Università della Svizzera italiana

Institute of Computational Science ICS

Numerical Computing **2020**

Student: Claudio Maggioni Discussed with: –

Solution for Project 1

Due date: Thursday, 8 October 2020, 12:00 AM

Numerical Computing 2020 — Submission Instructions
(Please, notice that following instructions are mandatory: submissions that don't comply with, won't be considered)

- Assignments must be submitted to iCorsi (i.e. in electronic format).
- Provide both executable package and sources (e.g. C/C++ files, Matlab). If you are using libraries, please add them in the file. Sources must be organized in directories called:
  Project_number_lastname_firstname
  and the file must be called:
  project_number_lastname_firstname.zip
  project_number_lastname_firstname.pdf
- The TAs will grade your project by reviewing your project write-up, and looking at the implementation you attempted, and benchmarking your code's performance.
- You are allowed to discuss all questions with anyone you like; however: (i) your submission must list anyone you discussed problems with and (ii) you must write up your submission independently.

The purpose of this assignment[1] is to learn the importance of numerical linear algebra algorithms to solve fundamental linear algebra problems that occur in search engines.

1. Page-Rank Algorithm

1.1. Theory [20 points]

1.1.1. Show that the order of convergence of the power method is linear, and state what the asymptotic error constant is.

First of all, we show the the sequence of vectors computed by power iteration indeed converges to λ_1 or the biggest eigenvector (we assume we name eigenvectors in decreasing order of magnitude, with $|\lambda_1| > |\lambda_i|$ for $i \in 2..n$).

We can express the seed for the eigenvector (i.e. the initial value of v of the power iteration) as a linear combination of eigenvalues:

$$v_0 = \sum_{i=1}^{n} a_i x_i$$

[1] This document is originally based on a SIAM book chapter from *Numerical Computing with Matlab* from Clever B. Moler.

We can then express the result of the n-th power method as

$$v_n = \gamma A v_{n-1} = A^n v_0 = \sum_{i=1}^{n} \gamma a_i \lambda_i^n x_i = \lambda_1^n \sum_{i=1}^{n} \gamma a_i \left(\frac{\lambda_i}{\lambda_1}\right)^n x_i = \gamma a_1 \lambda_1^n x_1 + \lambda_1^n \sum_{i=2}^{n} \gamma a_i \left(\frac{\lambda_i}{\lambda_1}\right)^n x_i$$

Here, γ is just a normalization term to make $||v_n|| = 1$. v_n clearly converges to x_1 since all the terms in the $\sum_{i=2}^{n}$ contain $\frac{\lambda_i}{\lambda_1}$, which is always less than 0 if $i > 1$ for the sorting of eigenvalues we did before. Therefore, these terms to the power of n converge to 0, and γ will cancel out $a_1 \lambda_1^k$ due to the normalization, thus making the sequence converge to λ_1.

To see if the sequence converges linearly we use the definitions of rate of convergence:

$$\lim_{n\to\infty} \frac{|x_{n+1} - \lambda_1|}{|x_n - \lambda_1|^1} = \mu$$

If this limit has a finite solution then the sequence converges linearly with rate μ.

$$\lim_{n\to\infty} \frac{\left|a_1 \lambda_1^{n+1} x_1 + \lambda_1^{n+1} \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^{n+1} x_i - \beta_{n+1} x_1\right|}{\left|a_1 \lambda_1^n x_1 + \lambda_1^n \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^n x_i - \beta_n x_1\right|^1} = \mu$$

To simplify calculations, we consider the sequence without the normalization factor γ that will converge to a denormalized version of x_1, named βx_1. We can then simplify the $a_1 \lambda_1^i x_1$ terms in the sequences with $\beta_i x_1$ since β_i can be set freely.

Now we consider that if $|\lambda_2| > |\lambda_i| \; \forall i \in 3 \ldots n$ (since we sorted the eigenvalues), then $\left(\frac{\lambda_i}{\lambda_1}\right)^n$ for $i > 2$ will always converge faster to 0 than $\left(\frac{\lambda_2}{\lambda_1}\right)^n$ thus all terms other than $i = 2$ can be ignored in the limit computation. Therefore, the limit has finite solution and the convergence rate is

$$\mu = \frac{\lambda_2}{\lambda_1}$$

1.1.2. What assumptions should be made to guarantee convergence of the power method?

The first assumption to make is that the biggest eigenvalue in terms of absolute values should (let's name it λ_1) be strictly greater than all other eigenvectors, so:

$$|\lambda_1| < |\lambda_i| \forall i \in \{2..n\}$$

Also, the eigenvector *guess* from which the power iteration starts must have a component in the direction of x_i, the eigenvector for the eigenvalue λ_1 from before.

