Working Paper Series
Department of Economics
University of Verona

The importance of Perron-Frobenius Theorem in ranking problems

Alberto Peretti

WP Number: 26 December 2014

ISSN: 2036-2919 (paper), 2036-4679 (online)

The importance of Perron–Frobenius Theorem in ranking problems

Alberto Peretti
Department of Economics, University of Verona
Verona, Italy
e-mail: alberto.peretti@univr.it

December, 2014

Abstract

The problem of ranking a set of elements, namely giving a "rank" to the elements of the set, may be tackled in many different ways. In particular a mathematically based ranking scheme can be used and sometimes it may be interesting to see how different can be the results of a mathematically based method compared with some more heuristic ways. In this working paper some remarks are presented about the importance, in a mathematical approach to ranking schemes, of a classical result from Linear Algebra, the Perron–Frobenius theorem. To give a motivation of such an importance two different contexts are taken into account, where a ranking problem arises: the example of ranking football/soccer teams and the one of ranking webpages in the approach proposed and implemented by Google's PageRank algorithm.

Keywords. Ranking scheme, Linear transformation, Eigenvalue, Dominant eigenvalue

AMS Classification. 15A18

JEL Classification. C65, C69

Contents

1 Introduction

Many ranking schemes have been proposed in the different situations where a ranking problem arises. A ranking scheme is a model for getting an order relation on a given finite set. For example PageRank is an algorithm used by Google Search to rank websites in their search engine results. It was developed at Stanford University by Larry Page and Sergey Brin, the founders of Google, in 1996 as part of a research project about a new kind of search engine. In this working paper we want to focus on the importance of a classical result from linear algebra, the Perron–Frobenius theorem, in the ranking problem and in particular in the PageRank method and in a football/soccer teams ranking [1,2,3,4,5].

Actually the reasons why the Perron–Frobenius theorem fits well in the ranking problem are quite straightforward and do apply more in general in each setting where a linear approach is taken to get the ranking method.

The linear approach is not the only one possible, of course, and actually situations have been discovered where an initial linear approach has appeared to be unsatisfactory. For example the football/soccer ranking belongs to this category: in [1] some reasons why this is true are provided and a more general non linear approach is preferable. The PageRank approach is linear and this is why an important step in the Page and Brin iterated algorithm does rely on the application of Perron–Frobenius theorem.

PageRank is the first algorithm that was used by the company, and it is the best-known. It has been modified during the years and combined with other methods. It is a link analysis algorithm and it assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of "measuring" their relative importance within the set. The numerical weight that it assigns to any given element is referred to as the PageRank of that element.

The idea under PageRank, contained in the Page and Brin's original paper, is to measure the importance of website pages depending on how many and how much "important" other pages are linked to them. According to Google: "*PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites.*"

In the football teams ranking things are even more straightforward: we can get the input for the ranking by taking informations from the matches that teams have played the ones against the others.

It is easy to understand why the specific scene where the ranking is applied is not so important, meaning with this for example either the website or the football/soccer teams ranking, because situations may be different but in any case we have a set of elements interacting one another with the reasonable property that someone's rank is high if it interacts with high ranked elements. Of course this is not enough on its own to get a high rank: what makes it enough may be in the football/soccer example to defeat high ranked teams in the direct match and to be pointed by high ranked websites in the web ranking example. It appears evident that the definition of the interactions and the way we translate interactions in a mathematical formula is a crucial step in the ranking scheme and this is going to deeply affect the ranking result.

2 Why the Perron–Frobenius theorem?

In [1] an interesting in-depth analysis on possible ranking methods of football teams may be found, mainly in terms of a possible mathematical background in the ranking procedure. Polls often are taken to discover what are the best teams, or what people think are the best, and newspapers publish additional indices that rank the top teams. Sometimes behind these indices there is some mathematics and the author observes that these are not understood or accepted by the general public as easily as the polls, just because they are based on mathematics.

