1-c)\lambda x v^H)$$

$$= (t-c\lambda)\det(tI-cA) - (t-c\lambda)(1-c)\lambda v^H \mathrm{adj}(tI-cA)x$$

$$= (t-c\lambda)\det(tI-cA) - (1-c)\lambda \det(tI-cA) v^H x$$

$$= (t-\lambda)\det(tI-cA) \qquad (\because v^H x = 1)$$

$$(t-c\lambda)P_{A(c)}(t) = (t-\lambda)\,P_{cA}(t) \longrightarrow (2)$$

where first equality follows from def'n of characteristic polynomial, the second equality from 2 partition of A (Thm 0.8.5.11 in Horn & Jhonson), the third equality from equation (1), the fourth equality from $v^H x = 1$, which is given in problem.

In eqn (2), zeros of LHS : $c\lambda$ and n eigen values of $A(c)$

zeros of RHS : λ and n eigen values of cA i.e.

$\lambda, c\lambda, c\lambda_2, \ldots, c\lambda_n$

$\Rightarrow$ Eigen values of $A(c)$ are

$\lambda, c\lambda_2, \ldots, c\lambda_n$.

<u>Note</u>: If $c = 0$, $P_{cA}(t) = \cancel{(t-\lambda)} \det(tI) = t^n$

and $P_{A(c)}(t) = (t-\lambda)\, t^{n-1}$

$\Rightarrow$ Eig. values of $A(c)$ are $\lambda, 0, \ldots, 0$.

and $c \neq 0$ case is explained above.

② If $A \in M_n$ has distinct eig. values $\lambda_1, \ldots, \lambda_n$ and there is exactly 1 eig. value λ_n of maximum modulus. If $x_0 \in \mathbb{C}^n$ is not orthogonal to left eig. vector associated with λ_n ~~i.e. $x_0^H A \neq \lambda_n x_0^H$~~ show that

$$x_{k+1} = \frac{A x_k}{(x_k^H x_k)^{1/2}} \quad \text{for } k = 0, 1, 2, \ldots$$

Converges to an eig. vector of A, and the ratios of a given non-zero entry in the vectors $A x_k$ and x_k converge to λ_n. What happens when $x_0 \in \mathbb{C}^n$ is orthogonal to a left eig. vector?

<u>proof</u>:

Since all eigen values of A are distinct, there exists 'n' independent eig. vectors $y_1, \ldots y_n$ of A.

Define $S = [\, y_1 \;\ldots\; y_n \,] \in M_n$

Since $|A| = (\lambda_1) \ldots (\lambda_n) \neq 0 \Rightarrow$ columns of A form Basis of $\mathbb{C}^n$.

We can write $x_0 = z_1 y_1 + \cdots + z_n y_n = S z$

where $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ is a vector $\in \mathbb{C}^n$.

~~Let~~ Suppose that last entry in z is zero, so that $x_0 = z_1 y_1 + \cdots + z_{n-1} y_{n-1}$. Let y denote a left eig. vector of A associated with λ_n.

(using Thm 1.4.7 : $Ax = \lambda x$, $y^H A = \mu y^H$,
@ If $\lambda \neq \mu$, then $y^H x = 0$)

we get $y^H y_i = 0$ for $i = 1, \ldots, (n-1)$.

Thus $y^H x^{(0)} = z_1 y^H y_1 + \cdots + z_{n-1} + z_n y^H y_n$

$= 0 + z_n y^H y_n = 0$ when $z_n = 0$

$\Rightarrow$ We can conclude that $z_n \neq 0$ if $x^{(0)}$ is not orthogonal to some left eigen ~~value~~ Vector associated with λ_n. $\rightarrow$ (1)

- A has n distinct eigen Values $\Rightarrow$ A is diagonalizable.

