Sports Ranking

In Practice and In Principle

Ian Hamilton

University of Warwick

8th June 2021

Definitions

Ranking - the ordered position of each team i.e. a positive integer from one to the number of teams in the tournament.

Rating - a parameter (or a set of parameters) representing the quality (or qualities) of a team, that in some way allows for ranking.

Points - a value awarded to a team due to a contest result.

Score - the in-game accumulations on which a result is based.

Ranking examples

Premier League standings

MATCHES | NEWS | TABLE | STATS

Season
2020-21

Club		MP	W	D	L	GF	GA	GD	Pts
1	Man City	38	27	5	6	83	32	51	86
2	Man United	38	21	11	6	73	44	29	74
3	Liverpool	38	20	9	9	68	42	26	69
4	Chelsea	38	19	10	9	58	36	22	67
5	Leicester City	38	20	6	12	68	50	18	66

Premiership Rugby standings

MATCHES | NEWS | TABLE

Team		P	W	D	L	PD	B	Pts
1	Bristol	21	16	1	4	182	15	81
2	Exeter	21	16	0	5	267	14	78
3	Sale	21	16	0	5	137	9	73
4	Harlequins	21	12	1	8	111	14	66
5	Northampton	21	11	0	10	18	9	55

NBA standings

GAMES | NEWS | TABLE

Season
2020-21

Eastern Conference

Team		W	L	Pct
1	76ers	49	23	.681
2	Nets	48	24	.667
3	Bucks	46	26	.639
4	Knicks	41	31	.569
5	Hawks	41	31	.569

Points per match

Points per match $= \frac{\text{Total Points}}{\text{Matches Played}}$

Example: Ligue 1 2019/20

Problem solved: Different number of matches

Adjusted Points per match

Home/Away adjusted points per match

$$= \frac{1}{2} \times \left(\frac{\text{Total Home Points}}{\text{Home Matches Played}} + \frac{\text{Total Away Points}}{\text{Away Matches Played}}\right)$$

Example: English rugby (exc. Premiership) 2019/20

Problem solved: Different number of matches; Proportion home/away

Schedule Strength

But what do we do when the schedules are not balanced?

There are three main factors to consider:

1. Number of matches
2. Proportion home vs away
3. Strength of opposition

So what might we do to address strength of opposition?

Rating Percentage Index

$$\begin{aligned}\text{RPI} = 25\% \times \text{Win Percentage} \\ + 50\% \times \text{Opposition's Win Percentage} \\ + 25\% \times \text{Opposition's Opposition's Win Percentage.}\end{aligned}$$

Example: NCAA basketball pre-2018

Problem solved: Different number of matches; Strength of schedule (??)

Working Example

	A	B	C	D	E
A	0	1	1	1	0
B	0	0	1	1	1
C	0	0	0	1	1
D	0	0	0	0	1
E	1	0	0	0	0

Table: Wins

	A	B	C	D	E
A	0	89-64	91-90	84-81	76-78
B	64-89	0	91-86	78-72	81-78
C	90-91	86-91	0	78-68	79-55
D	81-84	72-78	68-78	0	65-48
E	78-76	78-81	55-79	48-65	0

Table: Scores

Massey ratings

Idea: $r_i - r_j = y_k$

In matrix notation: $X\boldsymbol{r} = \boldsymbol{y}$, where:

- $X_{m \times n}$ is a matrix where each row is with respect to a match with a 1 in the column of the winner and a -1 in column of the loser and 0 elsewhere
- $\boldsymbol{r}_{n \times 1}$ is a rating vector
- $\boldsymbol{y}_{m \times 1}$ is a net score vector

Massey ratings

$$X = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}, \; \boldsymbol{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{pmatrix}, \text{ and } \boldsymbol{y} = \begin{pmatrix} 25 \\ 1 \\ 3 \\ 5 \\ 6 \\ 3 \\ 10 \\ 26 \\ 17 \\ 2 \end{pmatrix}$$

