CS 370 Spring 2017: Assignment 4

Due Tuesday July 25, 5:00 PM, in the Assignment Boxes, 4th Floor MC.

Instructor: G. Labahn Office: DC3629 e-mail: *glabahn@uwaterloo.ca*
Lectures: MWF 1:30-3:30 MC2035 Office Hours: Tues 11:00-12:00

1. **(10 marks)** Find the $PA = LU$ factorization by hand using row pivoting with maximal pivot for the following matrix:
$$A = \begin{bmatrix} 2.0 & 1.0 & 4.0 \\ -4.0 & -2.0 & 6.0 \\ 1.0 & 1.0 & 2.0 \end{bmatrix}.$$
Use this factorization to solve $Ax = b$ where
$$b = \begin{bmatrix} 2.0 \\ 1.0 \\ 9.0 \end{bmatrix}.$$

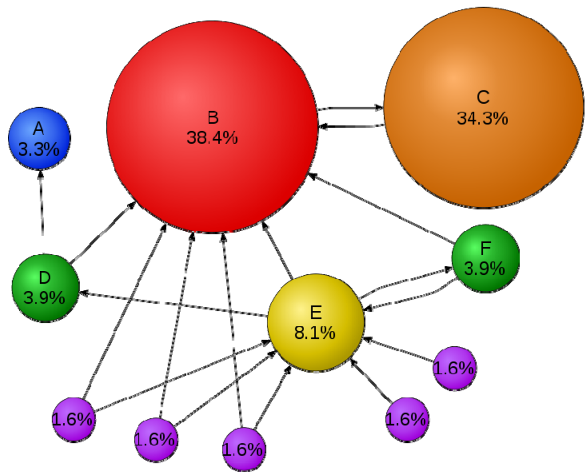
2. **(20 marks)** Computing a natural spline with interpolation points having Δx_i equal for all i requires solving a linear system $A\vec{x} = \vec{b}$ where
$$A = \begin{bmatrix} 2 & 1 & 0 & & \cdots & 0 & 0 \\ 1 & 4 & 1 & & & 0 & 0 \\ 0 & 1 & 4 & 1 & & 0 & 0 \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ 0 & 0 & 0 & & 1 & 4 & 1 \\ 0 & 0 & 0 & & & 1 & 2 \end{bmatrix}_{n \times n}$$
and $\vec{x}$ is the vector of derivatives of the spline. In this question you will investigate the solution of the linear system and determine its cost.

   (a) The factors of $A = LU$ have the form
$$L = \begin{bmatrix} 1 & 0 & & \cdots & 0 & 0 \\ l_2 & 1 & 0 & & & 0 \\ 0 & l_3 & 1 & & & \\ & & \ddots & \ddots & & \\ 0 & 0 & & & 1 & 0 \\ 0 & 0 & & & l_n & 1 \end{bmatrix} \qquad U = \begin{bmatrix} d_1 & u_1 & & \cdots & 0 & 0 \\ 0 & d_2 & u_2 & & & 0 \\ 0 & 0 & d_3 & u_3 & & \\ & & & \ddots & \ddots & \\ 0 & 0 & & & d_{n-1} & u_{n-1} \\ 0 & 0 & & & 0 & d_n \end{bmatrix}$$
   Find the recurrence equations which determine l_i, d_i and u_i.

   (b) What is the cost (that is, $O(n^k)$, for some k) of finding an LU decomposition of A? Justify your answer.

   (c) What would the cost be to compute a natural spline when all the Δx_i are the same?

3. **(20 marks)** Consider the small web given by



The probabilities displayed are the output of the pagerank algorithm for $\alpha = 0.85$. You can use these to check the correctness of your pagerank algorithm.

(a) Construct the Google matrix M for this web.

(b) Run the PageRank algorithm for 15 iterations to find a ranking vector $\vec{x}$. Use $\alpha = 0.75$.

(c) Verify that the ranking vector $\vec{x}$ approximately satisfies $M\vec{x} = \vec{x}$.

You can use either Maple or Matlab for this question.

4. **(10 marks)** A matrix $Q = [q_{ij}]$ is a positive Markov matrix if $0 < q_{ij} < 1$ and $\sum_i q_{ij} = 1$. Show that the Google matrix

$$M = \alpha(P + \frac{1}{R}\mathbf{e}\ \mathbf{d}^T) + (1-\alpha)\frac{1}{R}\mathbf{e}\ \mathbf{e}^T$$

is a positive Markov matrix.

5. **(15 marks)** Consider the positive $n \times n$ positive Markov matrix M, having n linearly independent eigenvectors $\mathbf{x}_i$, associated with eigenvalues λ_i. The PageRank algorithm uses the iteration

$$\mathbf{p}^{k+1} = M\mathbf{p}^k\ , \tag{1}$$

with

$$\mathbf{p}^\infty = \lim_{k\to\infty}(M)^k\mathbf{p}^0\ .$$

(a) Show that

$$\sum_{\ell=1}^{n}[\mathbf{p}^{k+1}]_\ell = \sum_{\ell=1}^{n}[\mathbf{p}^k]_\ell\ .$$

(b) If $\mathbf{x} = [x_1, \cdots, x_n]^T$ is the unique eigenvector of M corresponding to $\lambda = 1$, show that $\sum_{i=1}^{n} x_i \neq 0$.

6. **(25 marks)**

(a) Write a function

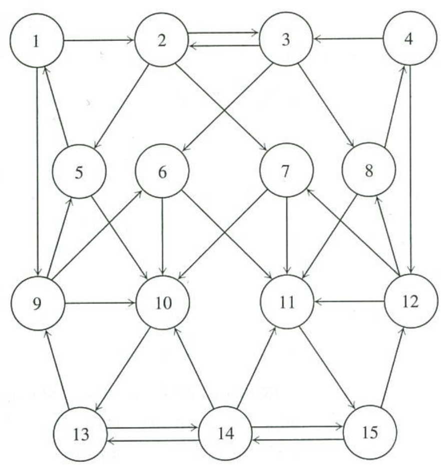
```
function [p, iter] = MyPageRank(G, alpha)
```

to compute the pagerank using the algorithm in Section 7.7 of the course notes. The inputs are the connectivity matrix G and the weight α. The output is the vector p of page ranks, and *iter* is the number of iterations that were required for the computation. Use a tolerance value of 10^{-7}.

Note: Let G_{ij} denote the (i, j)-th entry of the matrix G. $G_{ij} = 1$ indicates that there is a link from j to i. $G_{ij} = 0$ indicates no such link.

Note: You must take advantage of the sparsity of G. Avoid using additional loops (within the iteration loop) or creating full matrices (see Section 7.6 in the course notes).

(b) Run your code on the following small web using $\alpha = 0.75$:



(c) A connectivity matrix G and a list of URLs are provided in *math_uwaterloo.mat*. The data represents a network of 500 pages and was generated starting from the website `math.uwaterloo.ca`. Load the data and compute the pageranks with $\alpha = 0.85$. Use the following code to obtain the final ranking order and list the top twenty results:

```
[y I] = sort(p, 'descend');
for n = 1:min(length(I), 20)
    disp([num2str(n) ': ' U{I(n)}]);
end
```

(d) Experiment with the *math_uwaterloo.mat* using the following values of α: 0.15, 0.6, 0.75, 0.95. Report the number of iterations in each case. What do you notice about the relationship between α and the number of iterations? Explain.