It is evident that a good ranking scheme has a large numbers of potential users. As already noted before, the specific field is not so important: it is enough to adapt or specifically formalize the relations among the "agents" and the scheme may fit easily. Ranking schemes remove some, not all in general, subjectivity. Surely different ranking schemes may give totally different answers about who is the best, depending on the factors and aspects that are inserted and emphasized by the scheme.

As the author of [1] says, the topic can have a very important role in motivating students in the usefulness of non immediate mathematical notions: "*I use [these ranking methods] not because they solve with certainty the problems of which team is number one, but because the mathematics is fun and well motivated. These methods are excellent vehicles by which to introduce students to interesting and important mathematical ideas, including the Perron-robenius theorem, the power and inverse power methods for finding eigenvalues of a matrix, and fixed point theorems for nonlinear maps.*"

The ranking problem may be easily formulated as a linear eigenvalue problem and we give here the reason of this.

The context first: suppose there is a contest, a competition, with a number N of participants to which we want to assign a score. The score is based on the interactions with other participants and should depend on both the outcome of the interaction and strength of its rivals. In the football context the interaction is of course the match, where we have a final result saying who is the winner. In the webpage ranking there is no final winner, but we have just to find how we can interpret the interaction. In the football ranking we may state that a strong rival gives me some score even if I lost the direct match with it. In the webpage ranking it is reasonable that my page gets some score if important pages have a link to it.

In general we may define a vector of ranking values r, with positive components r_j, indicating the strength of the jth participant. The definition of the score is crucial in the model as, how it happens in general, it completely influences the model behaviour. By following [1],in the football case we may define a score for the ith participant as

$$s_i = \frac{1}{n_i} \sum_{j=1}^{N} a_{ij} r_j,$$

where N is the total number of participants in the contest, a_{ij} is a nonnegative number related to the outcome of the game between participant i and participant j, and n_i is the number of games played by participant i.

Remarks. The linear structure is evident. The ith score depends on all the other scores by means of a coefficient that characterizes the interaction between i and j. In the definition a sort of normalization is taken: the division by n_i is not important in the case all the participants play the same number of games, as in a regular soccer or football championship for example. It takes importance in case some additional games are possible. But, apart from this more technical aspect, the definition takes anyway to a classical linear model

$$s_i = \sum_{j=1}^{N} \frac{a_{ij}}{n_i} r_j = \sum_{j=1}^{N} b_{ij} r_j. \tag{1}$$

Some remarks on the coefficients a_{ij}: they are related to interactions between participants and they may take into account the different aspects of the specific field where we have the ranking problem. In the football case a possibility is to set $a_{ij} = 1$ if team i won the game against team j, $a_{ij} = \frac{1}{2}$ if the game ended in a tie and $a_{ij} = 0$ if team i lost the game against team j.[1]

For comparison, in the website ranking problem, the score that PageRank gives to sites is based on the concept of "link popularity": a certain webpage is important if, in addition to receive links from other high ranked pages, has a limited number of links to other pages. A formal representation of the concept is given by the formula

$$r(P) = \sum_{Q \to P} \frac{r(Q)}{|Q|}, \tag{2}$$

where P and Q are pages of the web and we indicate with $r(P)$ and $r(Q)$ the ranks of the pages, namely their importance, $|Q|$ is the number of external links of the page Q and the "$Q \to P$" under the summation means that the summation is extended over the pages Q that have a link to the page P.

Remark. Clearly the meaning is that for the ranking of P just the pages linked to P have relevance and the importance of these pages is reduced by the total number of links these pages have. The fewer external links a page has the better it is for the ranking of P.

We can write (2) in a way similar to (1). Let $P_1, P_2, \ldots, P_n$ are the n pages on the web. We may define the corresponding preference matrix in the following way: just rewrite (2) as

$$r(P_i) = \sum_{P_j \to P_i} \frac{r(P_j)}{|P_j|} = \sum_{P_j \to P_i} \frac{1}{|P_j|} \cdot r(P_j) = \sum_{j=1}^{N} a_{ij} r(P_j) \tag{3}$$

by setting

$$a_{ij} = \text{Prob}(P_j \to P_i) = \begin{cases} \frac{1}{|P_j|} & \text{if } P_j \to P_i \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

As indicated here the matrix A can be seen as a transition matrix A, whose elements are probabilities of a transition from a page to another page. Of course this setting assumes that if

[1]If we collect the coefficients a_{ij}, or $\frac{a_{ij}}{n_i}$ when the n_i's are relevant, we obtain a matrix, that is sometimes called the *preference matrix*.

