The Mathematics of Ranking Schemes or Should Utah Be Ranked in the Top 25?

J. P. Keener

Mathematics Department

University of Utah

Introduction

Problem: How to rank sports teams?

The Challenge:

Introduction

Problem: How to rank sports teams?

The Challenge:

- Uneven paired competition (not all teams play each other)

Introduction

Problem: How to rank sports teams?

The Challenge:

- Uneven paired competition (not all teams play each other)
- The data are sparse. (Each team plays 10 games and there are over 100 teams)

Introduction

Problem: How to rank sports teams?

The Challenge:

- Uneven paired competition (not all teams play each other)
- The data are sparse. (Each team plays 10 games and there are over 100 teams)
- There is no well-ordering (Team A beats Team B who beats Team C who beats team A).

Simple Idea

- Suppose team i has W_i wins and L_i losses, let

$$r_i = \frac{W_i}{W_i + L_i} = \frac{W_i}{N_i}.$$

Call $r = (r_i)$ the ranking vector.

Simple Idea

- Suppose team i has W_i wins and L_i losses, let

$$r_i = \frac{W_i}{W_i + L_i} = \frac{W_i}{N_i}.$$

Call $r = (r_i)$ the ranking vector.

- Alternate representation: Set $A = (a_{ij})$ where $a_{ij} = \frac{1}{N_i}$ if team i beat team j, $a_{ij} = 0$ otherwise. Then

$$r = Ar^0, \qquad r^0 = 1.$$

Basic Idea

- First modification: To get an indication of strength of schedule, let

$$r = A(Ar^0) = A^2 r^0.$$

Imagine the Possibilities
Mathematical Biology
University of Utah

Basic Idea

- First modification: To get an indication of strength of schedule, let

$$r = A(Ar^0) = A^2 r^0.$$

- Obvious generalization:

$$r = \frac{A^n r^0}{|A^n r^0|}$$

Basic Idea

- First modification: To get an indication of strength of schedule, let

$$r = A(Ar^0) = A^2 r^0.$$

- Obvious generalization:

$$r = \frac{A^n r^0}{|A^n r^0|}$$

- Question: What if we let $n \to \infty$?
- Answer: Define r to be eigenvector of A

$$Ar = \lambda r$$

Method 1:The Eigenvector Method

Suppose that team i has rank r_i, and a_{ij} is a measure of the result of the game between team i and team j where

- $0 \leq a_{ij} \leq 1$, $a_{ij} = 0$ if teams i and j have not played,
- $a_{ij} + a_{ji} = 1$ if teams i and j have played each other.

Assign a score s_i to team i

$$s_i = \frac{1}{N_i} \sum_j a_{ij} r_j$$

and then define r so that

$$\lambda r_i = s_i \qquad \text{or} \qquad Ar = \lambda r$$

Method 1:Existence Question

Do solutions of this eigenvalue problem exist?

The Perron Frobenius Theorem:

Imagine the Possibilities

Mathematical Biology
University of Utah

Method 1:Existence Question

Do solutions of this eigenvalue problem exist?

The Perron Frobenius Theorem:

- If A has nonnegative entries, then there exists nonnegative solution of $Ar = \lambda r$.

Method 1:Existence Question

Do solutions of this eigenvalue problem exist?

The Perron Frobenius Theorem:

- If A has nonnegative entries, then there exists nonnegative solution of $Ar = \lambda r$.
- If A is irreducible, then r is strictly positive and unique, and λ is simple and maximal.

Method 1:Existence Question

Do solutions of this eigenvalue problem exist?

The Perron Frobenius Theorem:

- If A has nonnegative entries, then there exists nonnegative solution of $Ar = \lambda r$.
- If A is irreducible, then r is strictly positive and unique, and λ is simple and maximal.
- r can be calculated using the power method:

$$r = \lim_{n\to\infty} \frac{A^n r^0}{|A^n r^0|}, \qquad r^0 = 1$$

Choosing a_{ij}

Ways to specify a_{ij}:

- $a_{ij} = 1$ if team i beat team j. (W-L record)

Choosing a_{ij}

Ways to specify a_{ij}:

- $a_{ij} = 1$ if team i beat team j. (W-L record)
- $a_{ij} = \frac{S_{ij}}{S_{ij}+S_{ji}}$, with $S_{ij} =$ points earned by team i vs team j.