1.1.3. What is a shift and invert approach?

The shift and invert approach is a variant of the power method that may significantly increase the rate of convergence where some application of the vanilla method require large numbers of iterations. This improvement is achieved by taking the input matrix A and deriving a matrix B defined as:

$$B = (A - \alpha I)^{-1}$$

where α is an arbitrary constant that must be chosen wisely in order to increase the rate of convergence. Since the eigenvalues u_i of B can be derived from the eigenvalues λ_i of A, namely:

$$u_i = \frac{1}{\lambda_i - \alpha}$$

the rate of convergence of the power method on B is:

$$\left|\frac{u_2}{u_1}\right| = \left|\frac{\frac{1}{\lambda_2-\alpha}}{\frac{1}{\lambda_1-\alpha}}\right| = \left|\frac{\lambda_1-\alpha}{\lambda_2-\alpha}\right|$$

By choosing α close to λ_1, the convergence is sped up. To further increase the rate of convergence (up to a cubic rate), a new α, and thus a new B, may be chosen for every iteration.

1.1.4. What is the difference in cost of a single iteration of the power method, compared to the inverse iteration?

Inverse iteration is generally more expensive than a regular application of the power method, due to the overhead caused by the intermediate matrix B. One must either recompute B every time α changes, which is rather expensive due to the inverse operation in the definition of B, or one must solve the matrix equation $(A - \alpha I)v_k = v_{k-1}$ in every iteration.

1.1.5. What is a Rayleigh quotient and how can it be used for eigenvalue computations?

The Railegh quotient is an effective way to either compute the corresponding eigenvalue of an eigenvector or the corresponding eigenvalue approximation of an eigenvector approximation. I.e., if x is an eigenvector, then:

$$\lambda = \mu(x) = \frac{x^T A x}{x^T x}$$

is the corresponding eigenvalue, while if x is an eigenvector approximation, for example found through some iterations of the power method, then λ is the closest possible approximation to the corresponding eigenvalue in a least-square sense.

1.2. Other webgraphs [10 points]

The provided PageRank MATLAB implementation was run 3 times on the starting websites `http://atelier.inf.usi.ch/ maggicl`, `https://www.iisbadoni.edu.it`, and `https://www.usi.ch`, with results listed respectively in Figure 1, Figure 2 and Figure 3.

One patten that emerges on the first and third execution is the presence of 1s in the main diagonal. This indicates that several pages found have a link to themselves.

Another interesting pattern, this time observable in all executions, is the presence of contiguous rectangular regions filled with 1s, especially along the main diagonal. This may be due to the presence of pages belonging to the same website, thus having a common layout and perhaps linking to a common set of internal (when near to the main diagonal) or external pages.

Finally, we can always observe a line starting from the top-left of G and ending in its bottom-left, running along a steep path slighly going right. This may be a side effect of the way pages are discovered and numbered: if new pages are continuously discovered, these pages will be added at the end of U and a corresponding vertical strip on 1s will appear in the bottomest non-colored region of G. This continues until n unique pages are visited and the line reaches the bottom edge of the connectivity matrix. The steepness of the line thus formed depends on the amount of new pages discovered in each of the first iterations of the `surfer(...)` function.

1.3. Connectivity matrix and subcliques [10 points]

The following ETH organization are following for the near cliques along the diagonal of the connectivity matrix in `eth500.mat`. The clique approximate position on the diagonal is indicated throgh the ranges in parenthesis.

- `baug.ethz.ch` (74-100)
- `mat.ethz.ch` (114-129)
- `mavt.ethz.ch` (164-182)

- `biol.ethz.ch` (198-216)
- `chab.ethz.ch` (221-236)
- `math.ethz.ch` (264-278)
- `erdw.ethz.ch` (321-337)
- `usys.ethz.ch` (358-373)
- `mtec.ethz.ch` (396-416)
- `gess.ethz.ch` (436-462)

1.4. Connectivity matrix and disjoint subgraphs [10 points]

1.4.1. What is the connectivity matrix G (w.r.t figure 5)?

The connectivity matrix G, with U being defined as $\{"alpha", "beta", "gamma", "delta", "rho", "sigma"\}$ is:

$$G = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

1.4.2. What are the PageRanks if the hyperlink transition probability p is the default value 0.85?

First we compute the matrix A, finding:

$$A = \frac{1}{40} \begin{bmatrix} 1 & 1 & 1 & 35 & 1 & 1 \\ 18 & 1 & 1 & 1 & 1 & 1 \\ 18 & 18 & 1 & 1 & 1 & 1 \\ 1 & 18 & 35 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 35 \\ 1 & 1 & 1 & 1 & 35 & 1 \end{bmatrix}$$