P_j is linked to P_i, the transition from P_j to P_i is just one of the $|P_j|$ possible transitions, each one with the same probability.

It is evident that the underlying mathematical structure is the same in the two contexts. The only difference is in the coefficients, in how they are defined. Of course the fact that in the web ranking the summation is just over a part of participants is irrelevant: it means that certain coefficients are zero.

Now let's come to the reason why eigenvalues are involved and take the football ranking example first. It is very reasonable to assume that the score of a team is proportional to its strength (its rank), as defined by (1). By calling λ the proportionality constant, and indicating for simplicity with a_{ij} what was before $\frac{a_{ij}}{n_i}$, we get

$$s_i = \lambda r_i,$$

that is

$$\sum_{j=1}^{N} a_{ij} r_j = \lambda r_i.$$

Writing the last equation in matrix form we get

$$Ar = \lambda r,$$

that is the usual equation that takes to the definition of an eigenvalue of the matrix A and the corresponding eigenvector r. The meaning of this is: if we assume the linear relation among scores and the proportionality between ranks and scores, then a possible rank vector must be an eigenvector of the coefficient matrix A, the preference matrix, and the corresponding eigenvalue is the scalar for the proportion.

Remark. It may sound unexpected to assume that scores are proportional to ranks, and not just that score are ranks. From a mathematical point of view the only difference is that in the latter case we "force" the eigenvalue to be one. The former assumption gives in some way some flexibility to the model. When the matrix A is written, either it has or not the eigenvalue one, no way for us to do anything. The λ is a way to get a more robust model.

Moving to the website ranking problem and the corresponding formula (2) that states how scores are defined in PageRank algorithm, the reason why a particular eigenvalue and its corresponding eigenvector are important is going to appear now. Suppose we collect the ranks of all the webpages in a vector r and suppose we want to compute it in an iterative way, starting from a previous evaluation of r itself. If $r^{(k)}$ is the kth iteration of vector r, at the beginning the most reasonable setting, not the only possible one, is to assume that all the pages have the same rank and, taking into account a sort of normalization, we may start from

$$r^{(0)} = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)^T .^{2}$$

In the context of website ranking the concept of probability already came out quite naturally in the definition of the matrix A. It is clear that we could interpret the rank vector too in a probabilistic way. High ranks mean high probability one is connected to those pages. The initial vector $r = r^{(0)}$ may be seen as the "uniform probability distribution vector": at the beginning of the ranking process all the webpages have the same probability.

We have already written the general equation (2) in terms of the n pages $P_1, P_2, \ldots, P_n$ of the web, in particular for the ith page P_i (see equation (3)). After the first transition we have:

$$r^{(1)}(P_i) = \sum_{P_j \to P_i} \frac{r^{(0)}(P_j)}{|P_j|} = \sum_{P_j \to P_i} \frac{1}{|P_j|} \cdot r^{(0)}(P_j) = \sum_{j=1}^{n} a_{ij} \cdot r^{(0)}(P_j) \quad i = 1, 2, \ldots, n.$$

In matrix/vector form the previous equation may be written as

$$r^{(1)} = Ar^{(0)}.$$

Clearly an iterative sequence is then defined, namely

$$r^{(k+1)} = Ar^{(k)} \quad k = 0, 1, \ldots \tag{5}$$

and this is the general iteration of the so called *power method.*

Remark. In order to summarize, we have seen that in the football ranking problem the definition of the scores takes to what in Linear Algebra is the matrix equation where eigenvalues and eigenvectors come from. In this equation, that involves the vector of ranks and the preference matrix, the eigenvalue is something like a convenient proportionality constant. In the website ranking instead the definition of the ranks takes to an iterative procedure involving the rank vector and again the preference matrix. Both the mathematical problems are related to the most important among the eigenvalues of a matrix, the so called *dominant* eigenvalue.

