MATRIX BASED RANKING METHOD WITH APPLICATION TO TENNIS

ORIGINAL PAPER BY: GEIR DAHL - UNIVERSITY OF OSLO
PUBLISHED IN: FEBRUARY 2012

08 April 2019

Summary by –
Abhishek Rawat
Pushkar Sathe
Ankush Karmarkar

Chennai Mathematical Institute

EARLY METHODS OF RANKING AND MATRIX BASED RANKING SYSTEMS

- In early 90s a web page was ranked according to maximum number of matches with the words in search on Google / any search engine
- If a page had CMI written a million times and no other useful information then the web page www.cmi.ac.in would not be at the top of search list
- So the web page ranking system needed a change. This change was brought by measuring popularity of a page (probability of a random surfer on internet visits a page)
- Reaching a web page from another web page can be considered as edge on a graph and weights of all the edges from different web pages together form a **vector** for web page
- Ranking web pages based on matrix of vectors of all web pages that match search words of a user is an example how newer ranking methods are applied

CURRENT TENNIS RANKING METHODOLOGY BY ATP(ASSOCIATION OF TENNIS PROFESSIONALS)

- Currently ATP ranks based on points scored by a player through points scored by winning matches in Grand slams, ATP Tour tournaments etc.
- Player with highest number of points in a year gets rank one and the least gets the last rank
- Below is a table that shows how points can be earned by a player to be ranked higher:

Category	W	F	SF	QF	R16	R32	R64	R128
Grand Slam	2000	1200	720	360	180	90	45	10
ATP Finals	+500 (1500 max)	+400 (1000 max)	200 for each match (600 max)					
Masters 1000	1000	600	360	180	90	45	10	
500 Series	500	300	180	90	45	20		

WHAT IS CURRENT METHOD NOT ACCOUNTING FOR?

- Imagine a situation that an unseeded player X is having a day of his life and he beats Roger Federer in a group stage match
- This match against Roger Federer is an important for X. The current system does not account for the weight of the opponent or the importance of a match for a player
- Our method in today’s discussion fills this exact gap. It adds bonus points to each match based on opponent and importance of match

LET US CONSIDER AN EXAMPLE OF NOVEMBER 2009 WORLD TOUR FINALS HELD IN LONDON

- The world tour finals was a final play-off of the best 8 players that participated in that year's ATP tournaments
- Andy Roddick who was at number 6 by highest ATP points backed out so Robin Söderling number 9 got to participate in the tournament
- The first round was a round robin in groups of 4 followed by semi finals and finals. The ATP points and ranks before the tournament were:

Ranking	Player	ATP points
1	Federer, Roger (SUI)	10,150
2	Nadal, Rafael (ESP)	9,205
3	Djokovic, Novak (SRB)	7,910
4	Murray, Andy (GBR)	6,630
5	Del Potro, Juan Martin (ARG)	5,985
6	Roddick, Andy (USA)	4,410
7	Davydenko, Nikolay (RUS)	3,630
8	Verdasco, Fernando (ESP)	3,300
9	Söderling, Robin (SWE)	3,010

