Google's secret and Linear Algebra

Pablo Fernández Gallardo
Departamento de Matemáticas
Universidad Autónoma de Madrid
pablo.fernandez@uam.es

28th July 2005

Abstract

In this note, we will analyze some of the mathematical aspects behind the (`PageRank`) algorithm used by Google to sort the results of the queries: a tasty cocktail of Linear Algebra, Graph Theory and Probability that makes life easier.

1 Introduction

Newspapers all around the world are considering these days the news about Google's plan to go public. It is not only the volume of the transaction[1], but also the special meaning of being the first trade of this kind since the dot-com "irrational exuberance" in the 90s.

But there is something else that explains the significance of this piece of news, and mainly related to the particular characteristics of the firm. A few decades ago, there was a complete revolution in Technology and Communications —and also (no doubt about it?) a cultural, sociological... revolution, namely the generalización of use and access to the Internet. Google's appearance has represented a revolution comparable to the former, as it became a tool which brought some order into this universe of information —not manageable before.

The design of a web search engine is a problem of *mathematical engineering*. Notice the adjective. First, a deep knowledge of the context is needed, in order to translate it into models, into Mathematics. But after this process of abstraction, of mathematization, and once the relevant conclusions have been drawn, it is essential to carry out a thorough, detailed and efficient design of the computational aspects inherent in this problem.

[1] Shares for a value of 2700 million dollars will be sold. Driven by the memory of past excesses and scandals of the financial bubble linked to high-tech firms in the past years, the offer will be articulated as an on-line auction, allowing many investors to have similar opportunities to buy shares. The ultimate purpose is to avoid important speculative deals... but one never can tell.

2 The Google engine

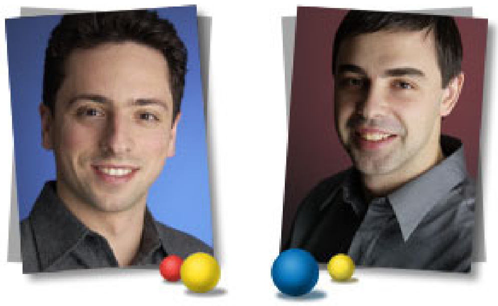


Figure 1: Brin (a la izquierda) y Page

The origin of Google search engine is well known. It was designed in 1998 by Sergei Brin and Lawrence Page, two Computer Science doctorate students at Stanford (Brin had a degree in Mathematics, and Page, in Computer Science). See them in the pictures[2]. Two young men, now in their thirties, who have become multimillionaires. The odd name of the engine is a variation from the term *googol*, the one who somebody[3] invented to refer to the overwhelming number 10^{100}. One of those numbers mathematicians are comfortable with but perhaps bigger than the number of particles in the whole Universe.

Although not so huge, the scale of the question we are concerned about is also immense. In 1997, when Brin and Page were to start working in Google's design, there were about 100 million web pages. Altavista, the most popular search engine in those days, attended 20 million of daily queries. Today, these figures have been multiplied: Google receives some hundred million of daily queries and indexes several billions of web pages.

Therefore, the design of a search engine must efficiently solve some computational aspects, namely, the way to store that enormous amount of information, how is it updated, how to manage the queries, the way to search in the databases, etc.

But, although interesting, we are not going to treat these questions here. The point of interest can be formulated in a simple manner: Let us suppose that, after a certain query, we have determined that, say, one hundred web pages enclose information that might be relevant, in some sense, for the user. Now,

in which order should they be displayed?

The objective, as explicitly posed[4] by Brin and Page (see [6]), is that, in a

[2]We swear we have nothing to do with that sort of Christmas balls decorating the pictures. They have been taken, just as they were, from Google's own web page.

[3]It is said that a nephew of the mathematician Edward Kasner. Kasner cheered up and also defined the googolplex, its value being 10^{googol}. Wow!

[4]This was not the only objective. They also pretended the search engine to be "resistant" to any kind of manipulation, for instance, to commercial-oriented attempts to place certain pages at the top positions on the list. Curiously enough, nowadays a new "sport", *Google bombing*, has become very popular: it is about trying to place a web page in the top positions, although usually with recreational purposes. The reader might try with queries such as "miserable failure". The result might surprise him. . . and even delight him!

sufficiently big number of times, al least one of, say, the first ten displayed pages contain useful information for the user.

We now ask the reader (most probably a *google-maniac*) to decide, from his own experience, whether Google fulfills this objective or not. We are sure the common response will be affirmative... and even amazingly affirmative! It seems to be Magic[5]... but it is just Mathematics.

And Mathematics requiring no more than tools of a first year graduate course, as we will see soon.

To tackle our task, we need an **ordering criterion**. Notice that (now we have already shifted to mathematical mode) if we label each web page with symbols $P_1, \dots, P_n$, all we want is to assign each P_j a number x_j, its **significance**. These numbers might be, for example, between 0 and 1.

Let us suppose then that, after a webcrawling, we have collected a complete list of web pages, and that we have assigned each of them a significance, no matter how. We can use this ordering each time we answer a query: the selected pages will be displayed ordered as prescribed by the list.

The explanation of how to build that list is still missing. Let us go for it.

NOTE. To complete the description, although we will not go into the details, we should say that there are a pair of elements used by Google, in combination with the general criterion we will explain here, when answering specific queries:

- On the one hand, as it is reasonable, Google does not give the same "score" to one term that is within the title of the page, either in boldface, in a small font, etc.
- For combined searches, it will be quite different if, within the document, the terms appear "close" or "distant" from each other.

3 The model

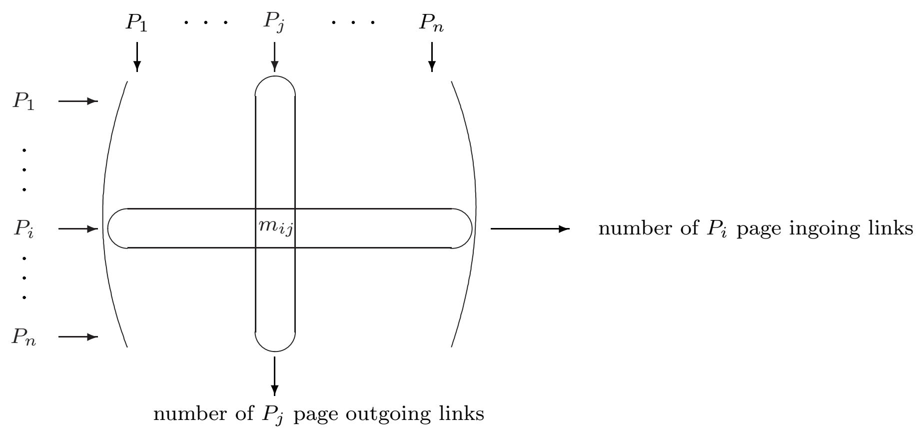
In the first step, we will describe, in an adequate way, the relevant elements for the question we are interested in, the significance assignment. Let us suppose that we have collected all the information about the web: sites, contents, links between pages, etc. In this first step of our process of modeling, we are only going to take into account the information referred to the web pages, labeled with $P_1, \dots, P_n$, and the links between them.

In these terms, the web can be modeled by a (directed) **graph** G. Each web page P_j is a vertex of the graph, and there will be an edge between vertices P_i and P_j whenever there is a link from page P_i to page P_j. It is a gigantic, overwhelming graph, whose real structure will deserve some consideration below (see section 8).