We then find the eigenvectors and eigenvalues of A through MATLAB, finding that the solution of $Ax = 1x$ is:

$$x' \approx \begin{bmatrix} 0.4771 \\ 0.2630 \\ 0.3747 \\ 0.4905 \\ 0.4013 \\ 0.4013 \end{bmatrix}$$

To obtain more easily interpretable PageRank values, we can require that the sum of all PageRanks should be one. Once we normalize in this way, the result is:

$$x = \frac{x'}{\sum_{i=1}^{6} x_i} \approx \begin{bmatrix} 0.1981 \\ 0.1092 \\ 0.1556 \\ 0.2037 \\ 0.1667 \\ 0.1667 \end{bmatrix}$$

Thus the pageranks are the components of vector x, w.r.t. the order given in U.

1.4.3. Describe what happens with this example to both the definition of PageRank and the computation done by pagerank in the limit $p \to 1$.

If p is closer to 1, then the probability a web user will visit a certain page randomly decreases, thus giving more weight in the computation of PageRank to the links between one page and another.

In the computation, increasing p decreases δ (which represents the probability of a user randomly visiting a page), eventually making it 0 when p is 1.

1.5. PageRanks by solving a sparse linear system [50 points]

1.5.1. Create `pagerank1(G)` by modifying `pagerank.m` to use the power method instead of solving thesparse linear system. What is an appropriate test for terminating the power iteration?

The MATLAB solution to this question can be found on `files_data/pagerank1.m`.

An appropriate termination condition is to stop iterating when the sequence of solution vectors stops converging. This can be achieved by checking when, for an iteration n

$$||x_n - x_{n-1}|| < ||x_{n-1} - x_{n-2}||$$

stops holding.

Since this condition might be too aggressive for a resonably approximated solution, a cutoff value for the norm (say 10^{-8}) can also be introduced.

1.5.2. Create `pagerank2(G)` by modifying `pagerank.m` to use the inverse iteration. Set α equal to 0.8, 0.9 and 1 and comment on the different number of iterations they take until convergence. Also, what should be done in the unlikely event that the backslash operation involves a division by zero?

The MATLAB solution to this question can be found on `files_data/pagerank1.m`.

The termination condition for the previous exercise is used also in this implementation.

Iterations are minimal with $\alpha = 1$. For example, for the graph in Figure 5 in the assignment, we have 5 iterations for $\alpha = 1$, 8 iterations for $\alpha = 0.9$, and 11 iterations for $\alpha = 0.8$.

To avoid divisions by 0 or near-0 numbers we check the reciprocal condition number of the matrix $A - \alpha * I$. If this number is below `eps`, α must be changed. My implementation simply increases α in increments of 10^{-2} until a condition number higher than `eps` is found.

1.5.3. Use your functions `pagerank1.m` and `pagerank2.m` (set $\alpha = 1$) to compute the PageRanks of the six-node example presented in Figure 1. Make sure you get the correct result from each of your three functions.

Here are the results for `pagerank1.m`:

```
Using power method implementation
     page-rank  in  out  url
   1   0.3210    2    2  alpha
   6   0.2007    2    1  sigma
   2   0.1705    1    2  beta
   4   0.1368    2    1  delta
   3   0.1066    1    3  gamma
   5   0.0643    1    0  rho
```

And here are the results for `pagerank2.m`:

```
Using inverse iteration implementation
     page-rank  in  out  url
```

```
  1   0.3210    2    2  alpha
  6   0.2007    2    1  sigma
  2   0.1705    1    2  beta
  4   0.1368    2    1  delta
  3   0.1066    1    3  gamma
  5   0.0643    1    0  rho
```

1.5.4. Use your functions `pagerank1.m` and `pagerank2.m` (set $\alpha = 1$) to compute the PageRanks of the six-node example presented in Figure 5. Discuss the differences between the results obtained now and the ones obtained in question 4.

Here are the results for `pagerank1.m`:

```
Using power method implementation
     page-rank  in  out  url
   4   0.2037    2    1  delta
   1   0.1981    1    2  alpha
   5   0.1667    1    1  rho
   6   0.1667    1    1  sigma
   3   0.1556    2    1  gamma
   2   0.1092    1    2  beta
```

And here are the results for `pagerank2.m`:

```
Using inverse iteration implementation
     page-rank  in  out  url
   4   0.2037    2    1  delta
   1   0.1981    1    2  alpha
   5   0.1667    1    1  rho
   6   0.1667    1    1  sigma
   3   0.1556    2    1  gamma
   2   0.1092    1    2  beta
```

Both results are almost identical, with differences in the order of 10^{-8} due to the approximation introduced by the termination conditions. This results are also approximately equal to the ones obtained manually once the normalization process by the sum of components is applied.

