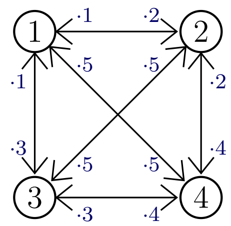


1. Jeremy is taking random walks on the digraph shown at right. At any vertex Jeremy has three choices as to where to go next:

- go to the vertex diagonally across the square.
- go to the adjacent vertex around the square in either a clockwise or an anticlockwise direction;

He takes the first option with probability one half but his probability for the other options is equal to one tenth of the destination vertex number. For example, if he is at vertex 1, his probabilities of next visiting vertices 2, 3 and 4 are 0.2, 0.3 and 0.5 respectively.

(a) List all the walks of length 2 from vertex 1 to vertex 2 and the probabilities associated with each of them. Hence find the probability that a walk of length 2 that starts at vertex 1 finishes at vertex 2.

$1—3—2 : p = \cdot 3 \times \cdot 5 = \cdot 15$
$1—4—2 : p = \cdot 5 \times \cdot 2 = \cdot 10$ $\qquad$ *Probability* $= \cdot 15 + \cdot 10 = \boxed{\cdot 25}$

(b) List all the walks of length 3 from vertex 1 to vertex 2 (there are a lot!) and the probabilities associated with each of them. Hence find the probability that a walk of length 3 that starts at vertex 1 finishes at vertex 2.

$1—2—1—2 : p = \cdot 2 \times \cdot 1 \times \cdot 2 = \cdot 004$
$1—2—3—2 : p = \cdot 2 \times \cdot 5 \times \cdot 5 = \cdot 050$
$1—2—4—2 : p = \cdot 2 \times \cdot 4 \times \cdot 2 = \cdot 016$

$1—3—1—2 : p = \cdot 3 \times \cdot 1 \times \cdot 2 = \cdot 006$
$1—3—4—2 : p = \cdot 3 \times \cdot 4 \times \cdot 2 = \cdot 024$

$1—4—1—2 : p = \cdot 5 \times \cdot 5 \times \cdot 2 = \cdot 050$
$1—4—3—2 : p = \cdot 5 \times \cdot 3 \times \cdot 5 = \cdot 075$

Probability $= \cdot 004 + \cdots + \cdot 075 = \boxed{\cdot 225}$

(c) Compile Jeremy's transition matrix T. Check that it is stochastic.

$$T = \begin{bmatrix} 0 & \cdot 2 & \cdot 3 & \cdot 5 \\ \cdot 1 & 0 & \cdot 5 & \cdot 4 \\ \cdot 1 & \cdot 5 & 0 & \cdot 4 \\ \cdot 5 & \cdot 2 & \cdot 3 & 0 \end{bmatrix}.$$

All entries are non-negative.
Each row sums to 1.
So T is stochastic.

(d) Calculate T^2 (by hand or computer) and use it to check your answer to (a).

$$T^2 = \begin{bmatrix} 0 & \cdot 2 & \cdot 3 & \cdot 5 \\ \cdot 1 & 0 & \cdot 5 & \cdot 4 \\ \cdot 1 & \cdot 5 & 0 & \cdot 4 \\ \cdot 5 & \cdot 2 & \cdot 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & \cdot 2 & \cdot 3 & \cdot 5 \\ \cdot 1 & 0 & \cdot 5 & \cdot 4 \\ \cdot 1 & \cdot 5 & 0 & \cdot 4 \\ \cdot 5 & \cdot 2 & \cdot 3 & 0 \end{bmatrix} = \begin{bmatrix} \cdot 30 & \boxed{\cdot 25} & \cdot 25 & \cdot 20 \\ \cdot 25 & \cdot 35 & \cdot 15 & \cdot 25 \\ \cdot 25 & \cdot 10 & \cdot 40 & \cdot 25 \\ \cdot 05 & \cdot 25 & \cdot 25 & \cdot 45 \end{bmatrix}.$$

(e) Calculate T^3 (by hand or computer) and use it to check your answer to (b). If the answers don't agree, you probably missed some of the walks.