$\Rightarrow A = S \Lambda S^{-1}$

where $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$

$$x_1 = \frac{Ax_0}{\|x_0\|} = \frac{S\Lambda S^{-1} S z}{\|Sz\|} = \frac{1}{\|Sz\|} S\Lambda z$$

$$x_2 = \frac{Ax_1}{\|x_1\|} = \frac{S\Lambda S^{-1} S\Lambda z}{\|S\Lambda z\|} = \frac{S\Lambda^2 z}{\|S\Lambda z\|}$$

($\because$ Assume $\|x_0\| = 1$ without loss of generality)

$\vdots$

$$x_k = \frac{S\Lambda^k z}{\|S\Lambda^{k-1} z\|} \text{ for } k = 0, 1, 2, \ldots$$: Induction Hypothesis

Then by induction we have that

$$x_{k+1} = \frac{A x_k}{\|x_k\|} = \frac{1}{\|x_k\|} \cdot \frac{1}{\|S\Lambda^{k-1}z\|} S\Lambda S^{-1} S \Lambda^k z$$

$$= \frac{S\Lambda^{k+1} z}{\frac{\|S\Lambda^k z\|}{\cancel{\|S\Lambda^{k-1}z\|}} \cdot \cancel{\|S\Lambda^{k-1} z\|}}$$

$$= \frac{S\Lambda^{k+1} z}{\|S\Lambda^k z\|}$$

Thus Induction formula is correct, i.e. we have

$$x_k = \frac{1}{\|S\Lambda^{k-1}z\|} \cdot S\Lambda^k z \quad \text{for } k = 1, 2, \ldots$$

We can write $S\Lambda^k z = \lambda_1^k z_1 y_1 + \cdots + \lambda_n^k z_n y_n \rightarrow (1)$

Assume without loss of generality that $\lambda_i \neq 0 \ \forall i$
(if $\lambda_i = 0$, simply that i^{th} term drops out from below eq/n)

$$x_k = \frac{S\Lambda^k z}{\|S\Lambda^{k-1}z\|}$$

$$= \frac{\lambda_1^k}{\|S\Lambda^{k-1}z\|} z_1 y_1 + \cdots + \frac{\lambda_n^k}{\|S\Lambda^{k-1}z\|} z_n y_n$$

$$= \frac{\lambda_1^k}{\|S\Lambda^{k-1}z\|} z_1 y_1 + \cdots + \frac{\lambda_n^k}{\|S\Lambda^{k-1}z\|} z_n y_n$$

$$= \frac{\lambda_1^k}{|\lambda_1|^k} \left(\left\| \frac{1}{\lambda_1^k} S\Lambda^{k-1} z \right\| \right)^{-1} z_1 y_1 + \cdots + \frac{\lambda_n^k}{|\lambda_n|^k} \left(\left\| \frac{1}{\lambda_n^k} S\Lambda^{k-1} z \right\| \right)^{-1} z_n y_n$$

(2)

Note that from ① , we can write

$$\frac{1}{\lambda_i^k} S \Lambda^{k-1} z = \frac{1}{\lambda_i^k}\left(\lambda_1^{k-1} z_1 y_1 + \cdots + \lambda_n^{k-1} z_n y_n\right)$$

$$= \left(\frac{\lambda_1}{\lambda_i}\right)^k \frac{z_1 y_1}{\lambda_1} + \cdots + \left(\frac{\lambda_n}{\lambda_i}\right)^k \frac{z_n y_n}{\lambda_n}$$

λ_n is largest eig. value $\Rightarrow \frac{1}{\lambda_i}$ is large for $i = 0, 1, .. n-1$.

and $\left(\frac{\lambda_n}{\lambda_i}\right)^k$ will become arbitrarily large in magnitude

for $i = 1, \ldots, n-1$ & as $k \to \infty$, it become ∞.

