

# Notes for Research Project on Ranking

Sinan Aksoy, Fan Chung, Olivia Simpson

June 19, 2014

# Contents

<b>1</b>	<b>Background</b>	<b>3</b>
1.1	Relevant Matrices . . . . .	3
1.1.1	Positive matrices . . . . .	3
1.1.2	Positive reciprocal matrices . . . . .	4
1.1.3	Positive consistent (reciprocal) matrices . . . . .	4
1.2	Two ranking methods . . . . .	5
1.2.1	Perron eigenvalue method . . . . .	5
1.2.2	Logarithmic least squares method . . . . .	6
<b>2</b>	<b>Ranking with Offense and Defense</b>	<b>8</b>
2.1	Introduction . . . . .	8
2.2	Our methods . . . . .	8
2.2.1	Pairwise dominance . . . . .	8
2.2.2	A new ranking algorithm . . . . .	9
<b>3</b>	<b>PageRank and Heat Kernel PageRank</b>	<b>11</b>
3.1	PageRank . . . . .	11
3.2	Heat Kernel PageRank . . . . .	11

# Chapter 1

## Background

Here, we give some background on our problem. We start by listing some of the basic properties of the matrices we will be considering. Then, we briefly consider two popular ranking methods: the Perron eigenvalue method and the logarithmic least squares method.

### 1.1 Relevant Matrices

#### 1.1.1 Positive matrices

**Theorem 1.1.1** (Perron-Frobenius [1] [3]). *Let  $A > 0$ . Then:*

1. *There is a positive (column) vector  $w$  such that  $Aw = \rho(A)w$ . We call  $w$  the right Perron vector of  $A$ . Similarly, there is a positive (row) vector  $v$  such that  $vA = \rho(A)v$ . We call  $v$  the left Perron vector of  $A$ .*
2. *The right Perron vector  $w$  is orthogonal to all other eigenvectors in right eigenspace; similarly,  $v$  is orthogonal to all other eigenvectors in left eigenspace.*
3.  *$\rho(A)$  is an algebraically (and hence geometrically) simple eigenvalue of  $A$ .*
4.  *$\rho(A)$  is the unique eigenvalue of maximum modulus.*
5.  *$\rho(A)$  is bounded above by the maximum row sum of  $A$  and from below by the minimum row sum of  $A$ .*

As an immediate corollary, we can state limiting behavior of positive matrices  $A$  in terms of the right and left Perron vectors:

**Corollary 1.1.2** ([3]). *Let  $w$  and  $v$  be the right and left Perron vectors of  $A > 0$ . Assume  $w$  and  $v$  are scaled so that  $vw = 1$ . Then:*

$$\lim_{m \rightarrow \infty} \left( \frac{A}{\rho(A)} \right)^m = vw$$

Lastly, we note:

**Theorem 1.1.3** ([1]). *Let  $A > 0$ . The largest eigenvalue of  $A$ ,  $\rho(A)$ , increases as any element  $a_{ij}$  increases.*

### 1.1.2 Positive reciprocal matrices

**Definition 1.1.4.**  $A$  is a positive reciprocal matrix if  $a_{ij} > 0$  and  $a_{ji} = \frac{1}{a_{ij}}$ .

**Theorem 1.1.5** ([1]). *The eigenvalues of a positive reciprocal matrix satisfy:*

$$\sum_{\substack{j,k \\ j \neq k}} \lambda_j \lambda_k = 0$$

### 1.1.3 Positive consistent (reciprocal) matrices

**Definition 1.1.6.** A matrix  $A$  is said to be consistent if  $a_{ij}a_{jk} = a_{ik}$  for all  $i, j, k$ . In other words, a positive reciprocal matrix is consistent when pairwise dominance relations are transitive.

**Comment:** Using the above definition, along with the fact that  $a_{ji} = 1/a_{ij}$ , the consistency condition becomes  $a_{ij}a_{jk}a_{ki} = 1$ . So, one can also define consistency in a graph theoretic sense: if we think of our matrix  $A$  as describing a weighted adjacency matrix, then consistency means that the product of the edge weights in any 3-cycle is 1.

**Comment:** Note that all positive consistent matrices are necessarily reciprocal matrices. Assuming  $A$  is positive, the definition of consistency implies  $a_{ii} = 1$ , which, in turn, implies  $a_{ij} = a_{ii}/a_{ji} = 1/a_{ji}$ . Thus, a positive, consistent matrix is necessarily a consistent, positive reciprocal matrix (but the converse is not necessarily true).