Before going on with the study of convergence of (5) it is necessary to point out what the dominant eigenvalue is and the reason why the power method converges to it.

[2] Transposition is required in the next vector/matrix notations.

2.1 The dominant eigenvalue

Fundamental results hold regarding the dominant (or leading) eigenvalue, that is the eigenvalue with maximum module (absolute value).[3] The most famous among these results is probably the Perron–Frobenius theorem.

The dominant eigenvalue is important for many reasons, practical and theoretical. A theoretical fundamental aspect, that follows directly from the definition, is that the dominant eigenvalue gives a bound for all the eigenvalues of the matrix. In the general context of the eigenvalue definition, where they are complex numbers, the dominant eigenvalue gives the radius of the circle where all the eigenvalues are located. If we are interested in how large eigenvalues can be the straight answer is in the dominant eigenvalue.

A practical important matter is that we do not need to find all the eigenvalues in order to select the dominant one. There are methods that take us directly to the dominant eigenvalue. Moreover, as eigenvectors are important too and maybe more important than eigenvalues, these methods take us at the same time to find also the eigenvectors corresponding to the dominant eigenvalue (dominant eigenvectors). The most famous of these methods is the so called *power method*, that we are going to recall here in the following.

2.2 The power method

For the purpose of finding the dominant eigenvalue of a matrix together with a corresponding eigenvector a great variety of methods have been designed. The power method is an iterative method that takes to compute the dominant eigenvalue and a corresponding eigenvector.[4]

Let's consider a square $n \times n$ matrix A having n linearly independent eigenvectors associated

[3] The definition, in its simplicity, can easily be misunderstood. In order the dominant eigenvalue exists we want it is unique. Then there are matrices that do not have the dominant eigenvalue. If we said "the eigenvalues with largest absolute value are dominant eigenvalues" then every matrix would have some, but if we want it is unique things are different. For example a 2×2 matrix with eigenvalues 1 and -1 does not have the dominant eigenvalue. The same if a matrix has a double eigenvalue and this happens to be the largest in module.

[4] We recall that to each eigenvalue of a matrix an infinite number of corresponding eigenvectors are associated. That's why we are always forced to say "a corresponding" eigenvector. But often just one eigenvector is enough because the others are linearly dependent on that one. In particular this is the relevant case in what we are dealing with, because in the Perron–Frobenius theorem context the linear space associated to the dominant eigenvalue is one dimensional and then what we need is just one eigenvector. In the (possible) case the linear space of eigenvalues associated to a given eigenvalue is two-dimensional (or more) we need a pair (or more) of linearly independent eigenvectors associated to that eigenvalue.

to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ [5] and suppose

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|.$$

This means we are assuming that the dominant eigenvalue λ_1 has algebraic multiplicity one.[6] Let's call $v^1, v^2, \ldots, v^n$ the corresponding eigenvectors. This means

$$Av^i = \lambda_i v^i, \quad i = 1, 2, \ldots, n.$$

The power method, starting from any arbitrarily chosen vector $x^{(0)}$, builds up a sequence of vectors $\{x^{(k)}\}$ that converges to the eigenvector associated to the dominant eigenvalue.

Here is a detailed motivation for this convergence result. Suppose $x^{(0)}$ is an arbitrary vector in $\mathbb{R}^n$. We may write $x^{(0)}$ as a linear combination of $v^1, v^2, \ldots, v^n$ that, because of the hypothesis of linear independence, are a basis of $\mathbb{R}^n$.

$$x^{(0)} = \sum_{i=1}^{n} c_i v^i \quad \text{and suppose } c_1 \neq 0.^7$$

Starting from $x^{(0)}$ we may build up the sequence

$$x^{(1)} = Ax^{(0)} \quad , \quad x^{(2)} = Ax^{(1)} \quad , \quad \ldots \quad , \quad x^{(k)} = Ax^{(k-1)} \quad , \quad \ldots$$

The following result holds:

Theorem 1 *For the sequence $\{x^{(k)}\}_{k\in\mathbb{N}}$ we have that*

$$\lim_{k\to\infty} \frac{x_j^{(k+1)}}{x_j^{(k)}} = \lambda_1 \quad \textit{and} \quad \lim_{k\to\infty} \frac{x^{(k)}}{x_j^{(k)}} = cv^1. \tag{6}$$

where j is an index for which $x_j^{(k)} \neq 0$, for every value of k.