Now, in eq ② , $\frac{\lambda_i^k}{|\lambda_i|^k}\left(\left\| \frac{1}{\lambda_i^k} S \Lambda^{k-1} z \right\|\right)^{-1} z_i y_i \to 0$

as $k \to \infty$ for $i = 1, \ldots, n-1$ bcz $\frac{1}{\lambda_i^k} S \Lambda^{k-1} z \to \infty$

as $k \to \infty$.

$$\Rightarrow \lim_{k\to\infty} x_k = \lim_{k\to\infty}\left[\frac{\lambda_1^k}{|\lambda_1|^k}\left\|\frac{1}{\lambda_1^k} S \Lambda^{k-1} z\right\|^{-1} z_1 y_1 + \cdots + \frac{\lambda_n^k}{|\lambda_n|^k}\left\|\frac{1}{\lambda_n^k} S \Lambda^{k-1} z\right\|^{-1} . z_n y_n\right]$$

but as $k \to \infty$ $\frac{1}{\lambda_i} S \Lambda^{k-1} z \to \infty$ for $i = 0, \ldots, n-1$

$$\Rightarrow \lim_{k\to\infty} x_k = \lim_{k\to\infty} \frac{\lambda_n^k}{|\lambda_n|^k}\left\|\frac{1}{\lambda_n^k} S \Lambda^{k-1} z\right\|^{-1} z_n y_n$$

but $\lim_{k\to\infty}\left(\frac{1}{\lambda_n^k} S \Lambda^{k-1} z\right) = \lim_{k\to\infty}\left(\frac{\lambda_1}{\lambda_n}\right)^k \frac{z_1 y_1}{\lambda_1} + \cdots + \left(\frac{\lambda_n}{\lambda_n}\right)^k \frac{z_n y_n}{\lambda_n}$

$$= \frac{z_n y_n}{\lambda_n}$$

If λ_n is real & positive, then

$$\lim_{k\to\infty} x_k = \frac{\lambda_n}{|z_n||y_n|} \cdot z_n y_n = \lambda_n \frac{z_n}{|z_n|} \cdot \frac{y_n}{|y_n|}$$

$$= \left(\frac{\lambda_n}{|y_n|}\frac{z_n}{|z_n|}\right) y_n \Rightarrow \begin{cases} \lambda_n, & \text{if } y_n z_n > 0 \\ -\lambda_n, & \text{if } y_n z_n < 0. \end{cases}$$

where y_n – eig. vector of A associated with λ_n.
z_n – scalar

$\Rightarrow$ Thus x_k converges to a non zero multiple of y_n and therefore converges to to an eig. vector of A associated with λ_n.

$\therefore \lim_{k\to\infty} |x_k| = \frac{|\lambda_n||y_n|}{|y_n|}$ or equivalently $\lim_{k\to\infty} |x_{k+1}| = |\lambda_n|$

$\Rightarrow$ Every non zero entry of $x_{k+1} = \frac{Ax_k}{\|x_k\|}$ Converges to λ_n in magnitude

i.e. $\lim_{k\to\infty} \frac{Ax_k}{\|x_k\|} = \lambda_n \frac{x_k}{\|x_k\|}$ $\left(\because \text{assuming } \lambda_n \text{ is positive}\right)$

Note: If λ_n is negative or Complex, then x_k will not converge, however each term in x_k gets close to multiple of y_n, thus each term in the sequence x_k becomes arbitrarily close to an eig. vector associated with λ_n.

- If $x_0 \in \mathbb{C}^n$ is orthogonal to left eig vector y, then from (1) $\Rightarrow z_n = 0$ ~~$\Rightarrow \lim_{k\to\infty} x_k = \frac{\lambda_n z_n y_n}{|z_n||y_n|}$~~

i.e. $\lim_{k\to\infty} x_k = \cancel{\frac{\lambda_n}{|z_n|}}\, 0 + \frac{\lambda_n}{|\lambda_n|^k} \left\| \frac{1}{\lambda_n^k} S \Lambda^{k-1} z \right\| z_n y_n$

$= 0$ $(\because \text{From Eq } (2))$

$\Rightarrow x_k$ Converges to trivial solution, when x_0 is orthogonal to left eig vector of A.