$$T^3 = \begin{bmatrix} 0 & \cdot 2 & \cdot 3 & \cdot 5 \\ \cdot 1 & 0 & \cdot 5 & \cdot 4 \\ \cdot 1 & \cdot 5 & 0 & \cdot 4 \\ \cdot 5 & \cdot 2 & \cdot 3 & 0 \end{bmatrix} \begin{bmatrix} \cdot 30 & \cdot 25 & \cdot 25 & \cdot 20 \\ \cdot 25 & \cdot 35 & \cdot 15 & \cdot 25 \\ \cdot 25 & \cdot 10 & \cdot 40 & \cdot 25 \\ \cdot 05 & \cdot 25 & \cdot 25 & \cdot 45 \end{bmatrix} = \begin{bmatrix} \cdot 150 & \boxed{\cdot 225} & \cdot 275 & \cdot 350 \\ \cdot 175 & \cdot 175 & \cdot 325 & \cdot 325 \\ \cdot 175 & \cdot 300 & \cdot 200 & \cdot 325 \\ \cdot 275 & \cdot 225 & \cdot 275 & \cdot 225 \end{bmatrix}.$$

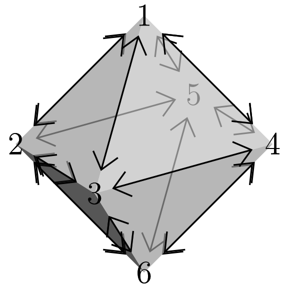
(f) Use T to verify that, if Jeremy keeps walking, then in the long run his visits to any vertex v will constitute a fraction $(v+5)/30$ of his visits to all vertices. (Thus, for example, 20% of his visits will be to vertex 1 and 30% to vertex 4.)

The claim is that the steady state vector is $S = \begin{bmatrix} 6/30 \\ 7/30 \\ 8/30 \\ 9/30 \end{bmatrix}$.

We need to check that $T'S = S$*:*

$$T'S = \begin{bmatrix} 0 & \cdot1 & \cdot1 & \cdot5 \\ \cdot2 & 0 & \cdot5 & \cdot2 \\ \cdot3 & \cdot5 & 0 & \cdot3 \\ \cdot5 & \cdot4 & \cdot4 & 0 \end{bmatrix} \frac{1}{30} \begin{bmatrix} 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = \frac{1}{30} \begin{bmatrix} 0 & \cdot1 & \cdot1 & \cdot5 \\ \cdot2 & 0 & \cdot5 & \cdot2 \\ \cdot3 & \cdot5 & 0 & \cdot3 \\ \cdot5 & \cdot4 & \cdot4 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = \frac{1}{30} \begin{bmatrix} 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = S.$$

2. Anton the ant is wandering around the structure shown at right, made from twelve length 1 struts joined at six vertices 1,...,6, making a regular octahedron with the struts as edges. He walks at a steady pace along the edges and at each vertex makes a random (equal-probability) choice as to which edge to walk next (this includes the possibility of retracing the previous edge).



(a) Anton starts at vertex 1. Explain why, after a walk of length two, he is twice as likely to be at vertex 6 than at vertex 5.

Since four edges meet at each vertex, the probability of Anton taking any one of them is $1/4$. *Thus all walks of length 2 starting at vertex 1 have probability* $(1/4)^2 = 1/16$. *But there are four walks of length 2 from vertex 1 to vertex 6 but only two from vertex 1 to vertex 5.*

(b) Again starting from vertex 1, what is the probability that after a walk of length 3 (not necessarily using distinct edges), Anton is back at vertex 1?

Walks of length 3 from vertex 1 back to vertex 1 are
1231, 1251, 1321, 1341, 1431, 1451, 1541, 1521.

So the probaility is $8(1/4)^3 = \boxed{1/8}$.

(c) Write out Anton's transition matrix T. Write it in the form $T = fM$ where f is a fraction and M is a symmetric matrix of zeros and ones.

$$T = \begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

(d) Making good use of the simple nature of M to save work, calculate T^2, writing it in similar simple form. Use T^2 to check your answer to (a).

$$T^2 = \frac{1}{16} \begin{bmatrix} 4 & 2 & 2 & 2 & 2 & 4 \\ 2 & 4 & 2 & 4 & 2 & 2 \\ 2 & 2 & 4 & 2 & 4 & 2 \\ 2 & 4 & 2 & 4 & 2 & 2 \\ 2 & 2 & 4 & 2 & 4 & 2 \\ 4 & 2 & 2 & 2 & 2 & 4 \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 2 & 1 & 1 & 1 & \boxed{1} & \boxed{2} \\ 1 & 2 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 1 & 1 & 1 & 2 \end{bmatrix}.$$