[5] This is a crucial hypothesis: from the theoretical point of view it means that the matrix A is similar to a diagonal matrix or, in an equivalent way, that the linear mapping associated to A is diagonal in a convenient system of coordinates. As not all matrices have this property, the hypothesis is relevant. From a practical point of view the aspect is even more crucial: the use of the power method is fundamental in the PageRank algorithm and this means that if we do not know the hypothesis is true or false we cannot be sure the convergence of the method gives us a significant result. We could have non convergence at all or we could have convergence to something that is not reliable as the dominant eigenvalue. The problem is that the "matrix of the web", the one that has to be used by PageRank, is an enormous matrix and it is impossible to perform an a priori numerical control testing the existence of n independent eigenvectors.

[6] As we shall see shortly this is "part" of the Perron–Frobenius statement, specifically the property is part of the thesis of the theorem. We can rely on the validity of the assumption if the Perron–Frobenius hypotheses are true.

[7] The condition means that we don't have to start from a point in the subspace spanned by the eigenvectors $v^2, v^3, \ldots, v^n$. We need $x^{(0)}$ to have a component in the subspace spanned by v^1.

Proof 1 *From the definition of the sequence* $\{x^{(k)}\}$ *we have that*

$$x^{(k)} = Ax^{(k-1)} = A^2x^{(k-2)} = \ldots = A^kx^{(0)} = A^k\sum_{i=1}^{n} c_iv^i = \sum_{i=1}^{n} c_iA^kv^i = \sum_{i=1}^{n} c_i\lambda_i^kv^i.$$

By collecting λ_1^k *we get*

$$x^{(k)} = \lambda_1^k\left(c_1v^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^k v^i\right)$$

and then

$$x^{(k+1)} = \lambda_1^{k+1}\left(c_1v^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} v^i\right).$$

Then for the indices j *for which* $x_j^{(k)} \neq 0$ *and* $v_j^1 \neq 0$ *we may write*

$$\frac{x_j^{(k+1)}}{x_j^{(k)}} = \lambda_1\frac{c_1v_j^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} v_j^i}{c_1v_j^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k} v_j^i}. \tag{7}$$

As $\left|\frac{\lambda_i}{\lambda_1}\right| < 1$ *for* $2 \le i \le n$*, we have*

$$\lim_{k\to\infty}\frac{x_j^{(k+1)}}{x_j^{(k)}} = \lambda_1.$$

Let's consider now the sequence of vectors $\left\{\frac{x^{(k)}}{x_j^{(k)}}\right\}$*, taking again, for each value of* k*, an* $x_j^{(k)}$ *component that is non zero. Then*

$$\frac{x^{(k+1)}}{x_j^{(k)}} = \lambda_1\frac{c_1v^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} v^i}{c_1v_j^1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k} v_j^i}$$

and again, taking the limit for $k \to \infty$*, we get*

$$\lim_{k\to\infty}\frac{x^{(k+1)}}{x_j^{(k)}} = \frac{v^1}{v_j^1},$$

that is a "normalization" of the eigenvector v^1*.*

Remark. The eigenvalue can be obtained also as the limit of a different sequence, namely as the

$$\lim_{k\to\infty}\frac{x^{(k)^T}Ax^{(k)}}{x^{(k)^T}x^{(k)}},$$

that is the limit of the Rayleigh quotients of the sequence $x^{(k)}$.

Remark. It is worthwhile to specify that in the practical implementations of the method some numerical problems are likely to arise. The method, if implemented as it has been presented, gives overflow/underflow problems. For this reason at each step it is convenient to normalize the vector $x^{(k)}$. The properties of the convergence are not modified and we prevent the norms becoming too large.