1.5.5. Use your functions `pagerank1.m` and `pagerank2.m` (set $\alpha = 1$) to compute the PageRanks of the three selected graphs from exercise 2. Report on the convergence of e.g. the power iteration for these subgraphs and summarize the advantage of the power method implemented in `pagerank2.m` against the original implementation in `pagerank.m`.

In the following outputs, `it = <number>` represents the number of iterations required to compute the PageRanks within the termination criteria described before.

Here are the results for `pagerank1.m` for starting website `http://atelier.inf.usi.ch/~maggicl`:

```
Using power method implementation

it = 73

     page-rank  in  out  url
 360   0.0869   31    1  https://creativecommons.org/licenses/by-sa/3.0
 204   0.0406    8    1  https://forum.gitlab.com
  82   0.0189  117   18  https://www.mediawiki.org
```

```
 81    0.0188  117    4  https://wikimediafoundation.org
 87    0.0150    6    1  https://docs.gitea.io
 78    0.0145  114    9  https://www.mediawiki.org/wiki/Special:MyLanguage
                             /How_to_contribute
 77    0.0132   77   13  https://foundation.wikimedia.org/wiki/Privacy_policy
217    0.0127   40    8  https://bugs.archlinux.org
 80    0.0115  107    6  https://foundation.wikimedia.org/wiki/Cookie_statement
215    0.0114   38    5  https://bbs.archlinux.org
216    0.0114   38    8  https://wiki.archlinux.org
218    0.0114   38    6  https://security.archlinux.org
428    0.0107    9    1  https://www.dnb.de/kataloghilfe
219    0.0102   38    7  https://aur.archlinux.org
359    0.0098    9    1  https://creativecommons.org/publicdomain/zero/1.0
366    0.0092   27   21  https://archive.org
446    0.0089   24    5  https://foundation.wikimedia.org/wiki/Terms_of_Use
 83    0.0079   78    0  https:\/\/schema.org
 85    0.0074   77    0  https:\/\/www.wikimedia.org\/static\/images
                             \/wmf-hor-googpub.png
181    0.0066    8    2  https://gitlab.com
 95    0.0062    2    1  https://www.enable-javascript.com
113    0.0061   13    1  https://www.britannica.com/topic/polenta
429    0.0058    8    1  https://www.dnb.de/EN/Home/home_node.html
432    0.0058    8    1  https://www.dnb.de/expertensuche
417    0.0058    8    1  https://www.dnb.de/DE/Home/home_node.html
379    0.0057   24    2  https://blog.archive.org
213    0.0057   32    7  https://www.archlinux.org
 99    0.0056    3    1  https://www.usi.ch/it
 19    0.0051    4    1  https://creativecommons.org/licenses/by-nc-sa/4.0
214    0.0050   31    5  https://www.archlinux.org/packages
220    0.0050   31    6  https://www.archlinux.org/download
```

Here are the results for `pagerank2.m` for starting website `http://atelier.inf.usi.ch/~maggicl`:

```
Using inverse iteration implementation

it = 7

     page-rank   in  out  url
360    0.0869   31    1  https://creativecommons.org/licenses/by-sa/3.0
204    0.0406    8    1  https://forum.gitlab.com
 82    0.0189  117   18  https://www.mediawiki.org
 81    0.0188  117    4  https://wikimediafoundation.org
 87    0.0150    6    1  https://docs.gitea.io
 78    0.0145  114    9  https://www.mediawiki.org/wiki/Special:MyLanguage
                             /How_to_contribute
 77    0.0132   77   13  https://foundation.wikimedia.org/wiki/Privacy_policy
217    0.0127   40    8  https://bugs.archlinux.org
 80    0.0115  107    6  https://foundation.wikimedia.org/wiki/Cookie_statement
215    0.0114   38    5  https://bbs.archlinux.org
216    0.0114   38    8  https://wiki.archlinux.org
218    0.0114   38    6  https://security.archlinux.org
428    0.0107    9    1  https://www.dnb.de/kataloghilfe
219    0.0102   38    7  https://aur.archlinux.org
359    0.0098    9    1  https://creativecommons.org/publicdomain/zero/1.0
```

```
 366   0.0092   27   21  https://archive.org
 446   0.0089   24    5  https://foundation.wikimedia.org/wiki/Terms_of_Use
  83   0.0079   78    0  https:\/\/schema.org
  85   0.0074   77    0  https:\/\/www.wikimedia.org\/static\/images
                             \/wmf-hor-googpub.png
 181   0.0066    8    2  https://gitlab.com
  95   0.0062    2    1  https://www.enable-javascript.com
 113   0.0061   13    1  https://www.britannica.com/topic/polenta
 429   0.0058    8    1  https://www.dnb.de/EN/Home/home_node.html
 432   0.0058    8    1  https://www.dnb.de/expertensuche
 417   0.0058    8    1  https://www.dnb.de/DE/Home/home_node.html
 379   0.0057   24    2  https://blog.archive.org
 213   0.0057   32    7  https://www.archlinux.org
  99   0.0056    3    1  https://www.usi.ch/it
  19   0.0051    4    1  https://creativecommons.org/licenses/by-nc-sa/4.0
 220   0.0050   31    6  https://www.archlinux.org/download
 214   0.0050   31    5  https://www.archlinux.org/packages
```