AND THE TOURNAMENT RESULTS WERE…

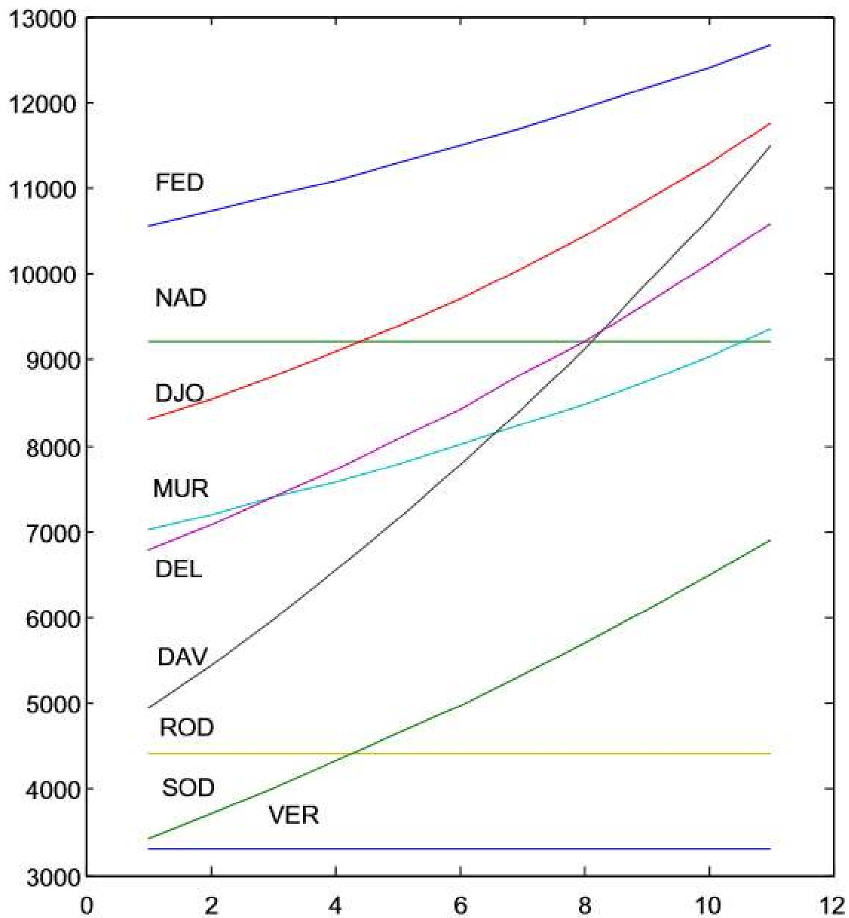
Group A		Group B	
Murray – del Potro:	6-3, 3-6, 6-2	Söderling – Nadal:	6-4, 6-4
Federer – Verdasco:	4-6, 7-5, 6-1	Djokovic – Davydenko:	3-6, 6-4, 7-5
del Potro – Verdasco:	6-4, 3-6, 7-6	Söderling – Djokovic:	7-6, 6-1
Federer – Murray:	3-6, 6-3, 6-1	Davydenko – Nadal:	6-1, 7-6
Murray – Verdasco:	6-4, 6-7, 7-6	Djokovic – Nadal:	7-6, 6-3
del Potro –Federer:	6-2, 6-7, 6-3	Davydenko – Söderling:	7-6, 4-6, 6-3

Semi-final	Davydenko – Federer	6-2, 4-6, 7-5
Semi-final	del Potro – Söderling	6-7, 6-3, 7-6
Final	Davydenko – del Potro	6-3, 6-4

COMPARING HOW RANKS IMPROVED BY ATP METHOD AND BY CONSIDERING WEIGHTS OF OPPONENTS

Bonus pts based on weight of opponent	Current ATP method	0.02	0.03	0.05	0.07	0.08	0.10	0.12	0.13	0.15	0.17
Federer	10,550	10,725	10,906	11,094	11,289	11,492	11,704	11,926	12,158	12,401	12,658
Nadal	9,205	9,205	9,205	9,205	9,205	9,205	9,205	9,205	9,205	9,205	9,205
Djokovic	8,310	8,554	8,816	9,098	9,400	9,725	10,074	10,450	10,854	11,290	11,759
Murray	7,030	7,203	7,386	7,581	7,788	8,008	8,242	8,491	8,758	9,042	9,346
del Potro	6,785	7,081	7,392	7,721	8,068	8,433	8,819	9,227	9,657	10,113	10,595
Roddick	4,410	4,410	4,410	4,410	4,410	4,410	4,410	4,410	4,410	4,410	4,410
Davydenko	4,930	5,442	5,980	6,547	7,144	7,773	8,437	9,137	9,877	10,661	11,490
Verdasco	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300
Söderling	3,410	3,706	4,011	4,325	4,650	4,987	5,338	5,703	6,085	6,484	6,904