We have seen that the limit of the sequence is an eigenvector associated to the dominant eigenvalue. The ranks of the webpages are the components of this limit vector, conveniently normalized.

Remark. A couple of further immediate remarks. Firstly it appears evident the extreme heaviness of the computations in the website ranking problem. The number of webpages exceeds a couple of billions and this is the length of the rank vector r and the order of the matrix A.

A more theoretical aspect is that the existence and the properties of the dominant eigenvalue depend on some properties of the matrix A, and for the moment we do not know if the matrix satisfies those properties. We shall go into these details in a later work.

2.3 The Perron–Frobenius theorem

In a ranking problem the Perron–Frobenius theorem is important because it gives us some sufficient conditions in order the problem to have a solution. Conditions are on the matrix A clearly. It is like to have properties that guarantee we can find, given the assumptions on the relations among participants, a proper ordering, a ranking, in the set.[8]

The Perron–Frobenius theorem is a classical "articulated" and somehow technical theorem. It requires some technical definitions about matrices and it can be found stated in many different equivalent ways. A very simple form may be the following:

Theorem 2 (Perron-Frobenius) *If the matrix A has nonnegative entries, then there exists an eigenvector r with nonnegative entries, corresponding to a positive eigenvalue ρ. Furthermore, if the matrix A is irreducible, the eigenvector r has strictly positive entries and the corresponding eigenvalue ρ is unique and simple and is the largest eigenvalue of A in absolute value.*

Remarks. A vector or matrix with nonnegative entries may be said a nonnegative vector or nonnegative matrix (we write $v \geq 0$ to say that v is a nonnegative vector) and similarly a vector or matrix with positive entries may be said a positive vector or positive matrix (we write $v > 0$ to say that v is a positive vector). We recall that these definitions give the possibility to obtain

[8]From a mathematical point of view ordering and ranking are actually different. An *order* in a set does not mean to give a "value" to each element of the set. It just mean to be able to compare any two elements (*total orders*). In the ranking problem our aim is to give a score to teams, having the possibility to appreciate how they are separated in the ranking. To get a ranking appears to be more difficult than to have an ordering.

a *partial order* in the set of nonnegative vectors by saying that, given two vectors v and u, $v \geq u$ whenever $v - u \geq 0$ and $v > u$ whenever $v - u > 0$. It is known that the order is not a *total order* because there are pairs of vectors that are non comparable, in the sense neither $v \geq u$ nor $u \geq v$ is true.

The definition of when a matrix is *irreducible* is quite technical and in the literature there is a great number of equivalent characterizations. We are giving some remarks in the next subsection.

Traditionally, and historically, the theorem may be given in two statements, one (the first) for positive matrices and the other for nonnegative matrices. Actually the first form is due to Perron. Frobenius, while trying to extend the theorem in the more general case of nonnegative matrices, obtained the second form, where the thesis is again true but with the need of the further assumption that the matrix is irreducible.

A possible statement of the Perron's theorem is the following. In addition to the main result, many other properties may be proved in the assumption. We give here some of the most known.

Theorem 3 (Perron's theorem for positive matrices) *Let A be an $n \times n$ positive matrix. Then the following statements hold.*

(i) There is an eigenvalue ρ (the so called Perron-Frobenius eigenvalue), that is real and positive, and for any other eigenvalue λ we have $|\lambda| < \rho$.

(ii) ρ is a simple eigenvalue, that is a simple root of the characteristic polynomial. In other words its algebraic multiplicity is one. As a consequence the eigenspace V_ρ associated with ρ is one-dimensional.

(iii) There exists a positive eigenvector r associated with ρ. Respectively, there exists a positive left eigenvector s.

(iv) There are no other positive eigenvectors of A, except (positive) multiples of r (respectively, left eigenvectors except (positive) multiples of s).