(e) Calculate $T^4 = (T^2)^2$. Use a computer if you like, but it it is not really necessary.

$$T^4 = (T^2)^2 = \frac{1}{64}\begin{bmatrix} 12 & 10 & 10 & 10 & 10 & 12 \\ 10 & 12 & 10 & 12 & 10 & 10 \\ 10 & 10 & 12 & 10 & 12 & 10 \\ 10 & 12 & 10 & 12 & 10 & 10 \\ 10 & 10 & 12 & 10 & 12 & 10 \\ 12 & 10 & 10 & 10 & 10 & 12 \end{bmatrix} = \frac{1}{32}\begin{bmatrix} 6 & 5 & 5 & 5 & 5 & 6 \\ 5 & 6 & 5 & 6 & 5 & 5 \\ 5 & 5 & 6 & 5 & 6 & 5 \\ 5 & 6 & 5 & 6 & 5 & 5 \\ 5 & 5 & 6 & 5 & 6 & 5 \\ 6 & 5 & 5 & 5 & 5 & 6 \end{bmatrix}.$$

(f) Suppose Anton has been wandering for a long time. By comparing T, T^2 and T^4, and by considering the symmetry inherent in the situation, estimate the proportion of Anton's vertex visits that have been to vertex 1.

The entries in the matrix are getting closer to one another, suggesting that higher powers of T will have entries essentially all equal. Since T must remain stochastic, row sums must be 1, so it appears that:

$$T^n \longrightarrow \frac{1}{6}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \textit{as } n \longrightarrow \infty, \qquad \textit{implying that } S = \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

This looks plausable in view of the complete symmetry of the graph. It suggests that $1/6$ of all Anton's visits will be to vertex 1.

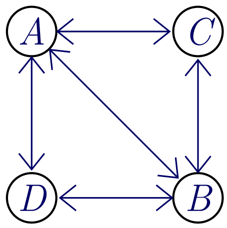
(g) Support your answer to (f) by demonstrating a steady state vector for T.

$$T'S = \frac{1}{4}\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = S.$$

3. A tiny web graph has four pages A, B, C, D. Every page has a link to every other page except for pages C and D, neither of which has link to the other.

(a) Draw the web graph.

(b) Write out the transition matrix, with probabilities assigned according to the Google PageRank algorithm.



$$T = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}.$$

(c) Calculate the page ranks of each page.
Hint: By symmetry, pages A and B will have the same rank, as will pages C and D. So there will only be two unknowns.

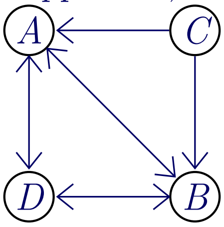
Suppose A, B have page rank p, C, D page rank q. Then $\boldsymbol{PR} = \begin{bmatrix} p \\ p \\ q \\ q \end{bmatrix}$.
We need to solve $T'\boldsymbol{PR} = \boldsymbol{PR}$.
We use $(T' - I)\boldsymbol{PR} = 0$ with the last row replaced by 1's:

$$\begin{bmatrix} -1 & 1/3 & 1/2 & 1/2 \\ 1/3 & -1 & 1/2 & 1/2 \\ 1/3 & 1/3 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} p \\ p \\ q \\ q \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \Longrightarrow \begin{array}{c} -2/3p+q=0 \\ -2/3p+q=0 \\ 2/3p-q=0 \\ 2p+2q=1 \end{array} \Longrightarrow \begin{array}{l} q=2/3p \\ p=1/2-q \\ \quad =1/2-2/3p \\ \therefore\ 5/3p=1/2 \textit{ so } p=\cdot 3,\ q=\cdot 2. \end{array}$$

Thus A, B have page rank $\cdot 3$ and C, D have page $\cdot 2$.

4. Repeat Question 3 with the links from A and B to C removed.
It will no longer be true that C and D have the same rank, so there will now be three equations in three unknowns. However one equation gives the value of one of the unknowns immediately, leading to a 2×2 system.

