Università della Svizzera italiana

Facoltà di scienze informatiche

Informatics Bachelor Course Numerical Computing **Academic Year 2017/2018**

Instructor: Prof. Olaf Schenk TA: Fabio Verbosio, Aryan Eftekhari

---

Assignment 2 - Social Networks

Due date: Tuesday 10 October 2017, 10:30am

Name: Vanessa Braglia

Discussed with: Riccardo Gabriele, Camillo Malnati

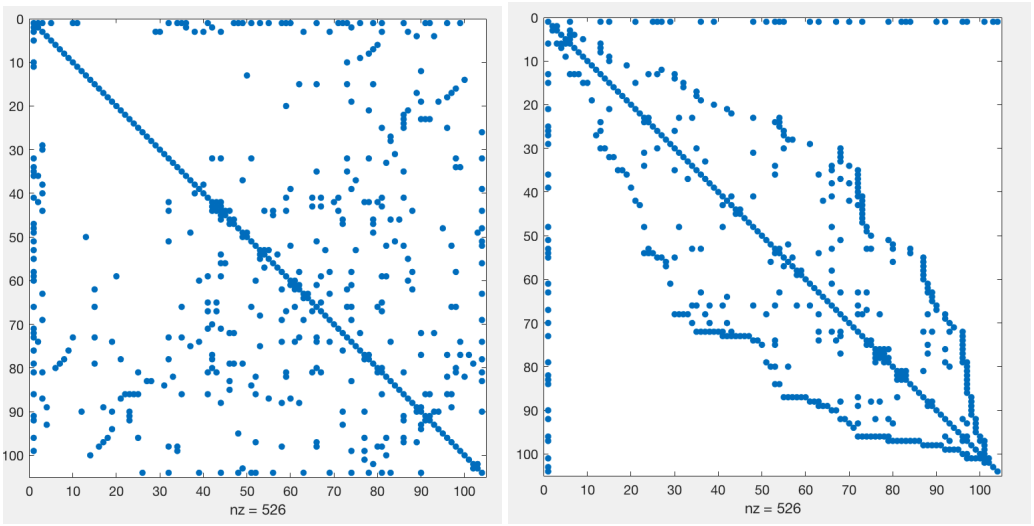
The purpose of this assignment[1] is to learn the importance of numerical linear algebra algorithms to solve fundamental linear algebra problems that occur in search engines.

Solve the following Social Networks problems:

1. The Reverse Cuthill McKee Ordering [20 points]

The first line $r = symrcm(A(2:end, 2:end))$; returns the symmetric reverse Cuthill-McKee ordering of the submatrix of A. This is a permutation matrix such that if we than look at the matrix A(r,r) we will have nonzero elements closer to the diagonal.
Looking deeper in our example, we apply the permutation only in a subset of matrix A: we want to keep the first line and the first column fixed because the first author will stay in its position at the center of the circle. Doing this we re-organized the author names such that co-authors will be as near as possible to each other. If coauhtors are next to each other, than nonzero elements of the matrix will be near the diagonal.

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



The right one represent the starting matrix, the left one the matrix after permutation.

We can see that the number of nonzero elements doesn't change since this process consists only in changing the order or rows and columns.

In the circle of author graph we will have more lines near the border of the circle and less at the center; this is because the coauthors are put ordered and put as near as possible to each other.

2. Degree Centrality [10 points]

See es2.m file

Author	Degree	Coauthors
Golub	31	Wilkinson, TChan, Varah, Overton, Ernst, VanLoan, Saunders, Bojanczyk, Dubrulle, George, Nachtigal, Kahan, Varga, Kagstrom, Widlund, OLeary, Bjorck, Eisenstat, Zha, VanDooren, Tang, Reichel, Luk, Fischer, Gutknecht, Heath, Plemmons, Berry, Sameh, Meyer, Gill
Demmel	15	Edelman, VanLoan, Bai, Schreiber, Kahan, Kagstrom, Barlow, NHigham, Arioli, Duff, Hammarling, Bunch, Heath, Greenbaum, Gragg
Plemmons	13	Golub, Nagy, Harrod, Pan, Funderlic, Bojanczyk, George, Barlow, Heath, Berry, Sameh, Meyer, Nichols
Schreiber	12	TChan, VanLoan, Moler, Gilbert, Pothen, NTrefethen, Bjorstad, NHigham, Eisenstat, Tang, Elden, Demmel
...		