When dealing with graphs, we like to use drawings in the paper, in which vertices are points of the plane, while edges are merely arrows joining these

[5]Not to talk about the incredible capacity of the search engine to "correct" the query terms and suggest, yes!, the word one had indeed in mind. This leads us to envisage supernatural phenomena... well, let us give it up.

points. But, for our purposes, it is helpful to consider an alternative interpretation —matrices. Let us build, then, an $n \times n$ matrix $\mathbf{M}$, whose rows and columns are labeled with symbols $P_1, \ldots, P_n$, and with zero-one entries. The matrix entry m_{ij} will be *one* whenever there is a link between page P_j and page P_i; and *zero* otherwise:



The matriz $\mathbf{M}$ is, except for a transposition, the adjacency matrix of the graph. Notice that it is not necessarily symmetric, because we are dealing with a directed graph. Observe also that, as suggested in the picture, the sum of the entries for P_j's column is the number of P_j's outgoing links, while we get the number of ingoing links summing in rows.

We will assume that the significance of a certain page P_j "is related to" the pages linking to it. This sounds reasonable: if there are a lot of pages pointing to P_j, its information must have been considered as "advisable" by a lot of web-makers.

The above "related to" is still rather vague. A **first attempt**, perhaps a bit naïve, amounts to suppose that the significance x_j of each P_j *is proportional to the number of links entering* P_j. Let us note that, whenever we have the matrix $\mathbf{M}$ at our disposal, the computation of each x_j is quite simple —just sum the entries of each row P_j.

But this model does not adequately grasp a situation deserving attention, namely when a certain page is cited from few, *but very relevant* pages. Say, for example, from `www.microsoft.com`, `www.amazon.com`, etc. The previous algorithm would assign it a low significance, and this is not what we want. So we need to enhance our model in such a way that a strong significance is assigned both

- to highly cited pages;
- and to those that, although not cited so many times, have links from very "significant pages".

Following this line of argument, the **second attempt** asumes that the significance x_j of each page P_j *is proportional to the sum of the significance of* *the pages linking to* P_j. Notice that we have changed, from our first attempt, "number of pages linking to P_j" to "sum of the significance of the pages linking to P_j". This slight variation completely alters the features of the problem.

Suppose, for instance, that page P_1 is cited in pages P_2, P_{25} and P_{256}, that P_2 is only cited in pages P_1 and P_{256}, etc., and that, say, there are links to page P_n from P_1, P_2, P_3, P_{25} and P_{n-1}. Following the previous assignment, x_1 should be proportional to 3, x_2 to 2, etc., while x_n should be proportional to 5.

But now, our assignment $x_1, \ldots, x_n$ must verify that

$$\begin{aligned} x_1 &= K\left(x_2 + x_{25} + x_{256}\right), \\ x_2 &= K\left(x_1 + x_{256}\right), \\ &\vdots \\ x_n &= K\left(x_1 + x_2 + x_3 + x_{25} + x_{n-1}\right), \end{aligned}$$

where K is a certain proportionality constant. In this way, we face an enormous system of linear equations, whose solutions are all the admissible assignments $x_1, \ldots, x_n$.

Let us write this system of equations in a better way, using matrices:

$$\begin{matrix} & & P_1 & P_2 & & & & P_{25} & & & & P_{256} & & & P_{n-1} & \\ & & \downarrow & \downarrow & & & & \downarrow & & & & \downarrow & & & \downarrow & \end{matrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = K \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

And now, one of those "symbol shows" mathematicians are so fond of. Let us call the significance vector $\mathbf{x}$. The $n \times n$ matrix of the system is exactly the matrix $\mathbf{M}$ associated to the graph. So we can state that the significance assignment is a solution of

$$\mathbf{M}\mathbf{x} = \lambda \mathbf{x}\,.$$

We have already used the symbol λ for the constant of proportionality. This is so because, as anyone who has been exposed to a Linear Algebra first course will recognize, the question has become a problem of **eigenvalues** and **eigenvectors**: our yearned significance vector $\mathbf{x}$ is no more than an eigenvector of the matrix $\mathbf{M}$. You might recall that this matrix keeps all the information about the web structure —vertices and adjacency relations.

The reader will certainly find curious the fact that one of the basic elements of any Linear Algebra course allows us to describe a so "applied" question... but perhaps it is not enough to arouse his enthusiasm yet. All right, an eigenvalue. But which one? There are so many... And also: how could we compute it?, the matrix is inconceivable huge: remember, thousand of million rows (or columns).

Be patient.

For the time being, it sounds reasonable to demand the entries of our vector (the significance of the web pages) to be non-negative (or, al least, with the same sign). This will be written as $\mathbf{x} \geq 0$ —we ask the reader to excuse this abuse of notation.

But also, since we pretend the method to be useful, we need this hypothetical non-negative vector to be **unique** —and if there were more than one, which of them should be chosen?

4 The random surfer

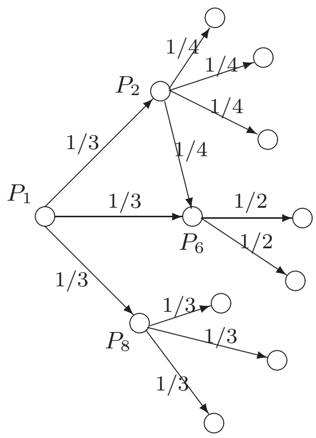
As an interlude, we are going to approach the question from a different point of view. Let us imagine a user surfing the web. At some moment, he will reach some page, say P_1. But, probably bored with the contents of P_1, he will jump to another page, following P_1's outgoing links (suppose there are N_1 possibilities). But, to which one?

Our brave navigator is a *random* surfer —and needless to say, also blond and suntanned. So, in order to decide his destination, he is going to use chance, and in the most simple possible way: with a regular (and virtual, we presume) dice, having so many faces as the number of outgoing links from P_1. In technical terms, the choice of destination follows a (discrete) *uniform* probability distribution in $[1, N_1]$.

Our model is no longer *deterministic*, but *probabilistic*: we do not know where he will be a moment of time later, but we do know *what are his chances* of being in each admissible destination. And it is a *dynamic* model as well, because the same argument may be applied to the second movement, and to the third one, etc. Our surfer is following what is known as a **random walk** in the graph.

We show, in the drawing, a possible situation: there are three edges leaving from P_1 to the vertices P_2, P_6 and P_8. So our navigator draws his destination, assigning probability 1/3 to each vertex. If the result is page P_2, then he is to draw again, but now with probability 1/4 for each possible destination from P_2.

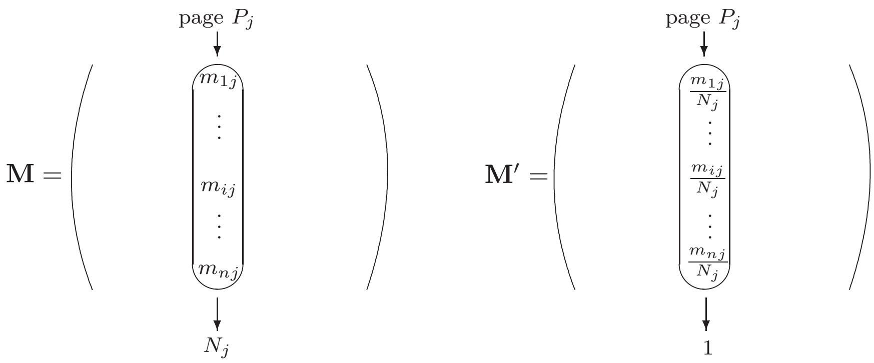
A very suggestive setting, although it is not clear how to formalize it: how could we, for instance, compute the probability of being at page P_{17} after five movements if the surfer leaves from P_1. More than that, we still have no suspicion about how all this is related to our assignment problem.