Suppose A, B have page rank p and C and D have rank q and r respectively.



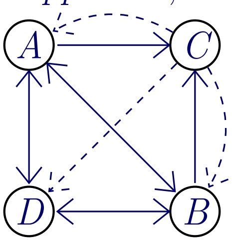
$$T = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix} \implies \begin{bmatrix} -1 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1 & 1/2 & 1/2 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p \\ p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \implies$$

$$\begin{aligned} -p/2+q/2+r/2&=0 \\ -p/2+q/2+r/2&=0 \\ -q&=0 \\ 2p+q+r&=1 \end{aligned} \implies \begin{aligned} q&=0 \\ \therefore\ p&=r \\ \therefore\ 3p&=1 \\ \therefore\ p&=1/3 = r. \end{aligned}$$

Thus C has rank 0 and A, B, D each have rank $1/3$. Looking at the diagram, this is what you should expect. There are no links into C so it has rank 0 and can be ignored. The remaining three pages all have links to each other, so none is more important than the others and hence they have equal rank.

5. Repeat Question 3 but this time with the links **to** A and B **from** C removed.
Again C and D will have different ranks, leading to a 3×3 system, but one equation involves only two unknowns, allowing one unknown to be eliminated from the other two equations, thus generating a 2×2 system.

There are no links from C so we introduce links from C to all other pages. Suppose A, B have page rank p and C and D have rank q and r respectively.



$$T = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix} \implies \begin{bmatrix} -1 & 1/3 & 1/3 & 1/2 \\ 1/3 & -1 & 1/3 & 1/2 \\ 1/3 & 1/3 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p \\ p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \implies$$

$$\begin{aligned} -2p/3+q/3+r/2&=0 \\ -2p/3+q/3+r/2&=0 \\ 2p/3 - q&=0 \\ 2p+q+r&=1 \end{aligned} \implies \begin{aligned} q&=2p/3 \\ \therefore\ r&=8p/9 \\ \therefore\ 2p+2p/3+8p/9&=1 \\ \therefore\ p=9/32,\ q=3/16,\ r&=1/4. \end{aligned}$$

Thus A, B each have rank $9/32$, C has rank $3/16$ and D has rank $1/4$.

6. Repeat Question 3 with 90% damping. This means solving the equation

$$(I-(1-\alpha)T')\boldsymbol{PR} = ({}^{\alpha}\!/\!{}_{n})\mathbf{1}$$

with $\alpha = {}^{1}\!/\!{}_{10}$. Symmetry still applies, so $\boldsymbol{PR}$ still involves only two unknowns, so this 4×4 system boils down to only 2×2, and hence can be solved without the need for Gaussian elimination. However the fractions get a little nastier so you may prefer to use a computer. You should find that the page rank fractions have denominator 64 (when expressed in lowest terms).

$$\left(\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} - \frac{9}{10}\begin{bmatrix}0&1/3&1/2&1/2\\1/3&0&1/2&1/2\\1/3&1/3&0&0\\1/3&1/3&0&0\end{bmatrix}\right)\begin{bmatrix}p\\p\\q\\q\end{bmatrix} = \frac{1/10}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix}1&-3/10&-9/20&-9/20\\-3/10&1&-9/20&-9/20\\-3/10&-3/10&1&0\\-3/10&-3/10&0&1\end{bmatrix}\begin{bmatrix}p\\p\\q\\q\end{bmatrix} = \begin{bmatrix}1/40\\1/40\\1/40\\1/40\end{bmatrix}\begin{matrix}\left.\right\}\frac{7}{10}p-\frac{9}{10}q=\frac{1}{40}\\ \left.\right\}-\frac{6}{10}p+q=\frac{1}{40}\end{matrix}$$

Solving the two equations in p *and* q *gives*

$$\begin{bmatrix}p\\q\end{bmatrix}=\begin{bmatrix}7/10&-9/10\\-3/5&1\end{bmatrix}^{-1}\frac{1}{40}\begin{bmatrix}1\\1\end{bmatrix}=\frac{1}{(7/10-27/50)40}\begin{bmatrix}1&9/10\\3/5&7/10\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\frac{5}{32}\begin{bmatrix}19/10\\13/10\end{bmatrix}=\begin{bmatrix}19/64\\13/64\end{bmatrix}$$

So A, B *have rank* $19/64$ *and* C, D *have rank* $13/64$.

