Problem 1. In the following directed graph $G = (V, E)$, every node $v \in V$ represents a webpage, and every edge $(v_1, v_2) \in E$ represents a hyperlink from $v_1$ to $v_2$. Consider the “Google random surfing” model with parameter $\alpha = 1/2$. Recall that the model can be regarded as a random walk on a complete graph, where each edge is attached a transition probability. Show this complete graph and give all the transition probabilities.

Solution.

Problem 2. Compute the exact page rank of every node in problem 1.

Solution. Let $M$ be the matrix describing the random walk on the above graph $G$. We know from the solution of Problem 1:

$$M = \begin{pmatrix}
0.125 & 0.375 & 0.375 & 0.125 \\
0.125 & 0.125 & 0.375 & 0.375 \\
0.125 & 0.125 & 0.125 & 0.625 \\
0.125 & 0.125 & 0.125 & 0.625
\end{pmatrix}$$

It is guaranteed that $M^T$ has an eigenvalue 1. $(0.125, 0.1563, 0.1953, 0.5234)^T$ is the only eigenvector of this eigenvalue satisfying the condition that all the components of the eigenvector sum up to 1. The components in this vector are the page ranks of the vertices $v_1, v_2, v_3, v_4$, respectively.
Problem 3. Define \( r_i \) as the page rank of \( v_i \) in problem 2; and let \( P = (r_1, r_2, r_3, r_4)^T \). Use the power method to compute an approximate page rank for every node by answering the following problems:

1. Recall that the power method is applied on a graph \( G' \) which may contain additional edges compared to the original graph \( G \) (that is, \( G' \) is the graph on which you can apply the equation of Slide 7 in our lecture notes).

2. Show all the steps of the power method until \( \text{Err}(t) \leq 0.01 \) (see the definition of \( \text{Err}(t) \) in Slide 23 of our notes).

Solution.

1. In \( G \), vertex \( v_4 \) has no out-going edges. Before performing the power method, we need to give it an out-going edge pointing to itself (i.e., a “self-loop”). The augmented graph \( G' \) is this:

2. Let \( p(v, t) \) be the approximate page rank of vertex \( v \) at \( t \)-th round.

Initially, \( t = 0 \). Set \( p(v_1, 0) = 1, p(v_2, 0) = p(v_3, 0) = p(v_4, 0) = 0 \). In each iteration, use the equation of Slide 8 of our notes to calculate \( p(v, t) \) for all vertices \( v \).

Iteration 1. We have

\[
\begin{align*}
p(v_1, 1) &= \frac{1 - \alpha}{4} = \frac{1 - \frac{1}{2}}{4} = \frac{1}{8} = 0.125 \\
p(v_2, 1) &= \frac{1 - \alpha}{4} + \alpha \cdot \frac{p(v_1, 0)}{d^+(v_1)} = \frac{1}{8} + \frac{1}{2} \cdot \frac{1}{2} = 0.375 \\
p(v_3, 1) &= \frac{1 - \alpha}{4} + \alpha \left( \frac{p(v_1, 0)}{d^+(v_1)} + \frac{p(v_2, 0)}{d^+(v_2)} \right) = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{1}{2} + \frac{0}{2} \right) = 0.375 \\
p(v_4, 1) &= \frac{1 - \alpha}{4} + \alpha \left( \frac{p(v_2, 0)}{d^+(v_2)} + \frac{p(v_3, 0)}{d^+(v_3)} + \frac{p(v_4, 0)}{d^+(v_4)} \right) = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{0}{2} + \frac{0}{1} + \frac{0}{1} \right) = 0.125 \\
\text{Err}(1) &= |0.125 - 0.125| + |0.1563 - 0.375| + |0.1953 - 0.375| + |0.5234 - 0.125| \approx 0.7969.
\end{align*}
\]
Iteration 2. Similarly, we get

\[ p_{(v_1, 2)} = \frac{1}{8} = 0.125, \]

\[ p_{(v_2, 2)} = \frac{1}{8} + \frac{1}{2} \cdot \frac{0.125}{2} \approx 0.1563, \]

\[ p_{(v_3, 2)} = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{0.125}{2} + \frac{0.375}{2} \right) = 0.25, \]

\[ p_{(v_4, 2)} = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{0.375}{2} + \frac{0.375}{1} + \frac{0.125}{1} \right) \approx 0.4688. \]

\[ \text{Err}(2) = |0.125 - 0.125| + |0.1563 - 0.1563| + |0.1953 - 0.25| + |0.5234 - 0.4688| \approx 0.1094. \]

Iteration 3.

\[ p_{(v_1, 3)} = \frac{1}{8} = 0.125, \]

\[ p_{(v_2, 3)} = \frac{1}{8} + \frac{1}{2} \cdot \frac{0.125}{2} \approx 0.1563, \]

\[ p_{(v_3, 3)} = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{0.125}{2} + \frac{0.1563}{2} \right) \approx 0.1953, \]

\[ p_{(v_4, 3)} = \frac{1}{8} + \frac{1}{2} \cdot \left( \frac{0.1563}{2} + \frac{0.25}{1} + \frac{0.4688}{1} \right) \approx 0.5234. \]

\[ \text{Err}(3) = |0.125 - 0.125| + |0.1563 - 0.1563| + |0.1953 - 0.1953| + |0.5234 - 0.5234| \leq 0.01. \]

Problem 4. Consider a new definition similar to \( \text{Err}(t) \):

\[ \text{Err}'(t) = \max_{i=1}^{n} |r_i - P(v_i, t)| \]

where the meanings of \( r_i \) and \( P(v_i, t) \) are the same as in Slide 23 of the lecture notes. Prove that, for any \( 0 < \epsilon \leq 1 \), the power method ensures \( \text{Err}'(t) \leq \epsilon \) after \( t = O(\log \frac{1}{\epsilon}) \) rounds.

Solution. In the lecture, we have proved

\[ \text{Err}(t) \leq \alpha \cdot \text{Err}(t - 1). \] (1)

Also:

\[ \text{Err}(0) = \sum_{i=1}^{n} |r_i - P(v_i, 0)| \leq \sum_{i=1}^{n} (r_i + P(v_i, 0)) \leq \sum_{i=1}^{n} r_i + \sum_{i=1}^{n} P(v_i, 0) = 2. \] (2)

From (1) and (2), we know that \( \text{Err}(t) \leq \epsilon \) for all

\[ t \geq \log_{\frac{1}{\epsilon}} \frac{2}{\epsilon}; \]

note that \( \log_{\frac{1}{\epsilon}} \frac{2}{\epsilon} = O(\log \frac{1}{\epsilon}) \).

Finally, since

\[ \text{Err}'(t) \leq \text{Err}(t) \]

holds for all \( t \geq 0 \), we conclude the proof.