Little by little. Recall that $\mathbf{M}$ is the matrix of the graph and that its entries are zeros and ones. Let N_j be the number of P_j's outgoing links (that is, the sum of the entries in the column labeled P_j in the matrix $\mathbf{M}$). Let us now build another matrix $\mathbf{M}'$ from the original $\mathbf{M}$ replacing each entry m_{ij} by

$$m'_{ij} = \frac{m_{ij}}{N_j}\,.$$

The entries of the new matrix $\mathbf{M}'$ will be non-negative numbers (between 0 and 1), in such a way that the sum of the entries for *each column* is 1. The next drawing shows the result of the transformation for a certain column:



The so built matrix $\mathbf{M}'$ is called a **stochastic (or Markovian) matrix** .

Let us say that the surfer is in page (vertex) P_k at the beginning. In order to put such a deterministic situation in a probabilistic setting, we say that he is in page P_k with probability 100%. We represent this initial condition with a vector having all its entries 0, but the one in position k, that has a 1. Recall that the surfer draws among the N_k destinations, assigning probability $1/N_k$ to each of them.

But when we multiply the matrix $\mathbf{M}'$ by this initial vector, we get

$$\begin{pmatrix} \cdots & \cdots & m'_{1k} & \cdots \\ \vdots & \ddots & \vdots & \vdots \\ \cdots & \cdots & m'_{kk} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \cdots & \cdots & m'_{nk} & \cdots \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} m'_{1k} \\ \vdots \\ m'_{kk} \\ \vdots \\ m'_{nk} \end{pmatrix}.$$

The result is still a vector with entries (numbers between 0 and 1, as the m'_{jk}s are either 0, or $1/N_k$) summing to 1, because there are exactly N_k non zero entries. But more than that: the vector we get exactly describes the probability of being, one moment later, in each page of the web, assuming he began at P_k.

What makes this model specially useful is that, in order to know the probabilities of being in each page of the web after *two* moments of time, it is enough to repeat the process. That is, to multiply $(\mathbf{M}')^2$ by the initial vector. And the same for the third movement, the fourth, etc.

Following the usual terminology, we consider a certain number of **states**, in our case being just the vertices of the graph G. The matrix $\mathbf{M}'$ is (appropriately) called the **transition matrix** of the system: each entry m'_{ij} describes the probability of going from state (vertex) P_j to state (vertex) P_i. And the entries of the successive powers of the matrix, the probabilities of going from each P_i to each P_j after several moments of time.

This solves our first question: how to calculate the transition probabilities between vertices as time goes by. But we have not discovered the relation with Google's ordering problem yet. We leave this for the moment (see subsection 9.1). But, for the impatient and well versed reader, let us advance that the non-negative vector we have been chasing for a long while turns out to be precisely the stationary state of this Markov chain.

> NOTE. It might happen that some pages had no outgoing links at all (being only zeros in the corresponding columns). It would not be a stochastic matrix. Google's solution: replace the zeros of each "bad" column by $1/n$. In this way, whenever the surfer reaches one of the pages from which there was no exit before, now he is allowed to quit them and go to any page of the web with the same probability. More on this, in section 8.

5 Qualifying for the playoffs

So many re-interpretations, so many re-formulations. . . we may forget the main question: Google's ordering algorithm. Let us go back to it. It must be said that the ideas behind Google's procedure can be traced back to the algorithms developed by Kendall and Wei in the fifties (oops, last century fifties!), see [11] and [16].

Now we will pose one more question, quite similar to the one we have dealt with, but with an entertainment taste. It will be illustrated with a simple example.

Let us consider a sport competition, such as the Professional Basketball League. The 30 teams of the NBA competition are divided (following geographical criteria) into two conferences, each one made up of three divisions:

- Eastern Conference:
  - Atlantic Division, Central Division and Southeast Division.
- Western Conference[6]:
  - Southwest Division, Northwest Division and Pacific Division.

Each team play the same number of games, 82, but not *the same number of games against each other*; it is customary they play more games against the teams from their own conference.

So we may ask the following question: once the regular season is finished, what 16 teams should classify for the playoffs? The standard system computes the number of wins to determine the final positions. But it is reasonable[7] to wonder about whether this is a "fair" system or not. After all, it might happen that a certain team could have obtained a lot of wins just because it was included in a very "weak" conference. What should be worthier: the number of wins, or their "quality"? And we face again Google's dichotomy!

[6] Where "our" Pau Gasol and Raúl López play.
[7] See [10].

In a certain competition there are, say, n teams, $E_1, \dots, E_n$. We register the regular season results in a matrix $\mathbf{A}$, its entries a_{ij} being the number of victories obtained by each team:

$$a_{ij} = \frac{\#\,\text{wins of } E_i \text{ over } E_j}{\#\,\text{matches of } E_i}\,.$$

As a normalization measure, we divide by the number of matches played by each team.

Our target is to assign each team E_j a number x_j, analogously called its significance, in such a way that the array $x_1, \dots, x_n$ means the final placings.

The first model, in which x_j is proportional to the number of wins, does not take into account the quality of the victories.

The second model amounts to decide that the significance x_j is proportional to the number of wins, *weighted* with the significance of the other teams:

$$x_j = K \sum_{k=1}^{n} a_{jk}\, x_k\,.$$

And this leads us, again, to

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}\,.$$

And once more, we want to find a non-negative eigenvector $\mathbf{A}$ (and a unique one, if possible).

5.1 For serious calculations: the computer

Let us consider a competition among six teams, $E_1, \dots, E_6$, divided into two conferences. Each team plays 21 games at all: 6 against each team from its own conference, 3 against the others. The following table includes the information about the results of the competition:

	E_1	E_2	E_3	E_4	E_5	E_6	
E_1	–	3/21	0/21	0/21	1/21	2/21	$\to$ 6/21
E_2	3/21	–	2/21	2/21	2/21	1/21	$\to$ 10/21
E_3	6/21	4/21	–	2/21	1/21	1/21	$\to$ 14/21
E_4	3/21	1/21	1/21	–	2/21	2/21	$\to$ 9/21
E_5	2/21	1/21	2/21	4/21	–	2/21	$\to$ 11/21
E_6	1/21	2/21	2/21	4/21	4/21	–	$\to$ 13/21

To the right, we have registered the number of wins of each team. This count suggests the following ordering:

$$E_3 \to E_6 \to E_5 \to E_2 \to E_4 \to E_1\,.$$

But notice, for example, that the leader team E_3 has collected a lot of victories against E_1, the worst team.

What happens with the second point of view? Some computation must be done. Even in such a simple example as this one, we need to use the computer.