TO BETTER UNDERSTAND THIS WE CAN ILLUSTRATE IT FROM A GRAPH BELOW



- As we increase α we see that the finalist Davydenko and Del Potro have ranked up to rank 3 and rank 4 respectively
- This reward of defeating higher ranked players is dependent on the weight we assign
- Thus, deciding the right weight (α) becomes important in this method
- We do not want a rank 7 player (more than 5000 points away from rank 1) to reach rank 1 based on just one tournament
- But we also want to reward these important matches for a player

COMPARISON OF ATP POINTS VS ALTERNATE METHOD WITH EXAMPLE FOR DAVYDENKO(WINNER)

- ATP points

Davydenko bt Nadal		
Ranking	Player	ATP Points
1	Fed	10,550
2	Nad	9,205
3	Djo	8,110
4	Mur	6,830
5	Del	6,185
6	Dav	3,830
7	Sod	3,410
8	Ver	3,300

Davydenko bt Soderling		
Ranking	Player	ATP points
1	Fed	10,550
2	Nad	9,205
3	Djo	8,310
4	Mur	7,030
5	Del	6,385
6	Dav	4,030
7	Sod	3,410
8	Ver	3,300

Davydenko bt Federer		
Ranking	Player	ATP Points
1	Fed	10,550
2	Nad	9,205
3	Djo	8,310
4	Mur	7,030
5	Del	6,785
6	Dav	4,430
7	Sod	3,410
8	Ver	3,300

Davydenko bt Del Potro		
Ranking	Player	ATP Points
1	Fed	10,550
2	Nad	9,205
3	Djo	8,310
4	Mur	7,030
5	Del	6,785
6	Dav	4,950
7	Sod	3,410
8	Ver	3,300

- Alternate Method

Davydenko bt Nadal		
Ranking	Player	Alternate Method
1	Fed	12,815
2	Nad	9,205
3	Djo	8,836
4	Mur	8,027
5	Sod	7,018
6	Del	6,845
7	Dav	5,671
8	Ver	3,300

Davydenko bt Soderling		
Ranking	Player	Alternate Method
1	Fed	12,815
2	Djo	10,877
3	Del	9,608
4	Nad	9,205
5	ur	8,887
6	Dav	7,275
7	Sod	7,018
8	Ver	3,300

Davydenko bt Federer		
Ranking	Player	Alternate Method
1	Fed	12,815
2	Del	11,412
3	Djo	10,877
4	Dav	10,238
5	Nad	9,205
6	Mur	8,887
7	Sod	7,018
8	Ver	3,300

Davydenko bt Del Potro		
Ranking	Player	Alternate Method
1	Dav	13,192
2	Fed	12,815
3	Del	11,412
4	Djo	10,877
5	Nad	9,205
6	Mur	8,887
7	Sod	7,018
8	Ver	3,300

**For the purpose of this experiment we select a very high bonus parameter. In real world application one may not choose such high parameter*

WE HAVE A PYTHON SIMULATION THAT COMPUTES ALL THE POINTS FOR PLAYERS FOR THE TOURNAMENT

- The input of python program is the weight that we decide for opponent player
- The two players and winner of each match i.e. player and winner of all matches in a tournament
- Classification of each match Group stage, Semi final, final
- Weight of each match. For example: 200 pts for group stage match, 400 pts for semi final and 500 pts for final

Let us do a quick demo now

THE RANKING METHOD AND ITS PROPERTIES

- Consider a graph $G = (V,E)$
  - Where –
    - V – set of tennis players
    - E – set of directed edges where edge (i,j) represents a match i won against j and edge (j,i) represents a match j won against i
- To account for more than one matches won by i against j we associate a non negative number a_{ij} with each edge (i,j)
- $a_{ij} =$ number of matches between i and j where i won
- Let $A_{ij} = [a_{ij}]$ be an $n \times n$ matrix
  - Where (i,j) is a_{ij}
  - Matrix A is called match matrix