3. The Connectivity of the Coauthors [10 points]

See es3.m file

```
>>  es3(A, 1, 42, name)
Golub and Moler has 2 common authors:
Wilkinson, VanLoan
```

```
>> es3(A, 1, 34, name)
Golub and Saunders has 1 common authors:
Gill
```

```
>> es3(A, 3, 104, name)
TChan and Demmel has 4 common authors:
Schreiber, Arioli, Duff, Heath
```

4. PageRankoftheCoauthorGraph[10points]

See pagerank.m file

	page-rank	in	out	author
1	0.0630	31	31	Golub
104	0.0312	15	15	Demmel
86	0.0269	13	13	Plemmons
44	0.0249	12	12	Schreiber
3	0.0236	10	10	TChan
81	0.0230	12	12	Heath
90	0.0207	9	9	Gragg
74	0.0203	10	10	Hammarling
66	0.0195	10	10	VanDooren
42	0.0171	8	8	Moler
79	0.0171	7	7	Gutknecht
32	0.0159	8	8	VanLoan
59	0.0150	7	7	Eisenstat
98	0.0147	7	7	Paige

5. Zachary's karate club: social network of friendships between 34 members [50 points]

(i) See es5_1.m file

```
>> es5_1
Node: 34          Degree: 17
Node: 1           Degree: 16
Node: 33          Degree: 12
Node: 3           Degree: 10
Node: 2           Degree: 9
```

(ii) See pagerank_5.m file

```
>> pagerank_5(karate, 0.85);
     page-rank  in  out
  34   0.1009   17   17
   1   0.0970   16   16
  33   0.0717   12   12
   3   0.0571   10   10
   2   0.0529    9    9
```

(iii) .

```
>> pagerank_5(karate, 0.85);
     page-rank  in  out
  34   0.1009   17   17
   1   0.0970   16   16
  33   0.0717   12   12
   3   0.0571   10   10
   2   0.0529    9    9
  32   0.0372    6    6
   4   0.0359    6    6
  24   0.0315    5    5
   9   0.0298    5    5
  14   0.0295    5    5
   6   0.0291    4    4
   7   0.0291    4    4
  30   0.0263    4    4
  28   0.0256    4    4
  31   0.0246    4    4
```

```
>> es5_1
Node: 34          Degree: 17
Node: 1           Degree: 16
Node: 33          Degree: 12
Node: 3           Degree: 10
Node: 2           Degree: 9
Node: 4           Degree: 6
Node: 32          Degree: 6
Node: 9           Degree: 5
Node: 14          Degree: 5
Node: 24          Degree: 5
Node: 6           Degree: 4
Node: 7           Degree: 4
Node: 8           Degree: 4
Node: 28          Degree: 4
Node: 30          Degree: 4
```

The top five nodes are the same, regardless if we use the degree or eigenvector pagerank centrality and this is because both measures for the importance of nodes. With pagerank code we can see that nodes with the same degree centrality have different eigenvector centralities so they are better ordered.

(iv) See es5_4.m file.

The spectral bisection is exactly equal to the real split observed by Zachary.

```
Group1: 17 6 7 5 11 12 1 13 22 18 4 8 2 14 20 3
Group2: 9 31 10 29 32 34 28 33 25 24 26 23 21 16 15 19 30 27
```

The matrix V contains all the eigenvector of A and we sort the authors with the second eigenvector:

$[ignore, p] = sort(V(:, 2));$

Exploring the *ignore* variable we can see that the first group contains all authors that correspond to negative values in the second eigenvector; instead the authors of the second group correspond all to positive values.