7. Repeat Question 4 with 80% damping. Considerations similar to those for the previous question regarding solution method will apply. You should find D has rank $43/140$.

$$\left(\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} - \frac{8}{10}\begin{bmatrix}0&1/2&1/2&1/2\\1/2&0&1/2&1/2\\0&0&0&0\\1/2&1/2&0&0\end{bmatrix}\right)\begin{bmatrix}p\\p\\q\\r\end{bmatrix} = \frac{2/10}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix}1&-\cdot4&-\cdot4&-\cdot4\\-\cdot4&1&-\cdot4&-\cdot4\\0&0&1&0\\-\cdot4&-\cdot4&0&1\end{bmatrix}\begin{bmatrix}p\\p\\q\\r\end{bmatrix} = \begin{bmatrix}\cdot05\\\cdot05\\\cdot05\\\cdot05\end{bmatrix}\begin{matrix}\left.\right\}\cdot6p-\cdot4q-\cdot4r=\cdot05\\ \left.\right\}q=\cdot05\\ \left.\right\}-\cdot8p+r=\cdot05\end{matrix}$$

Substituting the second equation on the right above into the first, and scaling up by a factor of 10*, leads to the* 2×2 *linear system* $\begin{matrix}6p-4r=\cdot7\\-8p+10r=\cdot5\end{matrix}$. *Hence*

$$\begin{bmatrix}p\\r\end{bmatrix}=\begin{bmatrix}6&-4\\-8&10\end{bmatrix}^{-1}\frac{1}{10}\begin{bmatrix}7\\5\end{bmatrix}=\frac{1}{(60-32)10}\begin{bmatrix}10&4\\8&6\end{bmatrix}\begin{bmatrix}7\\5\end{bmatrix}=\frac{1}{280}\begin{bmatrix}90\\86\end{bmatrix}=\begin{bmatrix}9/28\\43/140\end{bmatrix}$$

So the page ranks are A, B*:* $9/28\approx\cdot32$*;* C*:* $1/20=\cdot05$ *and* D*:* $43/140\approx\cdot31$.
[Compare the undamped ranks of $\cdot33$ *for* A, B, D *and* 0 *for* C*.]*

8. Again repeat Question 4 with 80% damping, but this time use two steps of the iterative method

$$P_0 = ({}^1/{}_n)\mathbf{1}; \quad P_k = \alpha P_0 + (1-\alpha)T'P_{k-1}, \ k \geq 1.$$

That is, calculate P_0, then P_1 and finally P_2, and use P_2 as an approximation for $\boldsymbol{PR}$. Compare the result to your answer to question 7.

$$P_0 = {}^1/{}_n\mathbf{1} = \frac{1}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}\cdot25\\\cdot25\\\cdot25\\\cdot25\end{bmatrix}.$$

$$P_1 = \frac{2}{10}\begin{bmatrix}\cdot25\\\cdot25\\\cdot25\\\cdot25\end{bmatrix} + \frac{8}{10}\begin{bmatrix}0 & {}^1/{}_2 & {}^1/{}_2 & {}^1/{}_2\\ {}^1/{}_2 & 0 & {}^1/{}_2 & {}^1/{}_2\\ 0 & 0 & 0 & 0\\ {}^1/{}_2 & {}^1/{}_2 & 0 & 0\end{bmatrix}\begin{bmatrix}\cdot25\\\cdot25\\\cdot25\\\cdot25\end{bmatrix} = \begin{bmatrix}\cdot05\\\cdot05\\\cdot05\\\cdot05\end{bmatrix} + \cdot8\begin{bmatrix}{}^3/{}_8\\{}^3/{}_8\\0\\{}^2/{}_8\end{bmatrix} = \begin{bmatrix}\cdot35\\\cdot35\\\cdot05\\\cdot25\end{bmatrix}.$$