Massey ratings

An ordinary least squares estimate may be obtained from the *normal equations*:

$$X^T X \boldsymbol{r} = X^T \boldsymbol{y}$$

In the present example

$$Z = X^T X = \begin{pmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{pmatrix}, \boldsymbol{s} = X^T \boldsymbol{y} = \begin{pmatrix} 27 \\ -11 \\ 30 \\ -2 \\ -44 \end{pmatrix}$$

i.e. $Z_{n \times n}$ is a symmetric matrix where z_{ii} is the number of matches played by team i and z_{ij} is the negative of the number of matches i has played against team j, and $\boldsymbol{s}_{n \times 1}$ is a vector of the aggregate score differential for each team.

Massey ratings

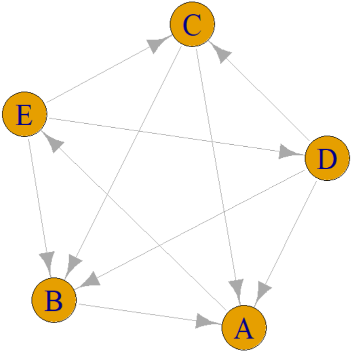
However the columns of Z are linearly dependent i.e. $\text{rank}(Z) < n$. Understood another way we need an identifiability constraint e.g. the sum of the ratings is zero

$$Z = \begin{pmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \boldsymbol{s} = X^T \boldsymbol{y} = \begin{pmatrix} 44 \\ -11 \\ 17 \\ -8 \\ 0 \end{pmatrix}$$

Gives ranking C, A, D, B, E

Glory seeker

Idea: Consider the tournament as a network. You randomly choose a team to start with then at each step you randomly choose from among the teams that has beaten them. Rank teams by the proportion of time you have spent supporting each team.



Glory seeker

$$C = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Define a column-normalised matrix $\tilde{C}$ where $\tilde{c_{ij}} = c_{ij} / \sum_{i=1}^{n} c_{ij}$.

$$\tilde{C} = \begin{pmatrix} 0 & 1 & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Then our rating vector is the stationary distribution of $\tilde{C}$ i.e. the leading eigenvector $\boldsymbol{r} = \tilde{C}\boldsymbol{r}$

Glory seeker

Consider PageRank - the stationary distribution for the matrix P

$$P = \alpha \tilde{C} + \frac{1-\alpha}{n} e e^T$$

where $e_{n\times 1}$ is a vector of 1s i.e. leading eigenvector of $r = Pr$

So the Glory seeker ranking is just an undamped PageRank.

Gives ranking A=E, B, C, D

Bradley Terry

Idea: The probability that team i beats team j is given by

$$p_{ij} = P(i \succ j) = \frac{r_i}{r_i + r_j}$$

where r_i is positive-valued, and can be thought of as a parameter reflecting the strength of team i.

Estimated using maximum likelihood

$$L = \prod_{i<j} \binom{m_{ij}}{c_{ij}} p_{ij}^{c_{ij}} (1 - p_{ij})^{m_{ij} - c_{ij}},$$

Gives ranking A=B, C, D=E

NET

In 2018 NCAA introduced a new metric to replace the RPI for college basketball:

"NET relies on game results, strength of schedule, game location, scoring margin, net offensive and defensive efficiency, and the quality of wins and losses. To make sense of team performance data, late-season games (including from the NCAA tournament) were used as test sets to develop a ranking model leveraging machine learning techniques. The model, which used team performance data to predict the outcome of games in test sets, was optimized until it was as accurate as possible. The resulting model is the one that will be used as the NET going forward." (NCAA, 2018).

