The Google matrix. Let $P_1, \ldots, P_n$ denote the web pages on the internet. For each i let x_i denote the 'importance' of page P_i. This is a positive real number. A page is considered important if many important pages link to it. However, a link from P_j to P_i is weighted with the total number of links n_j that page P_j contains. This leads to the equations

$$x_i = \sum_{P_j \to P_i} \frac{1}{n_j} x_j, \qquad \text{for every } i = 1, 2, \ldots, n.$$

Here the sum runs over the web pages P_j that contain a link to page P_i. Google's web crawlers determine the numbers n_j for each page A_j and keep the information up to date. Google's pagerank algorithm regularly determines the vector $(x_1, x_2, \ldots, x_n)$ of importance scores. We explain why the vector is unique and indicate how it can be computed.

The relations above can be expressed by saying that the vector with coordinates x_i is an eigenvector with eigenvalue 1 of the $n \times n$-matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

given by

$$a_{ji} = \begin{cases} 1/n_j, & \text{when there is a link from page } P_j \text{ to page } P_i; \\ 0, & \text{otherwise.} \end{cases}$$

Indeed, we have

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

If there is any link on page P_j at all, then the sum of the entries in the j-th column of A is equal to 1. However, if there are no links from page P_j to other pages, this definition implies that the j-th column of A contains only zeroes. Pages without links to other pages are very common. They are the so-called 'dangling nodes' of the graph of the internet. Their existence prevents the matrix from being stochastic. In order to get rid of the zero columns there are several strategies, one of which is to replace each of them by the vector all whose entries are equal to $1/n$.

Now the sums of the entries in each column of A are 1, but there still may be many zero entries. In order to eliminate those, we take a convex combination of M and a 'damping' matrix, which assigns the same importance to every page on the internet. In other words, we replace A by

$$M = (1 - \varepsilon)A + \varepsilon E,$$

where E denotes the matrix all of whose entries are equal to $1/n$ and ε is in the interval $(0, 1)$. In their original paper, Brin and Page suggest to take $\varepsilon = 0,15$.

Therefore the following theory applies . Let

$$M = \begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} \end{pmatrix}$$

be a stochastic $n \times n$-matrix. This means that the entries m_{ij} are positive real numbers and that the entries in each column sum up to 1.

$$\sum_{i=1}^{n} m_{ik} = 1, \qquad \text{for } k = 1, \dots, n.$$

For a vector $\mathbf{w} \in \mathbf{R}^n$ we write $\|\mathbf{w}\|$ for the sum of the absolute values of the coordinates of $\mathbf{w}$.

Theorem. *Let $f : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ be the linear map given by $f(\mathbf{x}) = M \cdot \mathbf{x}$. Let W be the subspace of $\mathbf{R}^n$ consisting of the vectors with sum of their coordinates equal to 0. Then*
(a) f preserves the sum of the coordinates of $\mathbf{x} \in \mathbf{R}^n$. In particular, f maps W to itself.
(b) $\lambda = 1$ is an eigenvalue of f. The corresponding eigenspace V has dimension 1. The coordinates of each non-zero vector in V are all non-zero and have the same sign.
(c) The vector space $\mathbf{R}^n$ is a direct sum of V and W.
(d) There is a constant c with $0 < c < 1$ and depending only on M for which

$$\|f(\mathbf{x})\| \leq c\|\mathbf{x}\|, \qquad \textit{for all } \mathbf{x} \in W.$$

Proof. (a) Writing x_k for the coordinates of $\mathbf{x}$, the i-th coordinate of $M(\mathbf{x})$ is $\sum_{k=1}^{n} m_{ik}x_k$. Their sum is

$$\sum_{i=1}^{n}\sum_{k=1}^{n} m_{ik}x_k = \sum_{k=1}^{n}\sum_{i=1}^{n} m_{ik}x_k = \sum_{k=1}^{n} x_k.$$

(b) Since $\lambda = 1$ is an eigenvalue of the transpose of M (with eigenvector $(1, 1, \dots, 1)$), it is also an eigenvalue of the matrix M. Let $\mathbf{v} \in \mathbf{R}^n$ denote an eigenvector. Writing v_k for the coordinates of $\mathbf{v}$, we have $v_i = \sum_{k=1}^{n} m_{ik}v_k$ for $i = 1, \dots, n$. If the signs of the coordinates v_i are not all the same or if some of them are zero, the fact that all m_{ij} are positive implies that $|v_i| < \sum_{k=1}^{n} m_{ij}|v_k|$. However, this implies

$$\sum_{i=1}^{n} |v_i| < \sum_{i=1}^{n}\sum_{k=1}^{n} m_{ik}|v_k| = \sum_{k=1}^{n}\sum_{i=1}^{n} m_{ik}|v_k| = \sum_{k=1}^{n} |v_k|.$$

Contradiction. Therefore the signs of the v_k are all the same. It follows that $\mathbf{v} \notin W$, so that the intersection of the eigenspace V with W is trivial. Since $V \neq \{0\}$, Grassmann implies that V has dimension 1 and hence that $V + W = \mathbf{R}^n$. This proves (b) and (c).