Here are the results for `pagerank1.m` for starting website `https://www.iisbadoni.edu.it/`:

```
Using power method implementation

it = 75

     page-rank   in  out  url
 411   0.0249   42    1  https://twitter.com/mozilla
  63   0.0248  145    1  https://twitter.com/firefox
  68   0.0203  142    1  https://www.instagram.com/firefox
 412   0.0164   37    1  https://www.instagram.com/mozilla
  62   0.0080   21    1  https://github.com/mozilla/kitsune
  81   0.0070  110    2  https://www.apple.com
 384   0.0064    5    1  https://www.xfinity.com/privacy/policy/dns
   4   0.0064   32    0  https:
 377   0.0059   19    1  https://abouthome-snippets-service.readthedocs.io
                             /en/latest/data_collection.html
 393   0.0059   19    1  https://www.adjust.com/terms/privacy-policy
 410   0.0057   16    1  https://wiki.mozilla.org/Firefox/Data_Collection
 400   0.0057   15    1  https://yandex.ru/legal/confidential
 396   0.0057   15    1  https://github.com/mozilla-mobile/firefox-ios
                             /blob/master/Docs/MMA.md
   5   0.0056   31    0  https://ssl
   3   0.0054   36    0  https://www.iisbadoni.edu.it/sites/default/files
                             /favicon.ico
   6   0.0054   36    0  https://www.iisbadoni.edu.it/sites/default/files
                             /logo.png
 208   0.0054  159    0  https://schema.org
  74   0.0052  178    5  https://foundation.mozilla.org
  72   0.0052   33   32  https://www.mozilla.org/privacy/websites/#cookies
  23   0.0051    2    1  https://www.iisbadoni.edu.it/mad
 300   0.0051  157    0  https://accounts.firefox.com
```

Here are the results for `pagerank2.m` for starting website `https://www.iisbadoni.edu.it`:

```
Using inverse iteration implementation
```

```
it = 7

     page-rank  in  out  url
 411   0.0249   42    1  https://twitter.com/mozilla
  63   0.0248  145    1  https://twitter.com/firefox
  68   0.0203  142    1  https://www.instagram.com/firefox
 412   0.0164   37    1  https://www.instagram.com/mozilla
  62   0.0080   21    1  https://github.com/mozilla/kitsune
  81   0.0070  110    2  https://www.apple.com
 384   0.0064    5    1  https://www.xfinity.com/privacy/policy/dns
   4   0.0064   32    0  https:
 393   0.0059   19    1  https://www.adjust.com/terms/privacy-policy
 377   0.0059   19    1  https://abouthome-snippets-service.readthedocs.io
                             /en/latest/data_collection.html
 410   0.0057   16    1  https://wiki.mozilla.org/Firefox/Data_Collection
 400   0.0057   15    1  https://yandex.ru/legal/confidential
 396   0.0057   15    1  https://github.com/mozilla-mobile/firefox-ios
                             /blob/master/Docs/MMA.md
   5   0.0056   31    0  https://ssl
   3   0.0054   36    0  https://www.iisbadoni.edu.it/sites/default/files
                             /favicon.ico
   6   0.0054   36    0  https://www.iisbadoni.edu.it/sites/default/files
                             /logo.png
 208   0.0054  159    0  https://schema.org
  74   0.0052  178    5  https://foundation.mozilla.org
  72   0.0052   33   32  https://www.mozilla.org/privacy/websites/#cookies
  23   0.0051    2    1  https://www.iisbadoni.edu.it/mad
 300   0.0051  157    0  https://accounts.firefox.com
```