So we ask a mathematical software, say[8] `Maple`, to perform the calculations. The following commands will sound familiar to the experienced reader; and not so hard to understand to the beginner. First, we load a required package:

```
> restart:with(linalg):
```

```
Warning, the protected names norm and trace have been redefined and unprotected
```

And after this brief (and perhaps deep!) answer. . . `Maple` in action!:

```
> A:=matrix(6,6,[0,3/21,0,0,1/21,2/21,3/21,0,2/21,2/21,2/21,1/21,6/21,
  4/21,0,2/21,1/21,1/21,3/21,1/21,1/21,0,2/21,2/21,2/21,1/21,2/21,4/21,0,
  2/21,1/21,2/21,2/21,4/21,4/21,0]);
```

$$A := \begin{bmatrix} 0 & \frac{1}{7} & 0 & 0 & \frac{1}{21} & \frac{2}{21} \\ \frac{1}{7} & 0 & \frac{2}{21} & \frac{2}{21} & \frac{2}{21} & \frac{1}{21} \\ \frac{2}{7} & \frac{4}{21} & 0 & \frac{2}{21} & \frac{1}{21} & \frac{1}{21} \\ \frac{1}{7} & \frac{1}{21} & \frac{1}{21} & 0 & \frac{2}{21} & \frac{2}{21} \\ \frac{2}{21} & \frac{1}{21} & \frac{2}{21} & \frac{4}{21} & 0 & \frac{2}{21} \\ \frac{1}{21} & \frac{2}{21} & \frac{2}{21} & \frac{4}{21} & \frac{4}{21} & 0 \end{bmatrix}$$

```
> autovects:=[evalf(eigenvectors(A))];
```

$$\begin{aligned} autovects := [&[0.012,\ 1.,\{[-0.062,\ -0.916,\ -2.131,\ 0.873,\ 0.731,\ 1.]\}],\\ &[0.475,\ 1.,\ \{[0.509,\ 0.746,\ 0.928,\ 0.690,\ 0.840,\ 1.]\}],\\ &[-0.111+0.117\,I,\ 1.,\{[-0.151-0.901\,I, 0.123+0.451\,I,\ -0.727+0.728\,I,\\ &\quad -0.435-0.128\,I,\ 0.192+0.377\,I,\ 1.]\}],\\ &[-0.126,\ 1.,\{[0.008,\ -0.685,\ 1.434,\ -1.071,\ 0.032,\ 1.]\}],\\ &[-0.139,\ 1.,\{[1.785,\ -3.880,\ 3.478,\ -5.392,\ 4.415,\ 1.]\}],\\ &[-0.111-0.117\,I,\ 1.,\{[-0.151+0.901\,I, 0.123-0.451\,I,\ -0.727-0.728\,I,\\ &\quad -0.435+0.128\,I,\ 0.192-0.377\,I,\ 1.]\}]] \end{aligned}$$

Once we become familiar with the way `Maple` displays the results (each line containing the eigenvalue first, its multiplicity and finally the associated eigenvectors), we find that there are six different eigenvalues: two complex numbers (one the conjugate of the other) and four real numbers. Their moduli are

```
> autovals:=[evalf(eigenvalues(A))]: modulos:=[seq(abs(autovals[j]),j=1..6)];
```

$$modulos := [0.012,\ 0.475,\ 0.161,\ 0.126,\ 0.139, 0.161]$$

We now discover that $\lambda = 0.475$ is the biggest (in modulus) eigenvalue, its associated eigenvector being

$$\mathbf{x} = (0.509,\ 0.746,\ 0.928,\ 0.690,\ 0.840, 1)\,.$$

And that this is the **only** eigenvector having real non-negative entries.

[8] No commercial intention in this: Matlab, Mathematica, Derive or even Excel could do the task.

So we have the answer we were looking for: the components of the vector suggest the following ordering:

$$E_6 \to E_3 \to E_5 \to E_2 \to E_4 \to E_1 \,.$$

And now E_6 is the best team!

Let us summarize. In this particular matrix with non-negative entries (that might be regarded as an small-scale version of the Internet matrix), we are in the best possible situation: there is a unique non-negative eigenvector, the one we need to solve the ordering question we posed.

Did this happen by chance? Or was it just a trick, an artfully choice of matrix to persuade the unwary reader that things work as they should?

The reader, far from being unwary, but curious, is now urgently demanding a categorical response. And he knows, no doubt about it, that it is time to welcome a new actor in this performance.

6 Mathematics enters the stage

It is time for Mathematics, the Science that deals with virtual, abstract objects and, through them, allows us to understand concrete realities. So let us put on mathematical mode, and let us distil the common essence of all the questions we have been dealing with.

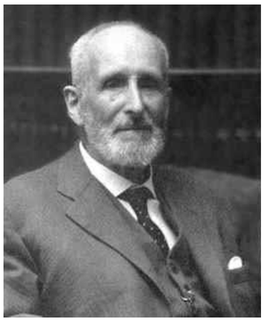


Figure 2: Perron

Doing so, we will discover that the only point shared by all our matrices (being stochastic or not) is that all their entries are **non-negative**. Not quite a lot of information, it seems. Neither symmetric matrices, nor positive definite, nor...

Nevertheless, as shown by Perron[9] at the beginning of the XX century, enough to obtain pleasant results:

Theorem 1 (Perron, 1907) *Let* $\mathbf{A}$ *be an square matrix with positive entries,* $\mathbf{A} > 0$*. Then,*

a) there exists a (simple) eigenvalue $\lambda > 0$ *such that* $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$*, where the corresponding eigenvector is* $\mathbf{v} > 0$*;*

b) this eigenvalue is bigger (in modulus) than the other eigenvalues;

c) any other positive eigenvector of $\mathbf{A}$ *is a multiple of* $\mathbf{v}$*.*

Stop! A pause for meditation: *now* we have a *theorem*. Is it not amazing? If you are able to check that the hypotheses are fulfilled (and no vagueness is allowed: they must be satisfied, exactly satisfied; no excuses, no exceptions), then the

[9] The German mathematician Oskar Perron (1880-1975), a conspicuous example of mathematical longevity, was interested in several fields such as Analysis, Differential Equations, Algebra, Geometry, Number Theory, etc., in which he published several text-books that eventually became classics.

conclusion is unappealable. In our case, no matter how the matrix is, whenever all its entries are positive, there will be a simple eigenvalue such that... an an unique eigenvector with... Wow!

If the reader kindly forgives this almost mystic trance, we will proceed. Let us analyze Perron's result. It points towards the direction we are interested in, but it is not enough, because the matrices we deal with might have zeros. So we need something else.

Figure 3: Frobenius

The following act of this performance was written by Frobenius[10], several years later, when he dealt with the general case of non-negative matrices. Frobenius observed that if we only have that $\mathbf{A} \geq 0$, then, although there is still a *dominant* (of maximum modulus) eigenvalue $\lambda > 0$ associated to an eigenvector $\mathbf{v} \geq 0$, there might be other eigenvalues of the same "size". Here goes his theorem:

Theorem 2 (Frobenius, 1908-1912) *Let* $\mathbf{A}$ *be a square matrix with non-negative entries,* $\mathbf{A} \geq 0$*. If* $\mathbf{A}$ *is **irreducible**, then*

(a) there exists a (simple) eigenvalue $\lambda > 0$ *such that* $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$*, where the corresponding eigenvector is* $\mathbf{v} > 0$*. Besides,* $\lambda \geq |\mu|$ *for any other eigenvalue* μ *of* $\mathbf{A}$*.*

(b) Any eigenvector ≥ 0 *is a multiple of* $\mathbf{v}$*.*

(c) If there are k *eigenvalues of maximum modulus, then they are the solutions of the equation* $x^k - \lambda^k = 0$*.*

NOTE. What does it mean the $n \times n$ matrix $\mathbf{A}$ to be *irreducible*? There are several ways to understand it:

1. There is no permutation (of rows and columns) transforming $\mathbf{A}$ into a matrix of the following type:

$$\left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{0} & \mathbf{A}_{22} \end{array} \right),$$

   where $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are square matrices.