NET

For the 2020/21 season, a new version was introduced with a reduced set of indicators:

> *"...the NCAA Evaluation Tool will be changed to increase accuracy and simplify it by reducing a five-component metric to just two. The remaining factors include the Team Value Index (TVI), which is a result-based feature that rewards teams for beating quality opponents, particularly away from home, as well as an adjusted net efficiency rating. The adjusted efficiency is a team's net efficiency, adjusted for strength of opponent and location (home/away/neutral) across all games played... No longer will the NET use winning percentage, adjusted winning percentage and scoring margin. The change was made after the committee consulted with Google Cloud Professional Services, which worked with the NCAA to develop the original NET." (NCAA, 2020)*

Comparison

	M	GS	BT	NET
A	2	1=	1=	?
B	4	3	1=	?
C	1	4	3	?
D	3	5	4=	?
E	5	1=	4=	?

Alternatives

There are a cornucopia of other alternatives including:

- Fair Bets
- Minimum Violation
- Trophic levels
- Colley matrix
- Keener method
- Elo
- Wei-Kendall
- etc.

So how should we choose?

What should we care about?

Q: Which of these do you think should be the most important factor in choosing a sports ranking methodology?

1. Transparency
2. Predictive ability
3. Principles

Please vote 1, 2 or 3 via the chat.

Bradley Terry

In the context of tournaments, the probability that team i beats team j is given by

$$P(i \succ j) = \frac{r_i}{r_i + r_j}$$

where r_i is positive-valued, and can be thought of as a parameter reflecting the strength of team i.

Zermelo (1929), Bradley & Terry (1952)

A principle-based approach

Maximise entropy

$$S(p) = -\sum_{i,j} p_{ij} \log p_{ij} \quad ,$$

subject to the retrodictive criterion,

$$\sum_j p_{ij} m_{ij} = \sum_j c_{ij} \quad , \tag{1}$$

where p_{ij} is the probability that i beats a , and $C + C^T = M = [m_{ij}]$ is the symmetric matrix where m_{ij} is the number of matches between i and j.

A principle-based approach

Then taking the Lagrangian as

$$\mathcal{L}(p, \boldsymbol{\lambda}) = S(p) - \sum_{i=1}^{n} \lambda_i \Bigg(\sum_{j=1, j\neq i}^{n} (m_{ij} p_{ij} - c_{ij}) \Bigg),$$

and setting $\frac{\partial \mathcal{L}}{\partial p_{ij}} = 0$ for all p_{ij} in the normal way gives that

$$\frac{\partial S(p)}{\partial p_{ij}} = \frac{\partial}{\partial p_{ij}} \sum_{r=1}^{n} \lambda_r \Bigg(\sum_{r=1, r\neq s}^{n} (m_{rs} p_{rs} - c_{rs}) \Bigg) \quad \text{for all } i, j.$$

So that for all i, j such that $m_{ij} \neq 0$,

$$-\log p_{ij} + \log(1 - p_{ij}) = \lambda_i - \lambda_j,$$

or equivalently

$$p_{ij} = \frac{r_i}{r_i + r_j},$$

where $r_i = \exp(-\lambda_i)$.

Extension to include ties

$$P(i \succ j) = \frac{r_i}{r_i + r_j + \nu\sqrt{r_i r_j}}$$

$$P(i \approx j) = \frac{\nu\sqrt{r_i r_j}}{r_i + r_j + \nu\sqrt{r_i r_j}}$$

Davidson (1970)

Extension to account for home advantage (order effects)

$$P(i \succ j) = \frac{r_i}{r_i + \gamma r_j + \nu\sqrt{r_i r_j}}$$
$$P(i \prec j) = \frac{\gamma r_j}{r_i + \gamma r_j + \nu\sqrt{r_i r_j}}$$
$$P(i \approx j) = \frac{\nu\sqrt{r_i r_j}}{r_i + \gamma r_j + \nu\sqrt{r_i r_j}}$$

Davidson & Beaver (1977)

Applying to 3 for a win, 1 for a draw

$$P(i \succ j) = \frac{r_i}{r_i + r_j + \nu(r_i r_j)^{\frac{1}{3}}}$$

$$P(i \approx j) = \frac{\nu(r_i r_j)^{\frac{1}{3}}}{r_i + r_j + \nu(r_i r_j)^{\frac{1}{3}}}$$

See: `https://alt-3.uk/`

Firth (2017)

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?