Here are the results for `pagerank1.m` for starting website `https://www.usi.ch`:

```
Using power method implementation

it = 66

     page-rank  in  out  url
  55   0.0741  354    1  https://www.instagram.com/usiuniversity
  53   0.0324  366    3  https://www.facebook.com/usiuniversity
 299   0.0248    6    1  https://twitter.com/usi_en
 329   0.0243    8    1  https://www.facebook.com/USIeLab
 308   0.0156    7    3  https://www.facebook.com/USIFinancialCommunication
  60   0.0155  316    2  https://www.swissuniversities.ch
 424   0.0144   96    1  https://it.bul.sbu.usi.ch
 330   0.0123    6    4  https://www.facebook.com/USI.ITDxC
 320   0.0122    7    1  https://www.facebook.com/usiimeg
  56   0.0107  320    0  https://www.youtube.com/usiuniversity
   5   0.0096  317   71  https://usi.ch
  62   0.0090  319   18  https://search.usi.ch
 337   0.0087    7    1  https://twitter.com/usisoftware
  63   0.0080  303   19  https://desk.usi.ch
 130   0.0077   25    0  https://www.swissuniversities.ch/it
  54   0.0072  208    0  https://twitter.com/USI_university
 323   0.0066    9    5  https://www.facebook.com/usiorientamento
 150   0.0062   12    1  https://www.innosuisse.ch/inno/it/home.html
```

```
 248   0.0061   10    1  https://www.facebook.com/usimdfc
 106   0.0060  132    8  https://newsletter.usi.ch/archive/en
 135   0.0057  201    0  https://schema.org
 326   0.0057    6    1  https://www.facebook.com/usialloggimendrisio
 322   0.0055    6    1  https://www.facebook.com/USImem
 366   0.0054    6    1  https://www.instagram.com/usi_ics_lugano
 212   0.0054   12    3  https://www.facebook.com/usimt
   7   0.0051  211   32  https://search.usi.ch/it
   6   0.0051  204    0  https://www.usi.ch/sites/all/themes/usiclean
                             /img/bollino-usi.svg
  14   0.0051  204   62  https://www.usi.ch/originalnode/342
  15   0.0051  204   57  https://www.usi.ch/originalnode/358
  16   0.0051  204   62  https://www.usi.ch/originalnode/343
  17   0.0051  204   57  https://www.usi.ch/originalnode/344
  18   0.0051  204   58  https://www.usi.ch/en/originalnode/12174
  20   0.0051  204   60  https://www.usi.ch/originalnode/349
  21   0.0051  204   62  https://www.usi.ch/originalnode/8996
  22   0.0051  204   60  https://www.usi.ch/originalnode/348
  23   0.0051  204   59  https://www.usi.ch/originalnode/351
  24   0.0051  204   58  https://www.usi.ch/originalnode/350
  25   0.0051  204   61  https://www.usi.ch/originalnode/353
  26   0.0051  204   58  https://www.usi.ch/en/originalnode/354
  27   0.0051  204   59  https://www.usi.ch/originalnode/8014
  61   0.0051  204    0  https://www.usi.ch/sites/all/themes/usiclean
                             /img/swissuniversities.svg
  57   0.0050  188    9  https://newsletter.usi.ch/archive
```

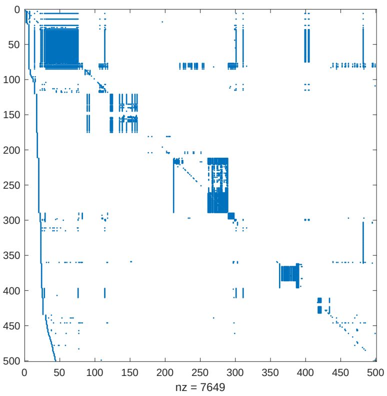
Here are the results for `pagerank2.m` for starting website `https://www.usi.ch`:

```
Using inverse iteration implementation

it = 7

     page-rank  in  out  url
  55   0.0741  354    1  https://www.instagram.com/usiuniversity
  53   0.0324  366    3  https://www.facebook.com/usiuniversity
 299   0.0248    6    1  https://twitter.com/usi_en
 329   0.0243    8    1  https://www.facebook.com/USIeLab
 308   0.0156    7    3  https://www.facebook.com/USIFinancialCommunication
  60   0.0155  316    2  https://www.swissuniversities.ch
 424   0.0144   96    1  https://it.bul.sbu.usi.ch
 330   0.0123    6    4  https://www.facebook.com/USI.ITDxC
 320   0.0122    7    1  https://www.facebook.com/usiimeg
  56   0.0107  320    0  https://www.youtube.com/usiuniversity
   5   0.0096  317   71  https://usi.ch
  62   0.0090  319   18  https://search.usi.ch
 337   0.0087    7    1  https://twitter.com/usisoftware
  63   0.0080  303   19  https://desk.usi.ch
 130   0.0077   25    0  https://www.swissuniversities.ch/it
  54   0.0072  208    0  https://twitter.com/USI_university
 323   0.0066    9    5  https://www.facebook.com/usiorientamento
 150   0.0062   12    1  https://www.innosuisse.ch/inno/it/home.html
 248   0.0061   10    1  https://www.facebook.com/usimdfc
 106   0.0060  132    8  https://newsletter.usi.ch/archive/en
```