2. All the entries of the matrix $(\mathbf{I} + \mathbf{A})^{n-1}$, where $\mathbf{I}$ stands for the $n \times n$ identity matrix, are positive.

3. If $\mathbf{A}$ is the adjancency matrix of a graph, then the graph is *strongly* connected (see section 8).

[10]Ferdinand Georg Frobenius (1849-1917) was one of the outstanding members of the Berlin School, along with distinguished mathematicians as Kronecker, Kummer or Weierstrass (his thesis supervisor), the leading mathematicians during the end of XIX century and the beginning of XX century. Strict prussian school: rigour, tradition, a "pure" mathematician, with no concession to Applied Mathematics. Changeable History would make many of his ideas on finite groups representations become the basis of Quantum Mechanics. He is well known for this contributions to Group Theory. His works on non-negative matrices belong to the last stage of his live.

Notice first that Frobenius' theorem is indeed a generalization of Perron's result, because if $\mathbf{A} > 0$, then $\mathbf{A}$ is ≥ 0 and irreducible.

Second, if $\mathbf{A}$ is irreducible (whether this is the case or not for the web matrix will be analyzed below), then the question is completely solved: there exists a unique non-negative eigenvector, associated to the positive eigenvalue of maximum modulus (and this will be very useful, as we will see in a moment).

These results, to which we will refer, from now on, as Perron-Frobenius Theorem, are widely used in other contexts (see section 9). Some people even talk about "Perron-Frobenius Theory", this theorem being one of its central results.

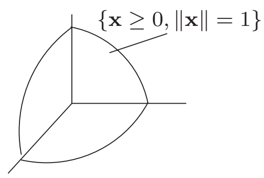
The proof is quite complicated, and here we will just sketch an argument with some of the fundamental ideas.

A "proof" of Perron-Frobenius Theorem (illustrated in the 3×3 case)

Let us start with a non-negative vector $\mathbf{x} \geq 0$. As $\mathbf{A} \geq 0$, the vector $\mathbf{Ax}$ is also non-negative.

In geometric terms, the matrix $\mathbf{A}$ maps the positive octant into itself. Let us consider now the mapping α given by

$$\alpha(\mathbf{x}) = \frac{\mathbf{Ax}}{\|\mathbf{Ax}\|}\,.$$



Notice that $\alpha(\mathbf{x})$ is always a unit length vector. The function α maps the set $\{\mathbf{x} \in \mathbb{R}^3 \,:\, \mathbf{x} \geq 0, \|\mathbf{x}\| = 1\}$, that is, the piece of the unit sphere we draw next to these lines, into itself.

Now, applying Brouwer Fixed Point Theorem, there exists a certain $\tilde{\mathbf{x}}$ such that $\alpha(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}$. Therefore,

$$\alpha(\tilde{\mathbf{x}}) = \frac{\mathbf{A}\tilde{\mathbf{x}}}{\|\mathbf{A}\tilde{\mathbf{x}}\|} = \tilde{\mathbf{x}} \quad \Longrightarrow \quad \mathbf{A}\tilde{\mathbf{x}} = \|\mathbf{A}\tilde{\mathbf{x}}\|\, \tilde{\mathbf{x}}\,.$$

Summing up, $\tilde{\mathbf{x}}$ is an eigenvector of $\mathbf{A}$ with non-negative entries associated to an eigenvalue > 0.

For all other details, such as proving that this eigenvector is (essentially) unique and the other parts of the theorem, we refer the reader to [1], [4], [13] and [14].

7 And what about the computational aspects?

The captious reader will be raising a serious objection: Perron-Frobenius' theorem guarantees the existence of the needed eigenvector for our ordering problem, but *says nothing* about how to compute it. Notice that the proof we sketched is not a constructive one. Thus, we still should not rule out the possibility that these results are not so satisfactory. Recall that Google's matrix is overwhelming. The calculation of our eigenvector could be a cumbersome task!

Let us suppose we are in an ideal situation, in those conditions that guarantee the existence of an positive eigenvalue λ_1 *strictly bigger* (in modulus) than the other eigenvalues. Let $\mathbf{v}_1$ be its (positive) eigenvector.

> NOTE. A matrix $\mathbf{A} \geq 0$ is said to be **primitive** if it has a dominant eigenvalue (bigger, in modulus, than the other eigenvalues). This happens, for instance, when, for a certain positive integer k, all the entries of the matrix $\mathbf{A}^k$ are positive.

We could, of course, compute *all* the eigenvalues and keep the one of interest. But even using efficient methods, the task would be excessive.

But the own structure of the problem helps us again and make the computation easy. A coincidence that comes in handy. It all becomes from the fact that the eigenvector is associated to the dominant eigenvalue.

Suppose, to simplify the argument, that $\mathbf{A}$ is diagonalizable. We have a basis of $\mathbb{R}^n$ with the eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, the corresponding eigenvalues being decreasing size ordered:

$$\lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \,.$$

We start, say, with a certain $\mathbf{v}_0 \geq 0$, that may be written as

$$\mathbf{v}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \,,$$

where the numbers $c_1, \ldots, c_n$ are the $\mathbf{v}_0$ coordinates in our basis. In what follows, *it is not going to be necessary* to compute them explicitly, and in fact we will not do it. It is enough to know that such numbers do exist. Now, we multiply vector $\mathbf{v}_0$ by matrix $\mathbf{A}$, to obtain

$$\mathbf{A}\mathbf{v}_0 = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_n\lambda_n\mathbf{v}_n \,,$$

because the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are eigenvectors of $\mathbf{A}$. We only *calculate* the product in the left —for our argument, it is enough to *know* that the right side is in that way. Excuse the reader being so insistent.

With good cheer, we repeat the operation:

$$\mathbf{A}^2\mathbf{v}_0 = c_1\lambda_1^2\mathbf{v}_1 + c_2\lambda_2^2\mathbf{v}_2 + \cdots + c_n\lambda_n^2\mathbf{v}_n \,.$$

And several more times, say k times:

$$\mathbf{A}^k\mathbf{v}_0 = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_n\lambda_n^k\mathbf{v}_n \,.$$

Let us suppose that $c_1 \neq 0$. Then,

$$\frac{1}{\lambda_1^k}\,\mathbf{A}^k\mathbf{v}_0 = c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n \,.$$

But $|\lambda_j/\lambda_1| < 1$ for each $j = 2, \ldots, n$ (recall that λ_1 was the dominant eigenvalue), so

$$\frac{1}{\lambda_1^k}\,\mathbf{A}^k\mathbf{v}_0 \xrightarrow{k\to\infty} c_1\mathbf{v}_1 \,.$$

Therefore, when repeatedly multiplying the initial vector by the matrix $\mathbf{A}$, we determine, more precisely each time, the *direction* of interest, namely the one given by $\mathbf{v}_1$. This numerical method is known as the **power method**, and its rate of convergence depends on the *ratio* between the first and the second eigenvalue (see in [8] an estimate for Google's matrix).