**Fact:** If  $A$  is a consistent positive reciprocal matrix, then the right and left Perron vectors of  $A$ ,  $w$  and  $v$ , are reciprocals of each other up to a multiplicative factor. That is, there exists  $c \in \mathbb{R}$  such that:

$$w_i = c \left( \frac{1}{v_i} \right)$$

Next, we note that consistency can be characterized in terms of the largest eigenvalue of a reciprocal matrix:

**Theorem 1.1.7** ([1]). *Let  $A$  be a positive reciprocal matrix. Then  $A$  is consistent if and only if  $\rho(A) = n$ .*

Consistent, positive reciprocal matrices also satisfy:

**Theorem 1.1.8** ([1]). *Let  $A$  be a consistent, positive reciprocal matrix. Then:*

$$A^k = n^{k-1}A$$

A natural question that arises is how to measure consistency in matrices that aren't consistent. Several metrics have been proposed - we mention a popular "consistency index" below defined by Saaty [1]:

**Definition 1.1.9.** The Consistency Index (C.I.) of a positive reciprocal matrix  $A$  is defined as:

$$\text{C.I.} = \frac{\rho(A) - n}{n - 1}$$

**Open Conjecture:** Let  $A \in M_n$  be a positive, reciprocal matrix with  $n \geq 4$ . The right and left Perron vectors of  $A$  are reciprocals of each other (up to a multiplicative factor) if and only if  $A$  is consistent. (Note: for  $n = 2$ , any positive reciprocal matrix is consistent. For  $n = 3$ , it has been proven that regardless of consistency, the right and left eigenvectors are reciprocals of each other. As far as I can tell, the conjecture is still unproven for  $n \geq 4$ .)

**Comment:** While it may not be helpful, we note the above conjecture can be alternatively stated in the language of graph theory: Let  $A$  be the weighted adjacency matrix of the complete weighted digraph on  $n$  vertices for  $n \geq 4$ . The products of the edge weights of every 3-cycle is 1 if and only if the right and left Perron vectors of  $A$  are reciprocals (up to a multiplicative factor) of each other.

## 1.2 Two ranking methods

### 1.2.1 Perron eigenvalue method

Saaty advocates use of the Perron eigenvalue method in which the ranking is given by the Perron vector of  $A$ . Saaty makes two key points to justify this method. The first is that, when  $A$  is consistent, the Perron vector  $w$  satisfies:

$$a_{ij} = \frac{w_i}{w_j}$$

This can be easily proved. Since every elements of  $A$  can be determined from the first row of  $A$ ,  $A$  has rank one and exactly one nonzero eigenvalue. Using the aforementioned fact that  $A^2 = nA$ , and denoting the columns of  $A$  as  $[a_1, a_2, \dots, a_n]$  we can see:

$$\begin{aligned} A^2 &= A[a_1, a_2, \dots, a_n] \\ &= [Aa_1, Aa_2, \dots, Aa_n] \\ &= [na_1, na_2, \dots, na_n] \\ &= nA \end{aligned}$$

which tells us the columns of  $A$  are all scalar multiples of the (unique) dominant eigenvector. So, for column  $a_k$ , the Perron vector  $w$  satisfies, for some scalar  $c$ :

$$\begin{aligned} w_i &= ca_{ik} \\ w_j &= ca_{jk} \end{aligned}$$

which, if we take the ratio and use the consistency definition, yields the desired result:

$$a_{ij} = \frac{w_i}{w_j}$$

The second key justification Saaty makes in favor of using the Perron eigenvalue method concerns perturbation. Namely, he shows that small changes in the entries of the matrix result in

small changes in the Perron vector. While these two arguments seem to be the most compelling, Saaty and others also have other arguments, including empirical arguments, on why the Perron vector should be used in ranking.

### Asymmetry in rankings derived from right and left Perron vectors

While Saaty and others have traditionally used the right Perron vector for ranking, there seems to be no justification for why the left eigenvector should not be used in place of the right eigenvector. Larger values in the right Perron vector signify a higher ranking for that item whereas smaller values in the left Perron vector signify a higher ranking for that item. In the case where  $A$  is consistent, given the above mentioned reciprocal relationship between the right and left Perron vectors, both vectors will lead to the same ranking. However, if  $A$  is inconsistent, then the right and left Perron vectors can lead to different rankings. In a recent survey of the Analytic Hierarchy Process, the author listed this issue as the largest theoretical dispute in the field. We give an example below from [4] to that illustrates the rank reversal issue:

$$A = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 1 & 3 & 1/3 & 1/2 \\ 1/3 & 1 & 1/6 & 2 \\ 3 & 6 & 1 & 1 \\ 2 & 1/2 & 1 & 1 \end{pmatrix} \end{matrix}$$

The right Perron vector is:  $w = [0.184; 0.152; 0.436; 0.227]$ , which leads to the ranking:  $c > d > a > b$  while the left Perron vector is  $v = [0.248, 0.338, 0.105, 0.259]$ , which leads to the different ranking of  $c > a > d > b$ .