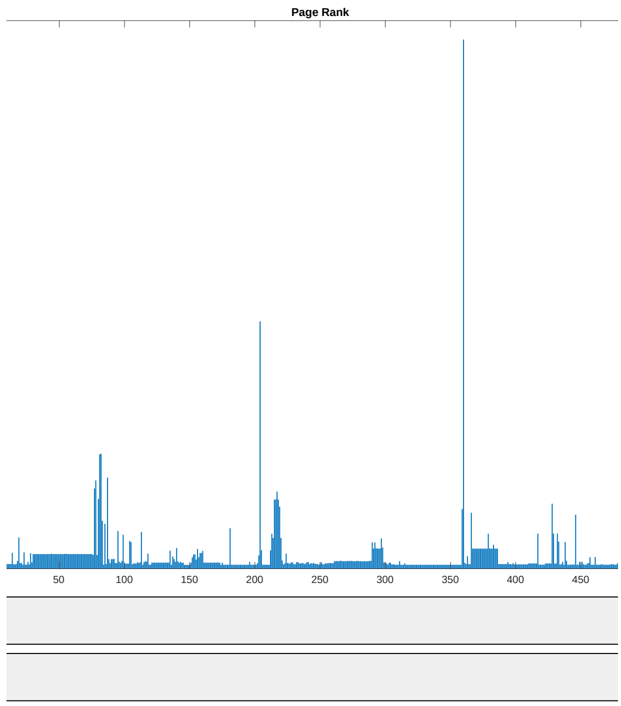
135	0.0057	201	0	https://schema.org
326	0.0057	6	1	https://www.facebook.com/usialloggimendrisio
322	0.0055	6	1	https://www.facebook.com/USImem
366	0.0054	6	1	https://www.instagram.com/usi_ics_lugano
212	0.0054	12	3	https://www.facebook.com/usimt
7	0.0051	211	32	https://search.usi.ch/it
14	0.0051	204	62	https://www.usi.ch/originalnode/342
27	0.0051	204	59	https://www.usi.ch/originalnode/8014
21	0.0051	204	62	https://www.usi.ch/originalnode/8996
22	0.0051	204	60	https://www.usi.ch/originalnode/348
25	0.0051	204	61	https://www.usi.ch/originalnode/353
61	0.0051	204	0	https://www.usi.ch/sites/all/themes/usiclean /img/swissuniversities.svg
16	0.0051	204	62	https://www.usi.ch/originalnode/343
23	0.0051	204	59	https://www.usi.ch/originalnode/351
15	0.0051	204	57	https://www.usi.ch/originalnode/358
17	0.0051	204	57	https://www.usi.ch/originalnode/344
24	0.0051	204	58	https://www.usi.ch/originalnode/350
6	0.0051	204	0	https://www.usi.ch/sites/all/themes/usiclean /img/bollino-usi.svg
20	0.0051	204	60	https://www.usi.ch/originalnode/349
26	0.0051	204	58	https://www.usi.ch/en/originalnode/354
18	0.0051	204	58	https://www.usi.ch/en/originalnode/12174
57	0.0050	188	9	https://newsletter.usi.ch/archive

The potential of the algorithm in `pagerank2.m` compared to the one in `pagerank.m` is that by using inverse iteration, thus theoretically avoiding a full matrix system solution process, the sparsity of matrix G can be maintained during the computation thus saving memory costs.

However, since my `pagerank2.m` implementation is a simplified implementation of the algorithm and indeed uses system solution (MATLAB's `mldivide`), this theoretical advantages are voided: my implementation is 3 times as long as the original one and uses approximately the same amount of memory. Advantages over `pagerank1.m` (due to the intentionally increased rate of convergence) are clear: in all the examples above, the inverse iteration implementation is at least 8 times faster than the "vanilla" power method.