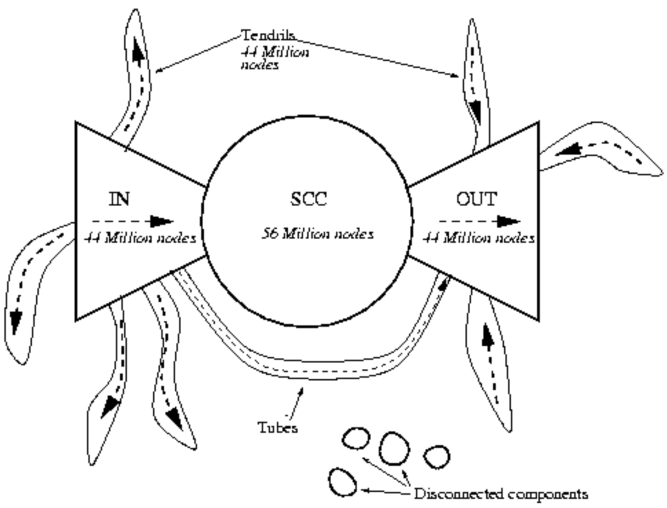
With this observation, our problem is finally solved, at least if we are in the best possible conditions (a non-negative irreducible matrix). The answer does exist, it is unique and we have an efficient method[11] to compute it at our disposal. But...

8 But, are we in an ideal situation?

To make things work properly, we need the matrix $\mathbf{M}$ associated to the web-graph G to be irreducible. In other words, we need G to be a strongly connected graph.

NOTE. Let us consider a directed graph G (a set of vertices and a set of directed edges). The graph is said to be *strongly connected* if, given any two vertices u and v, we are able to find a sequence of edges joining one to the other.
The same conclusion, but "erasing" the directions of the edges, lead us to the concept of *weakly connected* graph. Needless to say, an strongly connected graph is also a weakly connected graph —not necessarily in the other direction.

As the reader might suspect, it is not the case. A research developed in 1999 (see [7]) came to the conclusion that, among the 203 million pages under study, 90% of them laid in a gigantic (weakly connected) component, this in turn having a quite complex internal structure, as can be seen in the following picture, taken from [7]:



[11] According to Google's web page, a few hours are needed.

A quite peculiar structure, resembling a biological organism, a kind of colossal amoeba. Along with the central part (SCC, Strongly Connected Component), we find two more pieces[12]: the IN part is made up of web pages having links to those of SCC, and the OUT part is formed by pages pointed from the pages of SCC. Furthermore, there are a sort of tendrils (sometimes turning into tubes), comprising the pages not pointing to SCC's pages, nor accesible from them.

Notice, in any case, that the configuration of the web is something dynamic, evolving with time. And it is not clear whether this structure has been essentially preserved or not[13].

What should we do then? Or better: what does Google do? A standard trick: try to get the best possible situation in a reasonable way. For instance, *adding* a whole series of transition probabilities to all the vertices. That is, considering the matrix

$$\mathbf{M}'' = c\mathbf{M}' + (1-c)\begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} (1, \ldots, 1)\,,$$

where $p_1, \ldots, p_n$ is a certain probability distribution ($p_j \geq 0$, $\sum_j p_j = 1$) and c is a parameter between 0 and 1 (for Google, about 0.85).

As an example, we could choose a uniform distribution, $p_j = 1/n$ for each $j = 1, \ldots, n$ (and the matrix would have positive entries). But there are other reasonable choices, and this degree of freedom gives us the possibility of making "personalized" searches.

In terms of the random surfer, we are giving him the option (with probability $1-c$) to get "bored" of following the links and to jump to any web page (obeying a certain probability distribution).

9 Non-negative matrices in other contexts

The results on non-negative matrices we have seen in previous pages have a wide range of applications. The following two observations (see [13]) may explain their ubiquity:

[12]Researchers put forward some explanations: The IN set might be made up of newly created pages, with no time to get linked by the central kernel pages. OUT pages might be corporate web pages, including only internal links.

[13]A lot of interesting questions come up about the structure of the web graph. For instance, the average number of links per page, the mean distance between two pages, or the probability $\mathbf{P}(k)$ of a random selected page to have exactly k (say, ingoing) links. Should the graph be *random* (in the precise sense settled by Erdös and Rényi), then we would expect to have a binomial distribution (or a Poisson distribution in the limit). And we would predict that most pages will have a similar number of links. However, empirical studies suggest that the decay of the probability distribution is not exponential, but follows a *power law*, $k^{-\beta}$, where β is two something (see, for instance, [2]). This would imply, for example, that most pages have very few links, while a minority (even though very significant) have a lot of them. More than that, if we consider the web as an evolving system, to which new pages are added in succession, the outcome is that the trend gets reinforced: "the riches get richer". A usual conclusion in competitive systems (as in real life, we dare say!). We refer here to [3].

- In most "real" systems (from Physics, Economy, Biology, Technology, etc.), the measured interactions are positive, or at least non-negative. And matrices with non-negative entries are the appropriate way to encode these measurements.
- Many models involve linear iterative processes: starting from an initial state $\mathbf{x}_0$, the generic one is of the form $\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0$. The convergence of the method depends upon the size of $\mathbf{A}$'s eigenvalues, much better, upon the ratii between their sizes —particularly, between the biggest and all the others. And here is where Perron-Frobenius' theorem has something to say, as long as the matrix $\mathbf{A}$ is non-negative.

In the following subsections, we will explain, in some detail, Markov-chains based models, and briefly review some other models and extensions.

9.1 Probabilistic evolution models

Let us recall that a matrix $\mathbf{A}$ is said to be stochastic (or markovian) if $\mathbf{A} \geq 0$ and, for each column, the sum of the entries is 1. For these matrices, $\lambda = 1$ is always an eigenvalue.

> NOTE. The reason is that the $\mathbf{A}-\mathbf{I}$ columns add up to 0. So when summing all the $\mathbf{A}-\mathbf{I}$ rows we get the zero vector, and the rows are linearly dependant. This means that $\mathbf{A}-\mathbf{I}$ is singular. Thus, $\det(\mathbf{A}-\mathbf{I}) = 0$. And this implies, finally, that $\lambda = 1$ is an eigenvalue.
> Besides, it is not possible to have eigenvalues of modulus > 1, because $\mathbf{A}$ transforms probability vectors (their entries summing to 1) into probability vectors.

The probabilistic model of Markov chains is widely used in quite diverse contexts: Google's method is a nice example, but it is also used as a model for migrations, transmission of deseases, etc.

Let us consider another situation, perhaps not so familiar to the reader. Suppose we have some money to invest and that we decide not to buy stocks —some unpleasant experience might be the reason. So, in this truly conservative mood, we look for a "secure" investment: we will lend money to the State. It could be the Spanish State, or the Fed, but also Corte Inglés, Repsol, etc. There are several ways to do it: Treasury Bonds, corporate obligations, etc., although some of them are not at everybody's disposal. We might obtain a lower mean return, compared with the stock market, but we are sure we will not lose the value of our investment in exchange. Even if we are not so averse to the stock markets, we have heard around that "diversifying" is the golden rule for any investor, and this seems to be a good choice.