THE RANKING METHOD AND ITS PROPERTIES

- Let us begin with defining the variables we will use in understanding the ranking method
- **M : set of matches** in a tournament:

$$M = \bigcup_{i \neq j} M(i,j)$$

  (i, j) denoting that the **player** i **defeated player** j

- $\beta_m \geq 0$: the **weight** or importance associated with every match
- $\alpha \geq 0$: **bonus** parameter such that αx_j denotes the percentage of the score of the opponent j that player i defeated.
- A score vector $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$ satisfying the following system of linear equations:

$$x_i = \sum_{j \neq i,\, m \in M(i,j)} (\beta_m + \alpha x_j) \quad (i \leq n). \tag{1}$$

CLAIM: THE SCORE VECTOR IS UNIQUE

$$\begin{aligned} x_i &= \sum_{j \neq i,\ m \in M(i,j)} (\beta_m + \alpha x_j) \\ &= b_i + \alpha \sum_{j \neq i,\ m \in M(i,j)} x_j \end{aligned}$$

where $b_i = \sum_{j \neq i,\ m \in M(i,j)} \beta_m$ = total weight of the matches player i has won.

$$x = \alpha A x + b \tag{2}$$

Let $C_\alpha = I - \alpha A$

$$C_\alpha x = b.$$

Let $r_i(A) = i^{th}$ row sum = total number of matches won by player i.
Similarly, let $s_i(A) = i^{th}$ column sum = total number of matches lost by player i.

$$\gamma_A = \min\{\max_i r_i(A),\ \max_i s_i(A)\} \tag{3}$$

SOME GUIDELINES TO SELECT α FOR RANKING

Consider a player j who has lost s_j matches, and his contribution Δ_j to the other players is-

$$\Delta_j = \sum_i (\alpha a_{ij} x_j) = \alpha x_j \sum_i a_{ij} = \alpha x_j s_j$$

Now $\Delta_j < x_j \implies$ the player increases other players' scores by an amount less than his own score.

Thus, $\alpha x_j s_j < x_j \quad \implies \quad \alpha s_j < 1 \quad \implies \quad \alpha \max_j s_j < 1$

Hence, $\alpha < \frac{1}{(\max_j s_j)} \leq \frac{1}{\gamma_A}$.

Hence, to assure that no player contributes bonus points that are more than p times his score, α can be selected as

$$\alpha < \frac{p}{\gamma_A} \text{ ,where } p \leq 1.$$

THEOREM 1:

Theorem 1: Let $0 \leq \alpha < \frac{1}{\gamma_A}$. Then the following holds:

(i) C_α is an M-matrix, in particular, it is invertible and has a non negative inverse.

(ii) There is a unique solution x of the score equations (2), and this vector satisfies $x \geq b$.

Spectral Radius: Spectral Radius of a square matrix is the largest absolute value of its eigenvalues. It is represented by ρ.
Consider a square matrix H.

$$\rho(H) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } H\}$$

M-matrix: A matrix F is called an M-matrix if $F = sI + H$, where H is a non-negative matrix and $s > \rho(H)$.

THEOREM 1 PROOF:

(i) We know that $C_\alpha = I - \alpha A$ and $A \geq 0$. To show that C_α is an M-matrix, we need to estimate $\rho(A)$, where $\rho(A)$ is the spectral radius of A.

$\therefore \rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$

> Recall:
>
> *Gershgorin's disk theorem:*
>
> Let B be a square matrix. Let D_k denote the set of numbers that are less than or equal to $\sum_{j=1}^{m} |a_{kj}|$ and D'_k denote the set of numbers less than or equal to $\sum_{j=1}^{m} |a_{jk}|$.
>
> Let λ be an eigenvalue of B, then $\lambda \in D_k$ for some k and $\lambda \in D'_k$ for some k.