(a) Spy plot of connectivity matrix

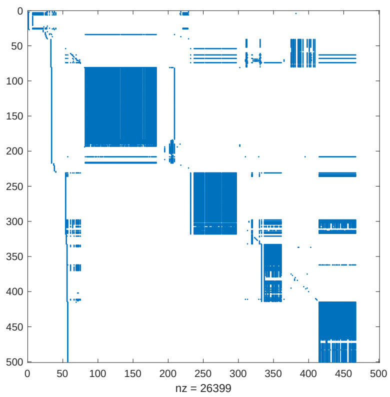


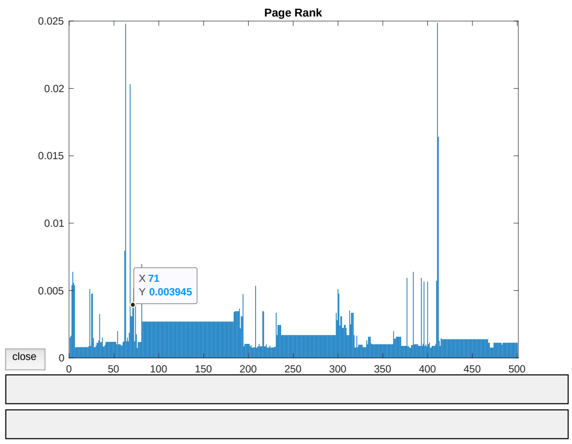
(b) Page rank bar graph

`360 0.0869 1 https://creativecommons.org/licenses/by-sa/3.0`
`204 0.0406 8 1 https://forum.gitlab.com`
`82 0.01897 https://www.mediawiki.org`
`81 0.01887 4 https://wikimediafoundation.org`
`87 0.0150 6 1 https://docs.gitea.io`
`78 0.01454 9 https://www.mediawiki.org/wiki/Special:MyLanguage/How_to_contribute`
`77 0.0132 https://foundation.wikimedia.org/wiki/Privacy_policy`
`217 0.0127 8 https://bugs.archlinux.org`
`80 0.01157 6 https://foundation.wikimedia.org/wiki/Cookie_statement`
`215 0.0114 5 https://bbs.archlinux.org`

(c) Top 10 webpages with highest PageRank

Figure 1: Results of first PageRank calculation (for starting website `http://atelier.inf.usi.ch/ maggicl/`)





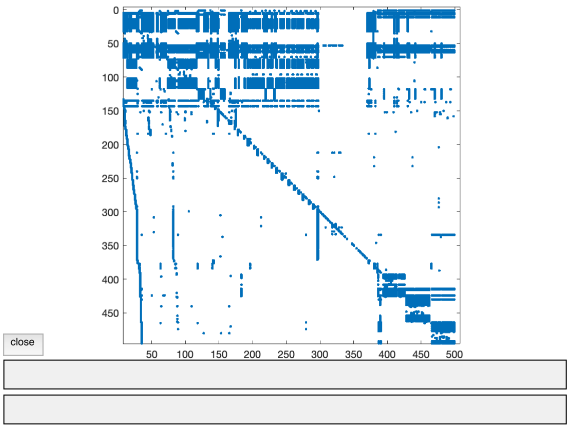
(a) Spy plot of connectivity matrix

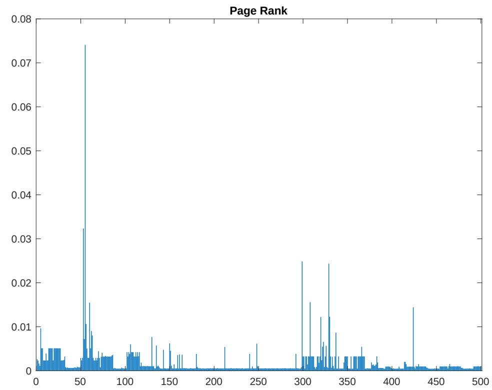
(b) Page rank bar graph

411	0.0249	1	https://twitter.com/mozilla
63	0.02485	1	https://twitter.com/firefox
68	0.02032	1	https://www.instagram.com/firefox
412	0.0164	1	https://www.instagram.com/mozilla
62	0.0080	1	https://github.com/mozilla/kitsune
81	0.00700	2	https://www.apple.com
384	0.0064	5	1 https://www.xfinity.com/privacy/policy/dns
4	0.0064	0	https:
377	0.0059	1	https://abouthome-snippets-service.readthedocs.io/en/latest/data_collection.html
393	0.0059	1	https://www.adjust.com/terms/privacy-policy
410	0.0057	1	https://wiki.mozilla.org/Firefox/Data_Collection

(c) Top 10 webpages with highest PageRank

Figure 2: Results of second PageRank calculation (for starting website https://www.iisbadoni.edu.it/)





(a) Spy plot of connectivity matrix

(b) Page rank bar graph

55	0.07414	1		https://www.instagram.com/usiuniversity
53	0.03246	3		https://www.facebook.com/usiuniversity
299	0.0248	6	1	https://twitter.com/usi_en
329	0.0243	8	1	https://www.facebook.com/USIeLab
308	0.0156	7	3	https://www.facebook.com/USIFinancialCommunication
60	0.01556	2		https://www.swissuniversities.ch
424	0.0144	1		https://it.bul.sbu.usi.ch
330	0.0123	6	4	https://www.facebook.com/USI.ITDxC
320	0.0122	7	1	https://www.facebook.com/usiimeg
56	0.01070	0		https://www.youtube.com/usiuniversity

(c) Top 10 webpages with highest PageRank

Figure 3: Results of third PageRank calculation (for starting website `https://www.usi.ch/`)