(v) $\displaystyle\lim_{k\to+\infty} \left(\frac{A}{\rho}\right)^k = rs^T$*, where the right and left eigenvectors are normalized, so that $s^T r = 1$. Moreover, the matrix rs^T is the projection onto the eigenspace V_ρ, the so called Perron projection.*

(vi) The Perron-Frobenius eigenvalue ρ satisfies the inequalities

$$\min_i \sum_j a_{ij} \leq \rho \leq \max_i \sum_j a_{ij}.$$

Remark. We recall that a positive (real) matrix is not guaranteed to have just real eigenvalues and these may be complex numbers in general. The absolute value is then intended in the complex field.

Here is a possible statement of the Frobenius theorem. It is usual to define as the *spectral radius* of a matrix A the maximum of the absolute values of its eigenvalues. In addition to the main result, here again we recall some other properties that can be proved in the assumptions.

Theorem 4 (Frobenius theorem for nonnegative irreducible matrices) *Let A be a non-negative irreducible $n \times n$ matrix with period p and spectral radius ρ. Then the following statements hold.*

(i) *ρ is positive and it is an eigenvalue of the matrix A, called the* Perron-Frobenius eigenvalue.

(ii) *ρ is simple. Both right and left eigenspaces associated with ρ are one-dimensional.*

(iii) *A has an eigenvector r and a left eigenvector s associated with ρ, whose components are positive for both and the only eigenvectors with all positive components are the ones associated with ρ.*

(iv) *The matrix A has exactly p (the period) complex eigenvalues with module ρ. Each of them is a simple root of the characteristic polynomial and is the product of ρ with a pth complex root of the unity.*

(v) $\lim_{k \to +\infty} \left(\frac{A}{\rho} \right)^k = rs^T$, *where the right and left eigenvectors are normalized, so that $s^T r = 1$. Moreover, the matrix rs^T is the projection onto the eigenspace V_ρ, the so called Perron projection.*

(vi) *The Perron-Frobenius eigenvalue ρ satisfies the inequalities*

$$\min_i \sum_j a_{ij} \leq \rho \leq \max_i \sum_j a_{ij}.$$

Remarks. The thesis cannot be obtained without the hypothesis that A is *irreducible*. The thesis is quite similar to the one we had before in the positive case. The only important difference is in *(iv)*: there is just one real dominant eigenvalue, but there are complex eigenvalues with the same absolute value, and the theorem states a complete description of those.

We have seen before why eigenvalues and eigenvectors are important in a ranking scheme. Now we have to state why in particular the Perron–Frobenius theorem is involved.

For the solution of the ranking problem eigenvalues are actually not so important. Inside the problem formulation the eigenvalue is just a proportionality constant that relates the score and the rank. The important part in the solution is the existence of the ranking vector, and of course ranks must be positive (nonnegative) numbers. Then the importance relies on having a positive (nonnegative) solution and this is the crucial step. The theorem, as we have seen here above, is about a positive eigenvalue and a corresponding positive eigenvector.

Clearly the uniqueness aspect is important too: it could be meaningless to end up with two possible rankings, maybe in contrast the one with the another.

Finally, the non uniqueness of the solution in the nonnegative case (see the Remarks following the theorem) is quite theoretical, because it is due to complex solutions. In a practical need of preparing a ranking scheme complex solutions are not interesting.

It is worthwhile to give some more details on the concept of irreducible matrices, given the importance it has in the matter: just think that in order to apply the positive version of Perron–Frobenius theorem, the one that does not need further assumptions, we must have a preference matrix A with just positive elements. In the examples we presented so far we do not have this, because in the football example both a victory or a defeat in a match take to a zero element in the matrix and the same happens in the website ranking example, where zero elements easily appear because webpages are not linked to all the webpages of the WWW for sure.

2.4 Irreducible matrices

There are many equivalent ways to characterize irreducible matrices. Some ways are algebraic ones and some are geometric. In [6] we report a set of characterizations and we briefly outline also the interesting relation between irreducible matrices and the graph theory. Sometimes it is easier to decide if the matrix is irreducible by giving a look at the corresponding graph.

This property, unlike the property of being nonnegative, that is evident, may not be evident at all. It is often related to the existence of something having some property in turn. Here are two possible conditions that have a simple meaning in ranking problems.