### 1.2.2 Logarithmic least squares method

The Perron vector is not the only continuous vector-valued function of positive reciprocal matrices that yields the correct scale when the matrix is consistent. Another popular method that also meets these criteria, introduced in [5], is the geometric mean method, or more commonly known as logarithmic least squares method (LLSM). This approach minimizes the multiplicative error; that is, if  $p$  denotes the LLSM vector, then this method aims to minimize  $e_{ij}$  below:

$$a_{ij} = \frac{p_i}{p_j} e_{ij}$$

That is, LLSM assumes the multiplicative error and aims to minimize the sums of these errors:

$$\sum_{i,j=1}^n \left( \ln a_{ij} - \ln \frac{p_i}{p_j} \right)^2$$

The normalized LLSM vector is given by:

$$p_i = \frac{\left(\prod_{j=1}^n a_{ij}\right)^{1/n}}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij}\right)^{1/n}}$$

The list of relative pros and cons of Perron eigenvalue method versus the LLSM remains long and is a topic of debate.

## Chapter 2

# Ranking with Offense and Defense

### 2.1 Introduction

Consider a network of players in a series of one-on-one competitions. Suppose the resources for these competitions are limited so that we do not expect round-robin style matches. For such a series of unevenly matched competitions we are interested in an overall ranking of the players. In particular, we want to avoid giving preference to an undefeated player if they played fewer games, or if they played only lower ranked players. We are interested in using a “pairwise dominance” relationship between each pair of players to compute a total ranking of the players.

The procedure for organizing the outcome data into a collection of pairwise dominance relationships results in a set of pairwise scores  $s_{uv}$  describing player  $u$ 's relative strength over player  $v$ . These scores can be organized in a matrix  $(S)_{uv} = s_{uv}$ .

Many algorithms have been proposed for summarizing the pairwise dominance scores into a single definitive ranking of all players. Ranking by the Perron vector was popularized by Saaty and is based on the following definition of ranking. Let  $r = \{r_i\}_{i=1}^n$  be an absolute score assignment over the  $n$  players in the network. Then  $r$  is a good ranking if  $r(u)/r(v)$  is close the pairwise dominance of player  $u$  over player  $v$ ,  $s_{uv}$ . Other single dimensional rankings have been proposed (LLSM, etc).

Given the natural asymmetry of the problem, however, we propose a new ranking which takes into account two “scores.” We will make these notions more precise in the next section.

### 2.2 Our methods

#### 2.2.1 Pairwise dominance

The first task is organizing a collection of outcomes into the set of pairwise dominance scores  $s_{uv}$ . Suppose players  $u$  and  $v$  played  $\alpha$  games against each other and that of these  $\alpha$  games, player  $u$  won  $w_u$  games and player  $v$  won  $w_v$  games. Then we will use the arithmetic mean of score differences of [6], and define pairwise dominance  $s_{uv}$  as such:

$$(2.2.1) \quad S(u, v) = s_{uv} = \frac{w_u - w_v}{\alpha}.$$

This can be modeled by a complete directed graph, where the edge  $(u, v)$  is weighted by the



score  $s_{uv}$ . Then the matrix  $S$  is the adjacency matrix of the network. We will refer to  $S$  as the *dominance matrix* of the network.

### 2.2.2 A new ranking algorithm

Given a dominance matrix  $A$  of a network, the goal is to construct a single dimensional hierarchy of players in the network, called a ranking. As mentioned, one possible ranking is a vector  $r$  such that  $r(u)/r(v)$  is a good approximation of  $S(u, v)$ . This intuitively suggests that, as the  $(u, v)^{th}$  entry in  $S$  represents the relative strength of member  $u$  to member  $v$ , an appropriate ranking should respect this relationship.

In this work, we suggest a new notion of ranking. Namely, we generalize the idea of dominance and consider facets of strength: offensive strength and defensive strength.

**Definition 2.2.1.** Let  $S$  be a dominance matrix of a network. Then we say that the network has a *total hierarchical ranking* if there exist vectors  $\omega, \delta : V \rightarrow \mathbb{R}$  such that

$$(2.2.2) \quad S(u, v) = c(\omega(u)\delta(v) - \delta(u)\omega(v)) + \epsilon_{uv},$$

for some constant  $c$  and some allowable error  $\epsilon_{uv}$ . We will refer to  $\omega$  and  $\delta$  as the *offensive* ranking and *defensive* ranking, respectively.

