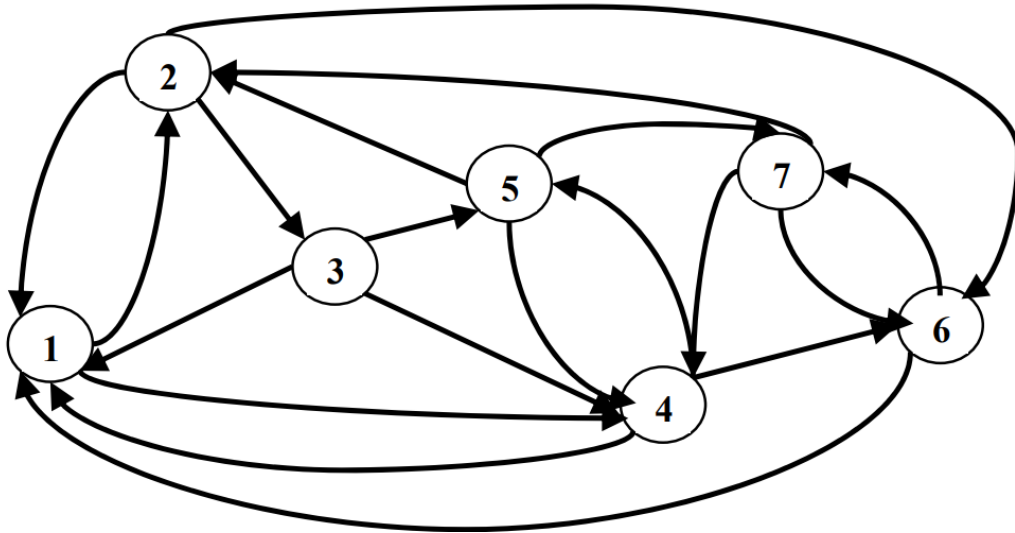


データサイエンス特論 第4回

203331 塩根風人



(1) 資料(p54 あたり)を参考に、Web ページ間の隣接行列(Adjacency Matrix) A を求めよ。
ただし、①,②,...,⑦の順で、行と列が並んでいるとする。

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

(2) Web ページ①,...,⑦までのそれぞれの PageRank 値を求めよ。(資料 pp.59-72 参照)

WebID	1	4	2	6	7	5	3
PageRank	0.222	0.194	0.175	0.159	0.108	0.084	0.058

(3) 上記の Web ページ①,...,⑦に対して HITS アルゴリズムを適用したときのオーソリティ(Authority)値が最大のページ、ならびにハブ(Hub)値が最大のページを求めよ。(資料 p.77-86 参照)

Authority Id: [4 1 6 2 5 7 3]

Authority: [0.550 0.482 0.406 0.398 0.287 0.214 0.115]

Hub Id: [7 3 4 5 2 1 6]

Hub: [0.459 0.447 0.398 0.394 0.340 0.322 0.236]

よって、Authority値が最大のページは4ページ、Hub値が最大のページは7ページである。