(i) A matrix A is irreducible if there is no permutation that transforms A into a block matrix of the form

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

where A_{11} and A_{22} are square matrices.

(ii) A nonnegative matrix A is irreducible if for any vector $r \geq 0$ we have $Ar > 0$.

Remarks. Both conditions have both an algebraic and a geometrical meaning. Clearly the conditions are far to be easy to use conditions, mainly if the matrix is a big one, as it is in practical problems. The two conditions can be seen in the ranking problem context. For example, in the case of football ranking, having defined the matrix as we did before (a_{ij} is 1 or 0 in case of loss or win) with condition (i) the matrix A is irreducible if we cannot partition the set of teams in two subsets S_1 and S_2 such that no team in S_1 plays any team in S_2 or every game between one team from S_1 and one team from S_2 resulted in a victory for the team in S_1. It is also interesting to observe that for this preference matrix to be irreducible there can be no winless teams.

With reference to the connections between the matrix and the associated graph (see [6]) it is easy to notice that condition (i) is equivalent to have a connected graph. In order to see the graph–matrix relation it is maybe more natural to take the website example, where connections among pages on the web correspond to non null elements in the matrix. Then a block of null elements means non connection between subsets of sites, in other words it means a disconnected graph. And of course the first thought one can have is that the web matrix is far to be associated to a connected graph. This in fact has been a problem for the designers of the PageRank method, who arranged strategies to modify the web matrix in order to obtain an irreducible matrix.

Condition (ii) may again be analysed with reference to the website example. If we interpret the elements a_{ij} of the preference matrix A as probabilities of transition from page j to page i (see (4)), and in the same way we give the r_j's the meaning of probability to be in the webpage P_j, then the ith component of vector Ar is

$$\sum_{j=1}^{N} a_{ij} r_j$$

that, starting from a probability distribution r, gives the probability of being in webpage P_i after considering all the interrelations. Then (ii) says that for every possible non trivial probability distribution, after considering the interrelations every page has a strictly positive probability. A similar interpretation holds if we take the r_i's in terms of rankings: every page gets a positive ranking after considering the links it receives from other pages whatever its initial score is. It is clear that condition (ii) has something to do with connection in the graph.

3 Conclusions and further studies

We gave an outline on how the Perron–Frobenius theorem is involved in the solution of a ranking scheme problem, taking a couple of examples from the football teams ranking and the websites

Google's PageRank method. There are many aspects behind the scene, some of them quite technical from the mathematical point of view.

For a future work it may be interesting to go on the details of these aspects. First of all how the choice of the preferences can affect the solution of the ranking problem and in particular the property of the matrix A of being irreducible. Probably this is a hard task if faced in general terms, and possibly an initial approach by means of some numerical simulation can be interesting in itself and can give some hints about what are the important properties.

Secondly, as far as it is possible to get the crucial informations from the web, how PageRank designers tackled the problem of a non irreducible matrix.

And finally, after a look at the literature, what are the possible results one can find by applying the power method in the case of a non irreducible matrix.

References

[1] J. P. Keener, The Perron–Frobenius theorem and the ranking of football teams, SIAM Review, Vol. 35, No. 1, 1993.

[2] The PageRank Algorithm, `http://pr.efactory.de/e-pagerank-algorithm.shtml`

[3] I. Rogers, The Google Pagerank Algorithm and How It Works, IPR Computing Ltd., `http://www.cs.princeton.edu/~chazelle/courses/BIB/pagerank.htm`

[4] PageRank Algorithm - The Mathematics of Google Search, `http://www.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html`

[5] David Austin, How Google Finds Your Needle in the Web's Haystack, Grand Valley State University, `http://www.ams.org/samplings/feature-column/fcarc-pagerank`

[6] A. Peretti, A. Roveda, On the mathematical background of Google PageRank algorithm, Working Paper Series, Department of Economics, University of Verona, 2014.

[7] G. H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, 1996.

[8] R. Horn, C. A. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[9] R. H. Horn, C. A. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.

[10] PageRank, at `https://it.wikipedia.org/wiki/PageRank` and related pages