---

#### Algorithm 1 TotalHierarchicalRanking( $S$ )

---

input: a dominance matrix  $S$

output: offensive ranking vector  $\omega$ , defensive ranking vector  $\delta$

```

 $\lambda = ir \leftarrow$  largest eigenvalue of  $S$ 
 $f = a + ib \leftarrow$  right eigenvector of  $S$  corresponding to  $\lambda$ 
 $\omega \leftarrow a, \delta \leftarrow b$ 
return  $\omega, \delta$ 

```

---

**Theorem 2.2.2.** Let  $G$  be a network with a total hierarchical ranking and let  $S$  be the dominance matrix of  $G$ . Then the vectors  $\omega, \delta$  returned by TotalHierarchicalRanking( $S$ ) are offensive and defensive ranking vectors.

*Proof.* We first note that the matrix  $S$  is skew-symmetric, i.e.,  $S = -S^T$ , by our definition of pairwise dominance (2.2.1). Thus, the eigenvalues of  $S$  are all of the form  $\lambda = \pm ir$  for  $r \in \mathbb{R}$ , and  $f$  must be a complex vector of the form  $f = a + ib$ . Let's consider the spectral decomposition of  $S$ ,

$$(2.2.3) \quad S = \sum_j \lambda_j f_j f_j^*,$$

where  $f^*$  denotes the conjugate transpose. Each term of the series is given by

$$\begin{aligned}
(\lambda f f^*)(u, v) &= ir((a + ib)(u)(a - ib)(v)) \\
&= ir(a(u)a(v) + b(u)b(v) - ia(u)b(v) + ib(u)a(v)) \\
&= r(a(u)b(v) - b(u)a(v)) + ir(a(u)a(v) + b(v)b(v)).
\end{aligned}$$

With this, (2.2.3) can be rewritten:

$$S(u, v) = \sum_j r_j (a_j(u)b_j(v) - b_j(u)a_j(v)) + ir_j (a_j(u)a_j(v) + b_j(v)b_j(v)).$$

Since eigenvalues come in conjugate pairs, we can simplify (2.2.3) using the above

$$\begin{aligned}
(2.2.4) \quad S(u, v) &= \sum_{j=1}^{n/2} \lambda_j f_j f_j^* - \lambda_j \bar{f}_j f_j^T \\
&= \sum_{j=1}^{n/2} 2r_j (a_j(u)b_j(v) - b_j(u)a_j(v)) \\
(2.2.5) \quad &= 2r_1 (a_1(u)b_1(v) - b_1(u)a_1(v)) + \sum_{j=2}^{n/2} 2r_j (a_j(u)b_j(v) - b_j(u)a_j(v))
\end{aligned}$$

where (2.2.4) is a result of the imaginary parts cancelling. Since we assume  $G$  has a total hierarchical ranking, we know there exist some vectors  $\omega, \delta$  such that (2.2.2) holds, so if we show that  $\sum_{j=2}^{n/2} 2r_j (a_j(u)b_j(v) - b_j(u)a_j(v))$  is reasonably small, we may choose  $\omega = a$  and  $\delta = b$  to complete the proof. □

## Chapter 3

# PageRank and Heat Kernel PageRank

(This chapter is forthcoming)

### 3.1 PageRank

**Definition 3.1.1.**

$$\alpha \sum_{k=0}^{\infty} (1 - \alpha)^k \frac{A^k e}{e^T A^k e}$$

### 3.2 Heat Kernel PageRank

**Definition 3.2.1.**

$$e^t \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{A^k e}{e^T A^k e}$$

# Bibliography

- [1] T. Saaty, *The analytic hierarchy process*, McGraw-Hill, New York, 1980.
- [2] T. Saaty, Rank according to Perron: a new insight, *Mathematics Magazine*, **60**(4) (1987) 211-212.
- [3] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1999.
- [4] C. Johnson, W. Beine, T. Wang, Right-left asymmetry in an eigenvector ranking procedure, *Journal of Mathematical Psychology*, **19** (1979) 61-64.
- [5] G. Crawford and C. Williams, A note on the analysis of subjective judgement matrices, *Journal of Mathematical Psychology* **29** (1985) 387-405
- [6] Xiaoye Jiang and Lek-Heng Lim and Yuan Yao and Yinyu Ye, Statistical ranking and combinatorial Hodge theory, *Mathematical Programming*, **127** 1, (2011) 203–244.