But things are not that easy. "Security" is not the most popular word among those that lent their money to the Argentine or Russian State, or more recently, among those who put their money in what seemed to be solid[14] firms such as Enron, Parmalat or Swiss Air.

[14] Solid, but with liquid products (Enron and Parmalat)... or gaseous! (Swiss Air).

There are no "secure" investments, although some of them are safer than others. So the markets try to determine, as precisely as possible, the **solvency** of each firm (in a wide sense, States also apply). If a certain firm is very reliable, in the sense that it will be able to pay back with very high probability the investors money (along with the corresponding interests), it might offer low returns. But if another firm is among those graded as "problematic" and it pretends to get investors, then it must offer higher returns. Where to put their money is left to the investors' discretion, depending upon their characteristics (being more or less risk averse —their speculative vein, in plain language).

This is a whole world, known as *Credit Risk*, in which the mathematical models play a central role. Let us describe one of them, namely the ratings (probabilistic evolution) model.

We have n ratings, say $S_1, \dots, S_n$. Following the jargon of the rating agencies (Standard & Poor's, Moody's, etc.), these states are AAA, BBB+, CC, etc., with that familiar sound of school marks.

After a detailed analysis, the probability of going from state S_i to estate S_j, the number a_{ij}, is determined. All this information is stored in a matrix $\mathbf{A}$, its entries being, once more, non-negative numbers. Moreover, it is a (row) stochastic matrix. A special state D (*default*) is usually included —all its entries are zeros, but the one in the diagonal; an absorbing state, in the Markov-chains language. The following matrix $\mathbf{A}$ may be an example[15]:

	AAA	*AA*	*A*	*BBB*	*BB*	*B*	*CCC*	*D*
AAA	90.58	8.36	0.84	0.09	0.13	0.00	0.00	0.00
AA	0.84	90.04	8.11	0.71	0.12	0.15	0.03	0.00
A	0.14	2.69	89.94	5.92	0.85	0.32	0.04	0.10
BBB	0.03	0.32	5.94	87.41	4.92	1.04	0.13	0.21
BB	0.02	0.12	0.62	7.49	81.93	7.88	0.92	1.02
B	0.00	0.09	0.29	0.65	6.96	81.60	4.43	5.98
CCC	0.17	0.00	0.36	1.09	2.35	10.37	64.06	21.60
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.0

For instance, there is a 90.58% probability that firms with the best rating remain with the same grade the following year.

We look at today's situation, the proportion of firms in each rating, and we build the initial vector $\mathbf{z}^{(0)} = (z_1^{(0)}, \dots, z_n^{(0)})$. According to this model, the vector of percentages will be, after k years, $\mathbf{z}^{(k)} = \mathbf{z}^{(0)}\mathbf{A}^k$.

It might be interesting, for instance, to study the *asymptotic behaviour* of the system —imagine we want to value long term investments, such as 20-30 year bonds. This means to estimate the percentage vector when $k \to \infty$, should $\mathbf{z}^{(\infty)}$ be called.

The most striking thing is that, if $\lambda = 1$ is the unique dominant eigenvalue (i.e., if the matrix is primitive), then, *no matter what the initial proportions are*, the stationary state $\mathbf{z}^{(\infty)}$ coincides with the eigenvector associated to $\lambda = 1$. Let us remember that this eigenvector is the one given by Perron-Frobenius Theory.

This have to be seen in order to be believed. So we go back to Maple again to perform the calculations for the following (column) stochastic matrix:

[15] It is the Standard & Poor's transition matrix for ratings over the 1981-96 period.

$$A := \begin{pmatrix} 0.4 & 0.1 & 0.1 & 0.2 \\ 0.4 & 0.6 & 0.3 & 0.1 \\ 0.15 & 0.2 & 0.4 & 0 \\ 0.05 & 0.1 & 0.2 & 0.7 \end{pmatrix}$$

We kindly request Maple to calculate its eigenvalues and eigenvectors:

```
> eigenvectors(A);
```

[1.0000000, 1, {[−0.37822566, −0.78653745, −0.35673556, −0.56304047]}],
[0.4999999982, 1, {[−0.2962451280, 0.4443676920, 0.4443676932, −0.5924902550]}],
[0.2292893218, 1, {[.2060703913, −0.7035683205, .6432117000, −.1457137708]}],
[0.3707106784, 1, {[−1.587596362, 0.9299924275, 1.780204107, −1.122600167]}]

Now we are no surprised at all by the output: the eigenvector associated to $\lambda = 1$ is the one and only to have same sign entries —observe Maple has displayed an eigenvector with negative entries, but you may change all its signs.

For our purposes, it is better to write it with positive entries summing to 1:

$$\mathbf{v} = (0.1814432986,\ 0.3773195881,\ 0.1711340205,\ 0.2701030929)\,.$$

Let us go now to the probabilistic evolution. We start from a certain vector $\mathbf{v}_0$ and let the matrix do the rest:

```
> v0:=matrix(4,1,[0.2,0.2,0.4,0.2]);
```

$$v0 := \begin{bmatrix} 0.2 \\ 0.2 \\ 0.4 \\ 0.2 \end{bmatrix}$$

```
> seq(evalm(A^n &* v0), n=1..20);
```

$$\begin{bmatrix} 0.18 \\ 0.34 \\ 0.230 \\ 0.250 \end{bmatrix}, \begin{bmatrix} 0.1790 \\ 0.3700 \\ 0.1870 \\ 0.2640 \end{bmatrix}, \begin{bmatrix} 0.18010 \\ 0.37610 \\ 0.175650 \\ 0.268150 \end{bmatrix}, \begin{bmatrix} 0.1808450 \\ 0.3772100 \\ 0.1724950 \\ 0.2694500 \end{bmatrix}, \begin{bmatrix} 0.18119850 \\ 0.37735750 \\ 0.171566750 \\ 0.269877250 \end{bmatrix},$$

$$\begin{bmatrix} 0.1813472750 \\ 0.3773516500 \\ 0.1712779750 \\ 0.2700231000 \end{bmatrix}, \begin{bmatrix} 0.1814064925 \\ 0.3773356025 \\ 0.1711836113 \\ 0.2700742937 \end{bmatrix}, \begin{bmatrix} 0.1814293771 \\ 0.3773264713 \\ 0.1711515388 \\ 0.2700926128 \end{bmatrix}, \begin{bmatrix} 0.1814380744 \\ 0.3773223566 \\ 0.1711403163 \\ 0.2700992527 \end{bmatrix},$$

$$\begin{bmatrix} 0.1814413475 \\ 0.3773206638 \\ 0.1711363090 \\ 0.2701016795 \end{bmatrix}, \begin{bmatrix} 0.1814425722 \\ 0.3773199980 \\ 0.1711348585 \\ 0.2701025713 \end{bmatrix}, \begin{bmatrix} 0.1814430288 \\ 0.3773197423 \\ 0.1711343288 \\ 0.2701029001 \end{bmatrix}, \begin{bmatrix} 0.1814431987 \\ 0.3773196455 \\ 0.1711341344 \\ 0.2701030214 \end{bmatrix}$$

$$\begin{bmatrix} 0.1814432617 \\ 0.3773196092 \\ 0.1711340626 \\ 0.2701030663 \end{bmatrix}, \begin{bmatrix} 0.1814432852 \\ 0.3773195957 \\ 0.1711340361 \\ 0.2701030830 \end{bmatrix}, \begin{bmatrix} 0.1814432937 \\ 0.3773195905 \\ 0.1711340263 \\ 0.2701030891 \end{bmatrix}, \begin{bmatrix} 0.1814432970 \\ 0.3773195887 \\ 0.1711340226 \\ 0.2701030913 \end{bmatrix}$$

$$\begin{bmatrix} 0.1814432982 \\ 0.3773195881 \\ 0.1711340214 \\ 0.2701030923 \end{bmatrix}, \begin{bmatrix} 0.1814432987 \\ 0.3773195878 \\ 0.1711340209 \\ 0.2701030926 \end{bmatrix}, \begin{bmatrix} 0.1814432988 \\ 0.3773195877 \\ 0.1711340207 \\ 0.2701030927 \end{bmatrix}.$$