$$P_2 = \begin{bmatrix}\cdot05\\\cdot05\\\cdot05\\\cdot05\end{bmatrix} + \cdot8\begin{bmatrix}0 & {}^1/{}_2 & {}^1/{}_2 & {}^1/{}_2\\ {}^1/{}_2 & 0 & {}^1/{}_2 & {}^1/{}_2\\ 0 & 0 & 0 & 0\\ {}^1/{}_2 & {}^1/{}_2 & 0 & 0\end{bmatrix}\begin{bmatrix}\cdot35\\\cdot35\\\cdot05\\\cdot25\end{bmatrix} = \begin{bmatrix}\cdot05\\\cdot05\\\cdot05\\\cdot05\end{bmatrix} + \cdot8\begin{bmatrix}\cdot325\\\cdot325\\0\\\cdot35\end{bmatrix} = \begin{bmatrix}\cdot31\\\cdot31\\\cdot05\\\cdot33\end{bmatrix}.$$

We see that P_2 is already within one or two percentage points of $\boldsymbol{PR}$.

9. The adjacency matrix A for a twelve page web graph is given at right.

$$\begin{bmatrix}
0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

(a) Write out the corresponding transition matrix T.

$$\begin{bmatrix}
0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5}\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3}\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\
\frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12}
\end{bmatrix}$$

Each 1 in A is replaced by the probability ${}^1/{}_{n_i}$ where n_i is the number of 1's in its row. For a row of 0's – only the last row for our A – each entry is replaced by ${}^1/{}_n$.

(b) Using a computer, find the page ranks of each page using the Google PageRank algorithm without damping.

We need to solve the equation $(T' - I)\boldsymbol{PR} = 0$, where $\boldsymbol{PR}$ is the PageRank (steady state) vector. The input to, and output from, the Matrix Reshish computer program that does this is discussed and presented below.

For a unique solution, we must use the fact that $\boldsymbol{PR}$ *is a probability vector. So we use the 'short cut' of replacing the bottom-row entries of the matrix equation by all 1's.*
We use the "Gauss-Jordan elimination" tool in Matrix Reshish to solve the resulting matrix equation. For this we need to enter the modified 12×12 *matrix* $(T' - I)$ *and an extra column for the modified right hand side vector of all zeros except for 1 at the bottom. We have the option to input the entries in fractional or decimal form. Using fractional form guarantees exact answers, but unfortunately there is no option to translate the fractional answers into decimal form, which is needed in order to compare magnitudes. There are several workarounds for this issue. One is to use a different computer application, such as MatrixCalc (as in part (c) below) or WolframAlpha which does have the option. Another is to enter fractions like* $1/3$ *and* $1/12$ *with lots of decimal places to reduce rounding error. Yet another, which I will use here, is to first multiply both sides of the equation by a constant factor to avoid the need for repeating decimals. (This clearly doesn't affect the solution of the equation.) For our matrix, a factor of 3 works. Matrix Reshish requires all decimal expressions to have at least one digit (even if it's just 0) before the decimal point, so our* 12×13 *augmented matrix input, and the corresponding output, looks like this:*

Decimal

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	b
1	-3	0	0	0	0	0	0	0	0	0	0	0.25	0
2	0.75	-3	0	0	0	0	0	0	0	0	0	0.25	0
3	0	0	-3	0	0	0	0	0	0	0	0	0.25	0
4	0.75	0	0	-3	0	0	0	0	0	0	0	0.25	0
5	0.75	1.5	0	1	-3	0	0	0	0	0	0	0.25	0
6	0	1.5	3	0	0.6	-3	0	0	0	0	0	0.25	0
7	0	0	0	1	0.6	0	-3	0	0	0	0	0.25	0
8	0.75	0	0	1	0.6	1.5	1	-3	0	0	0	0.25	0
9	0	0	0	0	0.6	1.5	0	1.5	-3	0	0	0.25	0
10	0	0	0	0	0	0	1	0	0	-3	0	0.25	0
11	0	0	0	0	0	0	0	1.5	0	3	-3	0.25	0
12	3	3	3	3	3	3	3	3	3	3	3	3	3

Solution set:

$x_1 = 0.025263157894736842104$
$x_2 = 0.031578947368421052629$
$x_3 = 0.025263157894736842104$
$x_4 = 0.031578947368421052629$
$x_5 = 0.057894736842105263156$
$x_6 = 0.077894736842105263155$
$x_7 = 0.047368421052631578946$
$x_8 = 0.10842105263157894736$
$x_9 = 0.13$
$x_{10} = 0.041052631578947368418$
$x_{11} = 0.12052631578947368420$
$x_{12} = 0.30315789473684210526$