$\therefore$ $|\lambda|$ is less than or equal to the minimum of maximum of sums of rows or columns of matrix A. i.e.

$$
\begin{aligned}
|\lambda| &\leq \min\{\max_i r_i \, , \max_i s_i\} \\
|\lambda| &\leq \gamma_A
\end{aligned}
$$

THEOREM 1 PROOF:

Now, consider

$$\begin{aligned}
\rho(\alpha A) &= \alpha\rho(A) \\
&= \alpha|\lambda| \\
&\leq \alpha\gamma_A \\
&< 1 && \text{because } 0 \leq \alpha < \frac{1}{\gamma_A}
\end{aligned}$$

Hence, as the coefficient of I is greater than $\rho(\alpha A)$, C_α is an M-matrix.

(ii) As C_α is an M-matrix, it is invertible and has a non-negative inverse. We know that

$$\begin{aligned}
C_\alpha x &= b \\
\implies \quad x &= C_\alpha^{-1} b \\
\because \quad C_\alpha^{-1} b &> 0 && \text{As both } C_\alpha^{-}1 \text{ and } b \text{ are non-negative} \\
\therefore \quad x &> 0
\end{aligned}$$

Consider (2)

$$x = \alpha A x + b$$

As all terms are unique and non-negative, there is a unique solution for square equations and this vector satisfies $x \geq b$. $\square$

COMBINATORIAL INTERPRETATIONS

Theorem 2: Assume that $\alpha\rho(A) < 1$ (which holds if $\alpha < \frac{1}{\gamma_A}$), so C_α is invertible, and let $C_\alpha^{-1} = [d_{ij}(\alpha)]$. Then

$$d_{ij}(\alpha) = \sum_{P \in \mathcal{P}_{ij}} \alpha^{l(P)} \qquad (i, j \leq n) \ ,$$

where -

P is a walk, i.e., a sequence $P\ : e_1\ e_2, ..., \ e_s$ of consecutive edges such that the terminal vertex of edge e_p is the initial vertex of edge e_{p+1} $(p < s)$,

$l(P)$ is the length of walk p which is the number of edges, and

$\mathcal{P}_{ij}$ denotes the set of walks from i to j in G.

THEOREM 2 PROOF:

Proof: Consider the coefficient matrix $C_\alpha = I - \alpha A$ of the score equations. If we assume that $\alpha\rho(A) < 1$, then C_α is invertible and its inverse is the sum of the Neumann series $I + \alpha A + \alpha^2 A^2 + ...$, i.e.

$$C_\alpha^{-1} = (I - \alpha A)^{-1} = I + \alpha A + \alpha^2 A^2 + \alpha^3 A^3 + ...$$

Let the $(i,j)^{th}$ entry of the power matrix A^k be denoted $a_{ij}^{(k)}$. $a_{ij}^{(k)}$ represents the number of distinct walks of length k from vertex i to vertex j in graph G.
Thus,

$$d_{ij}(\alpha) = I + \alpha A + \alpha^2 A^2 + \alpha^3 A^3 + ...$$

As the $(i,j)^{th}$ entry of A^k is $a_{ij}^{(k)}$ and $l(P)$ is the length of walk P, it follows that -

$$d_{ij}(\alpha) = \sum_{P \in \mathcal{P}_{ij}} \alpha^{l(P)}$$

$\square$

COROLLARY

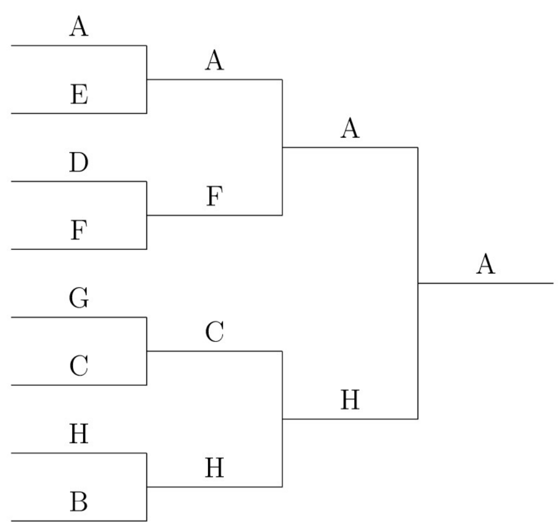
Corollary: Assume that $\alpha\rho(A) < 1$, and let $x = (x_1, x_2, ..., x_n)$ be the unique solution of the score equations (2). Then

