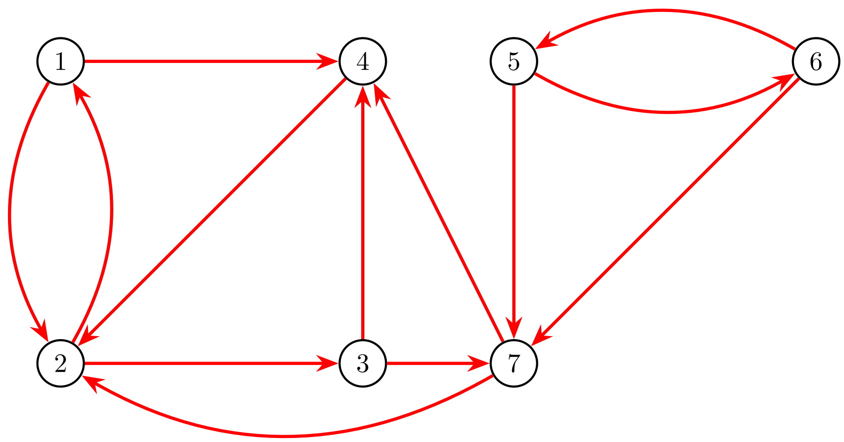
Application Activities - Module G Part 5 - Class Day 29

Activity 29.1
A $700,000,000,000 Problem:
In the picture below, each circle represents a webpage, and each arrow represents a link from one page to another.

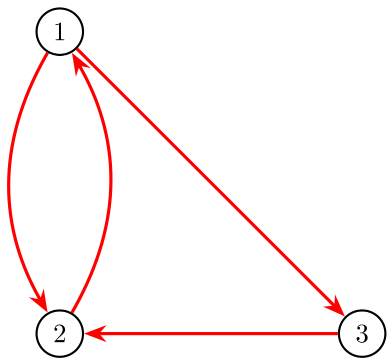


Based on how these pages link to each other, write a list of the 7 webpages in order from most imporant to least important.

Observation 29.2 The $700,000,000,000 Idea:
Links are endorsements.

1. A webpage is important if it is linked to (endorsed) by important pages.
2. A webpage distributes its importance equally among all the pages it links to (endorses).

Example 29.3 Consider this small network with only three pages. Let x_1, x_2, x_3 be the importance of the three pages respectively.

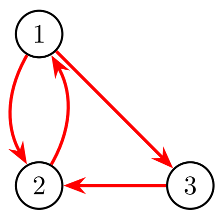


1. x_1 splits its endorsement in half between x_2 and x_3
2. x_2 sends all of its endorsement to x_1
3. x_3 sends all of its endorsement to x_2.

This corresponds to the **page rank system**

$$\begin{array}{rrl} & x_2 & = x_1 \\ \frac{1}{2}x_1 & + x_3 & = x_2 \\ \frac{1}{2}x_1 & & = x_3 \end{array}$$

Example 29.4



$$\begin{array}{rrl} & x_2 & = x_1 \\ \frac{1}{2}x_1 & + x_3 & = x_2 \\ \frac{1}{2}x_1 & & = x_3 \end{array}$$

We can summarize the left hand side of the system by putting its coefficients into a **page rank matrix** $A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}$, and store the right hand side of the system as the vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.
Thus, computing the imporance of pages on a network is equivalent to solving the matrix equation $A\mathbf{x} = \mathbf{x}$.

Activity 29.5 A **page rank vector** for a page rank matrix A is a vector $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{x}$. This vector describes the relative importance of webpages on the network described by A.
Thus, the $700,000,000,000 problem is what kind of problem?

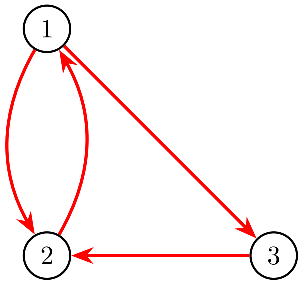
(a) A bijection problem

(b) A calculus problem

(c) A determinant problem

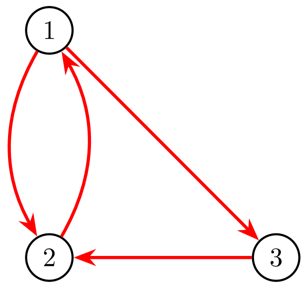
(d) An eigenvector problem

Activity 29.6 Find a page rank vector $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{x}$ (an eigenvector associated to the eigenvalue 1) for the following network's page rank matrix A.

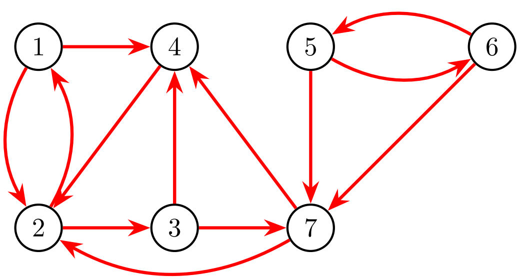


$$A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}$$

Observation 29.7 Row-reducing $A - I = \begin{bmatrix} -1 & 1 & 0 \\ \frac{1}{2} & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$ yields the basic eigenvector $\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$. Therefore, we may conclude that pages 1 and 2 are equally important, and both pages are twice as important as page 3.

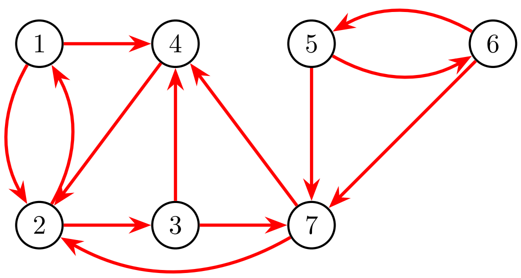


Activity 29.8 Compute the 7×7 page rank matrix for the following network.



For example, since website 1 distributes its endorsement equally between 2 and 4, the first column is $\begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

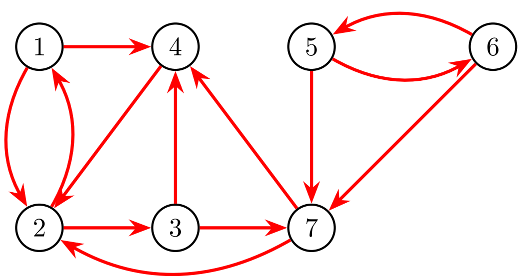
Activity 29.9 Find a page rank vector for the transition matrix.



$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Which webpage is most important?

Observation 29.10 Since a page rank vector for the network is given by $\mathbf{x}$, it's reasonable to consider page 2 as the most important page.



$$\mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ 2 \\ 2.5 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$