Only 20 steps are displayed, because, as the reader might check, the convergence to the vector $\mathbf{v}$ is very fast. The reader is free to do his own calculations, starting from (almost) any initial vector and he will end with the same result. No tricks, always the same result! Mathematical Magic.

9.2 Other models and extensions

As mentioned before, Perron-Frobenius Theory also plays a central role in many other contexts (we refer the reader to [13]). Let us mention just a pair:

- *Biological models*: A well known population model, in some sense a generalization of the one developed by Fibonacci, is encoded with the so called *Leslie matrices*. Their entries are non-negative numbers, related to the transition fractions between age classes and the survival rates. If λ_1 is the dominant eigenvalue, then the system behaviour (extinction, endless growth or oscillating behaviour) depends upon the precise value of λ_1 ($\lambda_1 > 1$, $\lambda < 1$ or $\lambda = 1$ being the three cases of interest).
- *Economic models*: In 1973, Leontief was awarded the Nobel Prize for the development of his *input-output model*. A certain country economy is divided into sectors, and the basic hypothesis is that the j-th sector's input of the i-th sector's output is proportional to the j-th sector's output. In these conditions, the existence of solution for the system depends upon the value of the dominant eigenvalue of the matrix that encodes the features of the problem.

Finally, there are several extensions of Perron-Frobenius Theory the reader might find interesting:

- *Cones in* $\mathbb{R}^n$: The key point of Perron-Frobenius' theorem is that any $n \times n$ matrix with non-negative entries preserve the "positive octant". But of course, there is a general version dealing with (proper convex) cones[16] (see, for example, [1,4]).
- *Banach spaces*: Those readers versed in Functional Analysis and Spectral Theory will be aware of the generalization to Banach spaces known as the Krein-Rutman theorem (see [12] and [5]). And those engaged in Partial Differential Equations will enjoy proving, using Krein-Rutman Theorem, that the first eigenfunction of the laplacian in the Dirichlet problem (in an open, connected and bounded set $\Omega \subset \mathbb{R}^n$) is *positive* (see the details in [9, Appendix to Chapter VIII]).

[16] A set $C \subset \mathbb{R}^n$ is said to be a *cone* if $a\mathbf{x} \subseteq C$ for any $\mathbf{x} \in C$ and for any number $a \geq 0$. It will be a *convex cone* if $\lambda\mathbf{x} + \mu\mathbf{y} \in C$ for all $\mathbf{x}, \mathbf{y} \in C$ and $\lambda, \mu \geq 0$. A cone is *proper* if (a) $C \cap (-C) = \{0\}$, (b) $\mathrm{int}(C) \neq \emptyset$; y (c) $\mathrm{span}(C) = \mathbb{R}^n$.

Coda

The design of a web search engine is a formidable technological challenge. But in the end, we discover that the key point is Mathematics: a wise application of theorems and a detailed analysis of the algorithm convergence. A new confirmation of the *unreasonable effectiveness* of Mathematics in the Natural Sciences, as Eugene Wigner used to say —as in so many other fields, we might add.

We hope that these pages will encourage the readers to explore for themselves the many problems he have briefly sketched here —and hopefully, they have been a source of good entertainment.

And a very fond farewell to Perron-Frobenius' theorem, which plays a so distinguished role in so many questions. Let us do it with a a humourous (but regretfully untranslatable[17]!) *coplilla manriqueña*:

Un hermoso resultado
que además se nos revela
indiscreto;
y un tanto desvergonzado,
porque de Google desvela
. . . su secreto.

To know more

[1] Bapat, R. B. y Raghavan, T. E. S.: *Nonnegative matrices and Applications*. Encyclopedia of Mathematics and its Applications **64**. Cambridge University Press, 1997.

[2] Barabási, A.-L.: *The Physics of the web*. Physics World (julio 2001). Available at the author's web page, `www.nd.edu/~alb`.

[3] Barabási, A.-L.: *Linked, the new Science of Networks. How everything is connected to everything else and what it means*. Plume Books, 2003.

[4] Berman, A. y Plemmons, R. J.: *Nonnegative matrices in the Mathematical Sciences*. Classics in Applied Mathematics **9**. Academic Press, 1979.

[5] Brézis, H.: *Análisis funcional. Teoría y aplicaciones*. Alianza Editorial, Madrid, 1984.

[6] Brin, S. y Page, L.: *The anatomy of a large-scale hypertextual web search engine*. `www-db.stanford.edu/~sergey/`

[7] Broder, A. et al.: *Graph structure in the web*. Available at the web page: `www9.org/w9cdrom/160/160.html`.

[17] More or less: "a beautiful result, which shows itself as indiscreet and shameless, because it reveals. . . Google's secret."

[8] Haweliwala, T. S. y Kamvar, S. D.: *The second eigenvalue of the Google matrix*. Available at the following web page: `www.stanford.edu/~taherh/papers/secondeigenvalue.pdf`.

[9] Dautray, R. y Lions, J.-L.: *Mathematical Analysis and Numerical Methods for Science and Technology. Volumen 3: Spectral Theory and Applications*. Springer-Verlag, Berlin, 1990.

[10] Keener, J. P.: The Perron-Frobenius Theorem and the ranking of football teams. *SIAM Review* **35** (1993), no. 1, 80–93.

[11] Kendall, M. G.: Further contributions to the theory of paired comparisons. *Biometrics* **11** (1955), 43–62.

[12] Krein, M. G. y Rutman, M. A.: Linear operators leaving invariant a cone in a Banach space. *Uspehi Matem. Nauk (N. S.)* **3** (1948), no. 1, 3–95 [Russian]. Translated to English in *Amer. Math. Soc. Transl.* **10** (1962), 199–325.

[13] MacLauer, C. R.: The Many Proofs and Applications of Perron's Theorem. *SIAM Review* **42** (2000), no. 3, 487–498.

[14] Minc, H.: *Nonnegative matrices*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, New York, 1988.

[15] Moler, C.: *The world's largest matrix computation*. Available at the following web page: `www.mathworks.com/company/newsletters/news_notes/clevescorner/ oct02_cleve.html`

[16] Wei, T. H.: *The algebraic foundations of ranking theory*. Cambridge University Press, London, 1952.

[17] Wilf, H. S.: *Searching the web with eigenvectors*. Available at the url: `www.cis.upenn.edu/~wilf/`.