Part (B) : Modeling

3) $A = [a_{ij}]_{n\times n}$

where $a_{ij} = \begin{cases} p\, g_{ij}/c_j + \frac{(1-p)}{n} & \text{if } c_j \neq 0 \\ \frac{1}{n} & \text{if } c_j = 0 \end{cases}$

$$\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n}\left(\frac{p\, g_{ij}}{c_j} + \frac{1-p}{n}\right)$$

$$= \frac{p}{c_j}\sum_{i=1}^{n} g_{ij} + \frac{1-p}{n}\left(\sum_{i=1}^{n} 1\right)$$

$$= \frac{p}{c_j}(c_j) + \frac{1-p}{n}(n) \qquad \left(\because \sum_{i=1}^{n} g_{ij} = c_j\right)$$

$$= p + 1 - p = 1$$

$\Rightarrow$ Column sum of A is 1

$\Rightarrow$ A^T has row sum $=1$ for each row.

$\Rightarrow$ A^T is stochastic or Markov matrix.

Thm: Every eigen value λ of a markov matrix A^T satisfies $|\lambda| \leq 1$ and one eigen value of A^T is 1.

proof:

$A^T \cdot 1 = 1 \cdot 1$ b(z row sum of A^T is 1.

where $1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

$\Rightarrow$ A^T has eigen value $= 1$.

- $A^T x = \lambda x$ for eig. pair (λ, x)

Let k be chosen such that $|x_i| \le |x_k|$ $\forall$ i in $1 \le i \le n$. Then equating the k^{th} Component of each side in $A^T x = \lambda x$, we get

$$\sum_{i=1}^{n} a_{ik} x_i = \lambda x_k \qquad \text{where } a_{ij} - \text{elements of } A$$

$$\Rightarrow |\lambda x_k| = |\lambda||x_k| = \left| \sum_{i=1}^{n} a_{ik} x_i \right| \le \sum_{i=1}^{n} a_{ik} |x_i|$$

$$\le \underbrace{\sum_{i=1}^{n} a_{ik}}_{1} |x_k|$$

bcz $|x_i| \le |x_k|$ $\forall$ $i \in [n]$.
but $\sum_{i=1}^{n} a_{ik} = 1$ $\forall$ $k \in [n]$.

$$\Rightarrow |\lambda x_k| = |\lambda| |x_k| \le |x_k|$$

$$\Rightarrow |\lambda| \le 1$$

$\therefore$ Every eigen value of A^T is one eigen Value $= 1$ and all other eig. values have absolute value smaller <u>or</u> equal to 1.

Using the fact that eigen values of A & A^T are same, we get

$\Rightarrow$ Eigen Values of A ~~are~~ satisfies $|\lambda| \le 1$ and atleast one eigen value $= 1$.

④ (i) The power method :

Step1: From G, find $c_j = \sum_{i=1}^{n} g_{ij}$: out-degree of j^{th} page

$b_j = \sum_{j=1}^{n} g_{ij}$: In-degree of j^{th} page

Step2: choose p arbitrarily in $(0,1)$, $p = 0.85$ is preferred.

Find Matrix A using $a_{ij} = \begin{cases} p\dfrac{g_{ij}}{c_j} + \dfrac{1-p}{n} & \text{if } c_j \neq 0 \\ \dfrac{1}{n} & \text{if } c_j = 0 \end{cases}$

Step3: choose $x_0 = 1$ - vector of all 1's

find $x_0 = \dfrac{x_0}{\|x_0\|}$ - normalized x_0 $(\because 1^T x_0 = 1)$

Iterate by using $x_{k+1} = \dfrac{A x_k}{(x_k^H x_k)^{1/2}} = \dfrac{A x_k}{\|x_k\|}$

Iterate until the Condition $\|x_{k+1}\| - \|x_k\| < \varepsilon$

or $\dfrac{\|x_{k+1}\|}{\|x_k\|} \approx 1$ (Very close to 1).

After N iterations it will converge to the page rank vector x, that satisfies $Ax = x$.

where N is not user defined, it is calculated using some Condition, for example $|x_{k+1}| - |x_k| < 10^{-4}$ or $|x_{k+1}| = |x_k|$ upto 6 decimal places etc...