A: Rugby union!

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?

A: Schools rugby!

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?

A: Daily Mail Trophy!

Motivation

Q: Wouldn't it be nice (for me, at least) if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some improvement?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some improvement?

A: Full house!

Rugby union scoring rule

League Points:

4 points for a win

2 points for a draw

0 points for a loss

1 bonus point for losing by less than seven points

1 bonus point for scoring four or more tries

Summary

Model	B-T	Davidson	Firth	Rugby
Points - win	1	2	3	4
Points - draw	NA	1	1	2
Points - other	NA	NA	NA	1 (try,losing)
Model - i win	r_i	r_i	r_i	???
Model - draw	NA	$(r_i r_j)^{1/2}$	$(r_i r_j)^{1/3}$	???
Model - other	NA	NA	NA	???

RASR (pronounced 'razor') - Ranking Algorithm for Schools Rugby

Part one: result outcome

$$
\begin{aligned}
P(\text{team } i \text{ beats team } j \text{ by wide margin}) &\propto \tau^4 r_i^4 \\
P(\text{team } i \text{ beats team } j \text{ by narrow margin}) &\propto \kappa\tau^3 r_i^4 r_j \\
P(\text{team } i \text{ draws with team } j) &\propto \nu r_i^2 r_j^2 \\
P(\text{team } j \text{ beats team } i \text{ by narrow margin}) &\propto \frac{\kappa r_i r_j^4}{\tau^3} \\
P(\text{team } j \text{ beats team } i \text{ by wide margin}) &\propto \frac{r_j^4}{\tau^4}
\end{aligned}
$$

RASR (pronounced 'razor') - Ranking Algorithm for Schools Rugby

Part two: try bonus outcome

$$\begin{aligned} P(\text{team } i \text{ and team } j \text{ both gain try bonus point}) &\propto \theta r_i r_j \\ P(\text{only team } i \text{ gains try bonus point}) &\propto \tau r_i \\ P(\text{only team } j \text{ gains try bonus point}) &\propto \frac{r_j}{\tau} \\ P(\text{neither team gains try bonus point}) &\propto \phi \end{aligned}$$

A principle-based approach

Maximise entropy

$$S(p) = -\sum_{i,j} \sum_{a,b} p^{ij}_{a,b} \log p^{ij}_{a,b} \quad ,$$

subject to conditions,

$$\sum_{a,b} p^{ij}_{a,b} = 1 \quad , \tag{2}$$

and

$$\sum_{j} \sum_{a,b} a p^{ij}_{a,b} = \sum_{j} \sum_{a,b} a m^{ij}_{a,b} \quad , \tag{3}$$

where $p^{ij}_{a,b}$ is the probability that i gains a points and j gains b points, and $m^{ij}_{a,b}$ is the number of matches that have resulted with i gaining a points and j gaining b points.

Potential models

Examples:

- Try bonus dependent on result outcome and opposition
- Try bonus independent of result outcome but dependent on opposition
- Try bonus independent of result outcome and opposition
- Offensive-defensive strengths
- Home-away strengths

To prior or not to prior?

Introduce a dummy $team_0$ against whom each other team wins one and loses one, then decide how much weight to give these matches.