So, to the nearest percentage point, the page ranks are:

Page:	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Rank:	*3%*	*3%*	*3%*	*3%*	*6%*	*8%*	*5%*	*11%*	*13%*	*4%*	*12%*	*30%*

Note how page 12 comes out with far higher rank than any of the other pages. The explanation for this can be found by examining the adjacency matrix A. Every row except the last has at least one nonzero entry, representing a link, and each of these is to the right of the main diagonal. This means that every link points to higher numbered page than the page it is on, and following links inevitably leads to page 12. If, instead of the introduced 'teleporting' links to every page, page 12 was just given one link to itself, then it would have rank 1 and all other pages would have rank 0.

(c) Repeat part (b) but using 60% damping.

This time we must solve the slightly more elaborate equation

$$(I-(1-\alpha)T')\boldsymbol{PR} = (\alpha/n)\mathbf{1},$$

with $(1-\alpha) = \cdot 6$ *(and so* $\alpha = \cdot 4$*) and* $n = 12$*. Fortunately the* $\cdot 6$ *factor already eliminates any repeating decimals in the LHS, and the necessity for* $\boldsymbol{PR}$ *to be a probability vector is built in to our equation. But* $\alpha/n = 1/30$ *introduces a repeating decimal. We could apply the same technique as in (b) and multiply both sides of the equation by 3, or we could just approximate* $1/30$ *with* $0 \cdot 033333$ *say. But instead I will take the opportunity to demonstrate the use of MatrixCalc.*

We have to solve a matrix equation of the form $\mathbf{Ax} = \mathbf{b}$*, where* $\mathbf{A}$ *is a* 12×12 *matrix, but as you know this is equivalent to solving a system of 12 linear equations in 12 unknowns, with the entries of* $\mathbf{A}$ *giving the coefficients of unknowns* $x_1, \ldots, x_{12}$ *and the entries of* $\mathbf{b}$ *giving the 12 right-hand-sides of the equations. To use MatrixCalc we must take this view and use its 'Solving Systems of Linear Equations' tool. First repeatedly click on the 'Cells' button until 12 blank equations are shown, then enter all the non-zero coefficients and right-hand-sides. There is no need to enter any zero coefficients. Also, MatrixCalc is less fussy than Matrix Reshish in that, for example, you can enter just '.95' rather than '0.95' and, most importantly, you can mix decimal and fractional entries. Furthermore you can set how many decimal places you want displayed in the answer (if you make no choice, answers are displayed as fractions). Irrespective of what you set for display, the actual calculations are done in exact arithmetic.*

Once all the data is entered, click the 'Solve by Gauss-Jordan elimination' button. This not only generates the answer but shows all the gory details of the elimination process - you have to scroll down a long way to see the answer! Here are copies of the essential parts of the input and output:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	=
1											-.05	1/30
-.15	1										-.05	1/30
		1									-.05	1/30
-.15			1								-.05	1/30
-.15	-.3		-.2	1							-.05	1/30
	-.3	-.6		-.12	1						-.05	1/30
			-.2	-.12		1					-.05	1/30
-.15			-.2	-.12	-.3	-.2	1				-.05	1/30
				-.12	-.3		-.3	1			-.05	1/30
						-.2			1		-.05	1/30
							-.3		-.6	1	-.05	1/30
				-.12		-.2		-.6		-.6	.95	1/30

$x_1 = 0.043$
$x_2 = 0.050$
$x_3 = 0.043$
$x_4 = 0.050$
$x_5 = 0.075$
$x_6 = 0.093$
$x_7 = 0.062$
$x_8 = 0.109$
$x_9 = 0.113$
$x_{10} = 0.056$
$x_{11} = 0.109$
$x_{12} = 0.198$

So, to the nearest percentage point, the new page ranks are:

Page:	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Rank:	*4%*	*5%*	*4%*	*5%*	*8%*	*9%*	*6%*	*11%*	*11%*	*6%*	*11%*	*20%*

As expected, the heavy damping factor of 60% has increased the lower ranks but significantly reduced the high rank of page 12.