(ii) solution to a linear system

step1: From G, find c_j using $c_j = \sum_i g_{ij}$

step2: Find A using $a_{ij} = \begin{cases} \frac{p\, g_{ij}}{c_j} + \frac{1-p}{n} & \text{if } c_j \neq 0 \\ \frac{1}{n} & \text{if } c_j = 0 \end{cases}$

step3: $AX = X \Rightarrow AX - X = 0$

$\Rightarrow$ Non-trivial solution of $(A-I)X = 0$ is required solution.

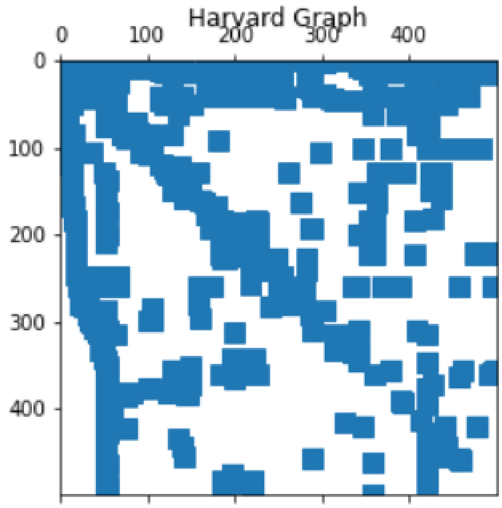
$\Rightarrow$ X = Basis $N(A-I)$ i.e. basis of nullspace $(A-I)$

or Simply x = null_space(A-eye(n)) will give us the solution vector i.e. page rank vector 'x'.

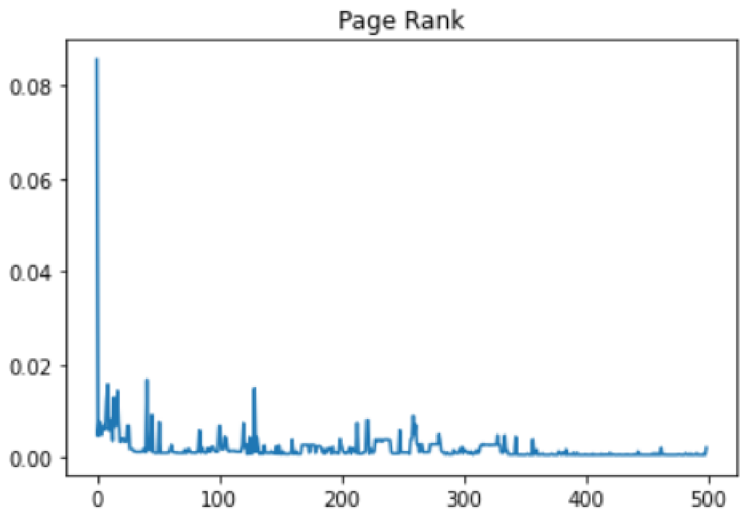
Note: we can use other methods, such as Gaussian eliminate method but null space method is simple and easier.

For p =0.85 :

Harvard Graph and its Pagerank:

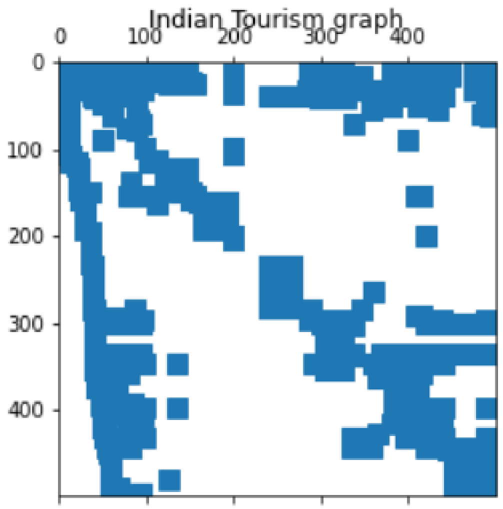


3

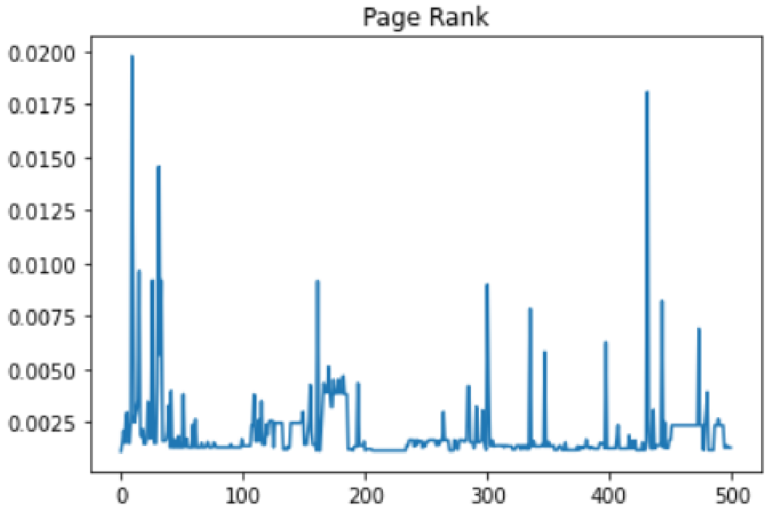


	Page-Rank	In	Out	URL
1	0.08564	195	26	[http://www.harvard.edu]
42	0.016699	42	0	[http://search.harvard.edu:8765/custom/query.h...
10	0.01568	21	18	[http://www.hbs.edu]
130	0.014853	24	12	[http://www.med.harvard.edu]
18	0.014367	45	46	[http://www.gse.harvard.edu]
15	0.012947	16	49	[http://www.hms.harvard.edu]

Indian Tourism Graph and its Pagerank:



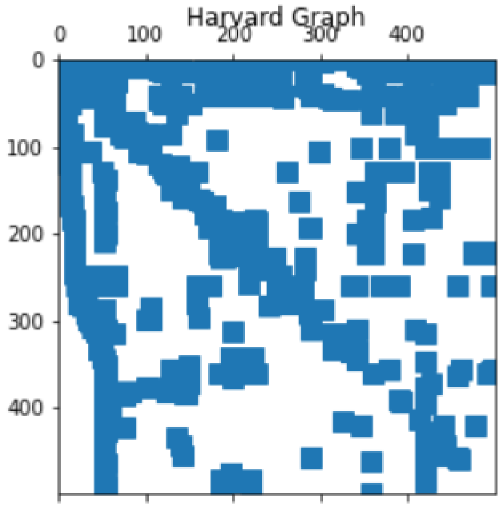
1



	Page-Rank	In	Out	URL
10	0.019776	69	2	[http://www.nic.in]
432	0.018096	7	1	[http://www.makeinindia.com]
32	0.014549	36	0	[http://india.gov.in]
11	0.010146	59	61	[http://cmf.gov.in]
16	0.009635	40	0	[http://drupal.org)]
27	0.009191	3	9	[http://nkn.gov.in/en]

For p = 0.7:

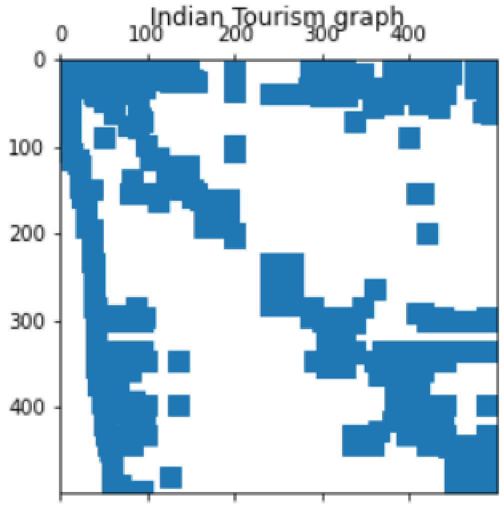
Harvard



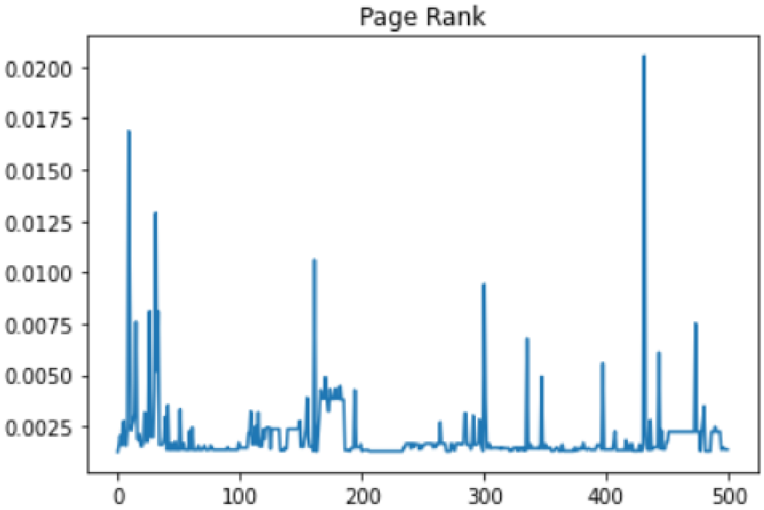
7

	Page-Rank	In	Out	URL
1	0.077561	195	26	[http://www.harvard.edu]
42	0.015022	42	0	[http://search.harvard.edu:8765/custom/query.h...
130	0.013117	24	12	[http://www.med.harvard.edu]
10	0.012927	21	18	[http://www.hbs.edu]
18	0.012432	45	46	[http://www.gse.harvard.edu]
15	0.010878	16	49	[http://www.hms.harvard.edu]

Indian Tourism:



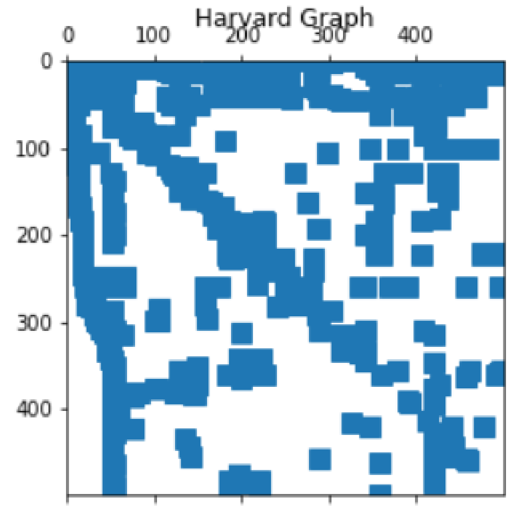
3



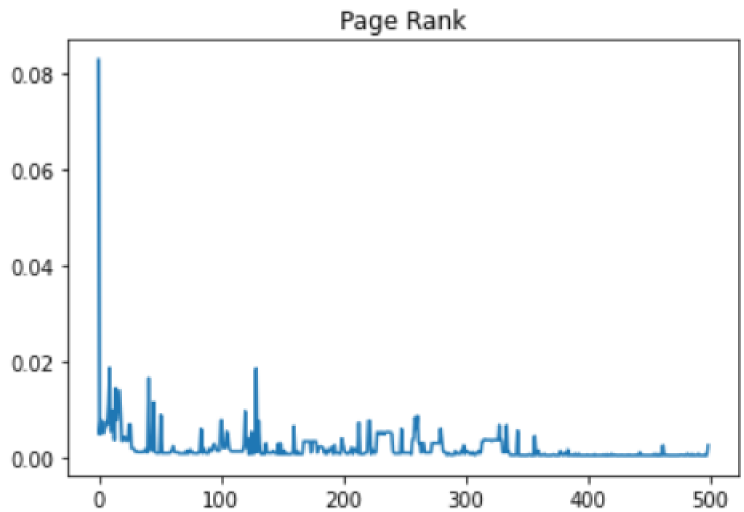
	Page-Rank	In	Out	URL
432	0.020544	7	1	[http://www.makeinindia.com]
10	0.016878	69	2	[http://www.nic.in]
32	0.012893	36	0	[http://india.gov.in]
162	0.010607	25	1	[http://subscribe.businessworld.in]
301	0.00943	22	1	[http://analytics.wrc.nic.in/cmfanalytics]
11	0.008816	59	61	[http://cmf.gov.in]