Pros:

- Ensures connectedness therefore rating from start of season
- Explicitly controls fairness in situations of varying fixture numbers
- Allows for estimation of structural parameters even with existence of 100% record

Cons:

- Might not match intuition / round robin outcomes

Intuitive Measure

Projected Points per Match

$$\text{PPPM}_i = \frac{1}{n-1} \sum_j \sum_{a,b} a p^{ij}_{a,b}$$

Intuitive measure that converges to the rating in round robin

Daily Mail Trophy

League Points per Match + Additional Points

Additional Points in the Daily Mail Trophy are awarded based on the ranking of the current season's opponents in the previous season's tournament:

Rank 1 to 25:	0.3
Rank 26 to 50:	0.2
Rank 51 to 75:	0.1
Otherwise:	0

Results 2015/16

School	DMT Rank	DMT	PPPM Rank	PPPM
Wellington College	1	6.46	7	3.73
Kirkham	2	6.44	1	4.41
Bedford	3	6.35	2	4.37
Bromsgrove	4	6.21	4	4.15
Sedbergh	5	6.10	5	3.99
Woodhouse Grove	6	5.65	19	3.31
Millfield	7	5.21	13	3.64
Clifton College	8	5.11	8	3.73
Solihull	9	5.10	11	3.67
St Paul's	9	5.10	14	3.58

Results 2016/17

School	DMT Rank	DMT	PPPM Rank	PPPM
Wellington College	1	7.22	3	4.37
Sedbergh	2	6.50	2	4.43
Harrow	3	6.34	6	4.22
St Peter's, York	4	6.23	8	4.06
Kirkham	5	6.15	1	4.61
Canford	6	6.10	9	4.02
Clifton College	7	6.00	5	4.25
Rugby	8	5.96	7	4.06
Brighton College	9	5.90	4	4.29
Woodhouse Grove	10	5.81	12	3.93

Results 2017/18

School	DMT Rank	DMT	PPPM Rank	PPPM
Sedbergh	1	7.41	1	4.65
Wellington College	2	7.18	7	4.18
Cranleigh	3	6.33	4	4.32
Harrow	4	6.20	3	4.33
Cheltenham College	5	6.16	8	4.07
St Peter's, York	6	5.83	6	4.19
Brighton College	7	5.63	20	3.59
Reed's	8	5.50	2	4.38
Clifton College	8	5.50	16	3.72
Haileybury	10	5.49	10	4.02

Transparency revisited

After all this analysis we recommended a ranking that we called Dapper (Damped and Adjusted Points Per match)

$$\text{Merit Points} = \frac{\text{League Points} + \text{Additional Points} + 9}{\text{Matches Played} + 3}$$

with Additional Points taken to be

Rank 1 to 25:	2.25
Rank 26 to 50:	1.5
Rank 51 to 75:	0.75
Otherwise:	0

Transparency revisited

Team	RASR	DMT	Dapper
Kirkham Grammar School	1	2	1
Bedford School	2	3	3=
Bromsgrove School	3	4	2
Sedbergh School	4	5	3=
Seaford College	5	12	6
Wellington College	6	1	7
Clifton College	7	8	5
QEGS, Wakefield	8	17	13
Tonbridge School	9	18	9
Solihull School	10	13	10

Table: 2015/16: Top ten comparison

Resources

Talks: RSS Merseyside Local Group: Statistics and Football
https://www.youtube.com/channel/UChNoOmvmV9KzB8KCxP2n9_w

Books: Who's #1? by Langville & Meyer; Contest Theory (ch 9,10) by Vojnovic

Conferences: http://www.nessis.org/index.html

Competitions: https://rss.org.uk/news-publication/news-publications/2021/section-group-reports/sports-section-euro-2020-prediction-competition/

Websites: https://alt-3.uk/; www.warwick.ac.uk/IanHamilton

WDSS Summer project: https://recruitment.wdss.io

Bibliography

NCAA (2018). *The NET explained*. https://www.ncaa.com/news/basketball-men/article/2018-11-26/net-explained-ncaa-adopts-new-college-basketball-ranking, Accessed November 4, 2019.

NCAA (2020). *College basketball's NET rankings, explained*. https://www.ncaa.com/news/basketball-men/article/2020-07-12/college-basketballs-net-rankings-explained, Accessed March 31, 2021.