(d) Let $\mathbf{x} \in W$. Then the i-th coordinate of $f(\mathbf{x})$ is $\sum_k m_{ik}x_k$. Writing ϵ_i for the sign of $\sum_k m_{ik}x_k$, we have that

$$\|f(\mathbf{x})\| = \sum_{i=1}^{n} \epsilon_i \sum_{k=1}^{n} m_{ik}x_k = \sum_{k=1}^{n}(\sum_{i=1}^{n} \epsilon_i m_{ik})x_k.$$

Since $f(\mathbf{x})$ is in W, its coordinates cannot all have the same sign ϵ_i. Since $\sum_{i=1}^n m_{ik} = 0$, it follows that the absolute value of $\sum_{i=1}^n \epsilon_i m_{ik}$ is at most $1-2\mu$, where μ is the minimum of all m_{ij}. Note that $0 < \mu \leq 1/n$. Put $c = 1-2\mu$. Then we have

$$\|f(\mathbf{x})\| \leq c\sum_{k=1}^{n} |x_k| = c\|\mathbf{x}\|,$$

as required. Note that $0 < c < 1$ unless, we have $n = 2$ and $\mu = 1/2$, in which case $c = 0$ and all entries of the 2×2-matrix are equal to $1/2$. If this happens, we have $f(W) = 0$ and the statement of part (d) is trivially true for any choice of c.

Corollary. *Let f be as in the theorem. Let $\mathbf{x} \in \mathbf{R}^n$ but $\mathbf{x} \notin W$. Then the sequence $f^k(\mathbf{x})$ converges to a non-zero multiple of the eigenvector $\mathbf{v}$ of the previous theorem.*

Proof. By part (c) of the theorem, $\mathbf{R}^n$ is the direct sum of W and the eigenspace $V = \mathrm{Span}(\mathbf{v})$. Therefore $\mathbf{x} = \alpha\mathbf{v} + \beta\mathbf{w}$ for some $\mathbf{w} \in W$ and $\alpha, \beta \in \mathbf{R}$. Since $\mathbf{x} \notin W$, the coefficient α is not zero. Since $f^k(\mathbf{v}) = \mathbf{v}$ we have

$$f^k(\mathbf{x}) - \alpha\mathbf{v} \;=\; \alpha f^k(\mathbf{v}) + \beta f^k(\mathbf{w}) - \alpha\mathbf{v} \;=\; \beta f^k(\mathbf{w}), \qquad \text{for every } k > 0.$$

By part (d) of the theorem we have that $\|f^k(\mathbf{x}) - \alpha\mathbf{v}\,\| \leq \beta c^k\|\mathbf{w}\|$. Since $0 < c < 1$, this tends to zero as $k \to \infty$.

Example. The matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1/3 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/3 & 0 \\ 1/2 & 1/2 & 1/2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 0 \end{pmatrix}$$

is associated to an internet consisting of six pages. The fourth page is a dangling node. Since almost every page links to page 5, that page looks the most important. On the other hand page 3 seems unimportant. Only page 2, which does not look very important itself, links to it. We do the computation to see whether this is what the pagerank algorithm finds. We replace the zeroes in the fourth column by $1/n = 1/6$ and get

$$A = \begin{pmatrix} 0 & 0 & 0 & 1/6 & 1/3 & 0 \\ 1/2 & 0 & 0 & 1/6 & 0 & 0 \\ 0 & 1/2 & 0 & 1/6 & 0 & 0 \\ 0 & 0 & 1/2 & 1/6 & 1/3 & 0 \\ 1/2 & 1/2 & 1/2 & 1/6 & 0 & 1 \\ 0 & 0 & 0 & 1/6 & 1/3 & 0 \end{pmatrix}$$

Since some entries of A are zero, the conditions of the theorem are not satisfied. Nevertheless we find a unique eigenvector $(x_1, x_2, x_3, x_4, x_5, x_6)$ of importance scores with positive coordinates:

$$(0.142,\ 0.111,\ 0.098,\ 0.184,\ 0.321,\ 0.142).$$

Using the damping matrix E hardly changes the outcome:

$$(0.144,\ 0.103,\ 0.082,\ 0.185,\ 0.340,\ 0.144).$$

As expected, page 5 is considered to be the most important, while page 3 has the lowest score. Computing $A^k(\mathbf{x})$ with $\mathbf{x}$ the vector all of whose coordinates are equal to $1/6$ yields for $k = 10$ already a very accurate approximation to the eigenvector:

$$(0.145,\ 0.102,\ 0.082,\ 0.185,\ 0.338,\ 0.145)$$