$$x_i = \sum_{j=1}^{n} b_j \sum_{P \in \mathcal{P}_{ij}} \alpha^{l(P)} \qquad\qquad (i \leq n).$$

This shows that the score x_i of player i is the sum of the player's match score b_i and bonus score. The bonus score is damped by the factor $\alpha^{l(P)}$ where the exponent is the length of the win sequence that the walk P represents.
The contribution of player j to the score of player i is $b_j d_{ij}(\alpha)$ where $d_{ij} = \sum_{P \in \mathcal{P}} \alpha^{l(P)}$.

ANOTHER EXAMPLE OF SINGLE ELIMINATION TOURNAMENT AND ITS APPLICATION

- A single elimination tournament consists of player participating in order of 2 and each match might result in elimination of a player
- So in a single elimination tournament the match matrix A will have a_{ij} strictly 0 or 1 as each player wins or losses exactly one match against a player
- We have a mock single elimination tournament to simulate and the match results of that tournament are:



SINGLE ELIMINATION TOURNAMENT AND ITS APPLICATION

- The corresponding match matrix of the above tournament will be:

$$
\begin{array}{c c}
 & \begin{array}{cccccccc} A & B & C & D & E & F & G & H \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \end{array} &
\begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

- Now we can rank all the player by solving a simple $Cx = b$ equation

METHODS OF COMPUTING POINTS/RANK OF EACH PLAYER

- The methods to compute any matrix equation can be majorly classified in two types:
  - Direct Methods and Iterative methods
  - The examples of direct methods are:
    - Single Value Decomposition (SVD), LU Factorization, etc.
  - The examples of Iterative methods are:
    - Jacobi, Gauss-Seidel, Successive over Relaxation (SOR) etc.
- Iterative methods are better when the matrix are large or sparse
- Key differences between direct methods and iterative methods are –
  - Entire matrix is stored in memory for direct methods
  - In direct method the exact solution is assured; For iterative methods the solution goes closer to solution depending on number of iterations and accuracy required

QUICK DEMO OF SVD AND GAUSS-SEIDEL

- SVD
  - The matrix transforms matrix A to $A = U\Sigma V^T$
    - where, U and V are unitary matrices, Σ is a non-negative diagonal matrix

So,

$$Ax = b$$

$$\Rightarrow U\Sigma V^T x = b$$

$$\Rightarrow \Sigma V^T x = U^T b$$

Let $C = U^T b$, $w = V^T x$

$$\Rightarrow \Sigma w = C$$

Since, Σ is diagonal, we can easily solve for w using the above equation.

Finally, x can be obtained by solving

$$V^T x = w$$

QUICK DEMO OF SVD AND GAUSS-SEIDEL

- Gauss-Seidel
  - The matrix is split in a lower triangular matrix and a strictly upper triangular matrix

$$A = L + U$$

  - The system of linear equations changes to $Lx = b - Ux$
  - The Gauss-Siedel iterates for value of x from its previous value using the equation:

$$x^{(k+1)} = L^{-1}(b - Ux^{(k)})$$

CLOSING WORDS

- The example discussed with tennis can be used by tennis professionals to track and compare their performances with all other players in the tournament after each match
- A tennis enthusiast may obtain his "personal ranking" by selecting a value of α
- The new ranking can be used as a way of showing which players are on they way up, or down, on the rankings
- This pair wise comparison and ranking can be used in many other application where weightage of the other objects/players/projects etc. has high importance

THANK YOU