HIROSHIMA UNIVERSITY 広島大学

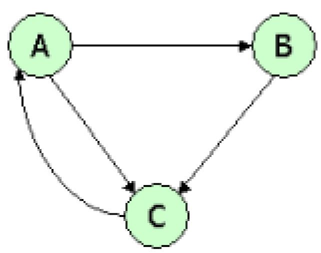
課題 3 PageRank (Homework 3)

Big Data KA218001
ビッグデータ KA218001

Submission Information

Date	Student ID	Name
21/1/2025 (DD/MM/YYYY)	C240424	Yousef Ibrahim Gomaa Mahmoud Mabrouk

答え:



Step 1. Construct the transition matrix M for the illustrated graph:

- where matrix M is stochastic (elements per column add up to 1),
- there are n column vectors where n is the number of pages,
- and each element j in each column i represents the probability of the surfer being at page j in the next time step. (page does not out-link to itself ~spider trap)

$$M = \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix} \end{array}$$

Step 2. Calculate the limiting distribution $v_i = Mv_{i-1}$*:* (v_i is the principal eigenvector)

- The graph is strongly connected.
- There are no dead ends. (and no teleporting ꞵ)
- Assuming initial probability of a surfer being at any of these pages (1/n), then:

$$v_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

$$\therefore v_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/6 \\ 1/2 \end{bmatrix},$$

  - Multiplying v over multiple steps by M recursively:

$$\therefore v = \begin{bmatrix} 1/2 \\ 1/6 \\ 1/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 1/4 \\ 0.4167 \end{bmatrix}, \begin{bmatrix} 0.4167 \\ 1/6 \\ 0.4167 \end{bmatrix}, \begin{bmatrix} 0.4167 \\ 0.2083 \\ 0.375 \end{bmatrix}, \dots, \begin{bmatrix} 0.4062 \\ 0.1979 \\ 0.3958 \end{bmatrix} \approx \begin{bmatrix} 2/5 \\ 1/5 \\ 2/5 \end{bmatrix}$$

  - Or using the power iteration method:

$$v_0 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$\therefore v = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$\because a + b + c = 1$$

$$\& \; c = a, \qquad \frac{1}{2}a = b, \qquad \frac{1}{2}a + b = c$$

    - Solving:

$$\therefore a = \frac{2}{5}, \qquad b = \frac{1}{5}, \qquad c = \frac{2}{5}$$

$$v = \begin{bmatrix} 2/5 \\ 1/5 \\ 2/5 \end{bmatrix}$$

  - Finally, vector v carries the values of final PageRank ~importance for each corresponding page.