For p = 0.95

Harvard

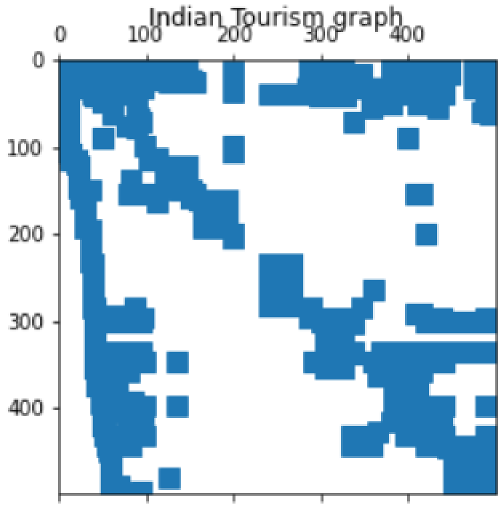


8

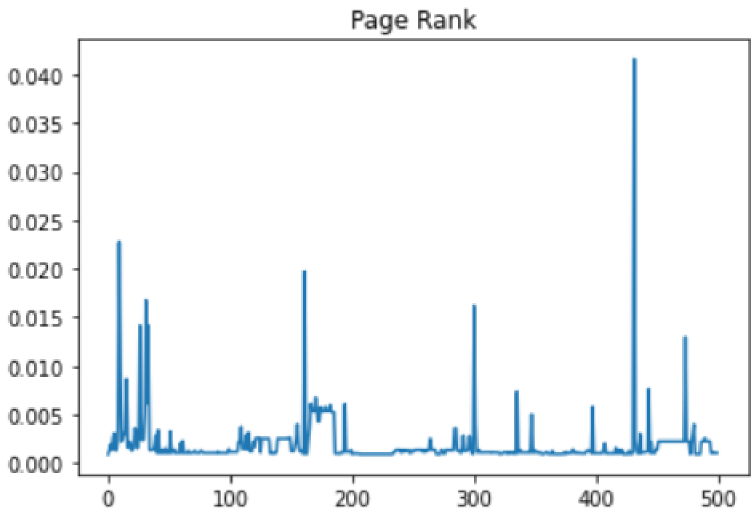


	Page-Rank	In	Out	URL
1	0.082902	195	26	[http://www.harvard.edu]
10	0.018641	21	18	[http://www.hbs.edu]
130	0.018448	24	12	[http://www.med.harvard.edu]
42	0.016473	42	0	[http://search.harvard.edu:8765/custom/query.h...
15	0.014363	16	49	[http://www.hms.harvard.edu]
18	0.013926	45	46	[http://www.gse.harvard.edu]

Indian Tourism



3



	Page-Rank	In	Out	URL
432	0.041599	7	1	[http://www.makeinindia.com]
10	0.022757	69	2	[http://www.nic.in]
162	0.019732	25	1	[http://subscribe.businessworld.in]
32	0.016807	36	0	[http://india.gov.in]
301	0.016194	22	1	[http://analytics.wrc.nic.in/cmfanalytics]
27	0.01421	3	9	[http://nkn.gov.in/en]

Comments:

We can observe that varying p value in Harvard graph ,the pageranks are varying and also the ranks of websites are changing but we are getting same websites in top 6 with shuffling in their positions.

For Indian Tourism graph, we are getting different pageranks by varying p value and this time we are getting different websites in top 6 for each p value.

From above observations ,we can infer that by varying p value, pageranks are varied, obviously graph is varied with p because we are plotting pageranks vs index of websites in U and there is no guarantee that same websites will stay in top positions if p is varied.