Choosing a_{ij}

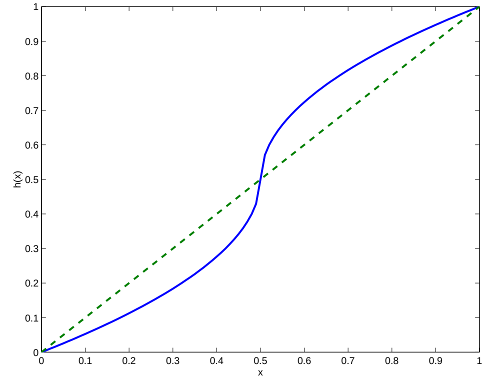
Ways to specify a_{ij}:

- $a_{ij} = 1$ if team i beat team j. (W-L record)
- $a_{ij} = \frac{S_{ij}}{S_{ij}+S_{ji}}$, with $S_{ij} =$ points earned by team i vs team j.
- $a_{ij} = \frac{S_{ij}+1}{S_{ij}+S_{ji}+2}$.

Choosing a_{ij}

Ways to specify a_{ij}:

- $a_{ij} = 1$ if team i beat team j. (W-L record)
- $a_{ij} = \frac{S_{ij}}{S_{ij}+S_{ji}}$, with S_{ij} = points earned by team i vs team j.
- $a_{ij} = \frac{S_{ij}+1}{S_{ij}+S_{ji}+2}$.
- $a_{ij} = h\left(\frac{S_{ij}+1}{S_{ij}+S_{ji}+2}\right)$.



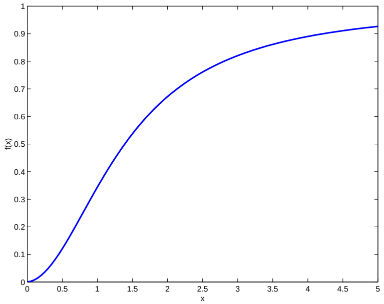
Method 2:The Nonlinear Fixed Point Method

Suppose $f(x)$ is positive, increasing, concave for $0 \leq x < \infty$.
Define r as solution of

$$r_i = F_i(r) = \frac{1}{N_i} \sum_j f(a_{ij} r_j)$$

where

$$f(x) = \frac{.05x + x^2}{2 + .05x + x^2}$$



Method 2:Finding r

Suppose $F(x) : R^n \rightarrow R^n$ is continuous,

Method 2:Finding r

Suppose $F(x) : R^n \rightarrow R^n$ is continuous,

- positive: $F(p) > 0$ for all $p > 0$,

Method 2:Finding r

Suppose $F(x) : R^n \rightarrow R^n$ is continuous,

- positive: $F(p) > 0$ for all $p > 0$,
- increasing: $F(p) > F(q)$ whenever $p > q$,

Method 2:Finding r

Suppose $F(x) : R^n \rightarrow R^n$ is continuous,

- positive: $F(p) > 0$ for all $p > 0$,
- increasing: $F(p) > F(q)$ whenever $p > q$,
- concave: $F(tp) > tF(p)$ for $0 < t < 1$, $p > 0$

Method 2:Finding r

Suppose $F(x) : R^n \rightarrow R^n$ is continuous,

- positive: $F(p) > 0$ for all $p > 0$,
- increasing: $F(p) > F(q)$ whenever $p > q$,
- concave: $F(tp) > tF(p)$ for $0 < t < 1$, $p > 0$

then the sequence of vectors $r^k = F(r^{k-1})$, with $r^0 = 1$ is a monotone decreasing sequence with $\lim_{n\rightarrow\infty} r^n = r$ where

$$r = F(r)$$

Remark: This is a generalization of the Perron Frobenius Theorem.

Method 3:The Probability Method

Suppose the probability that team i beats team j is

$$\pi_{ij} = \frac{r_i}{r_i + r_j} \approx \frac{S_{ij}}{S_{ij} + S_{ji}}.$$

Then

$$S_{ij} r_j \approx S_{ji} r_i.$$

Thus, let r minimize

$$\sum_{i,j} (S_{ji} r_i - S_{ij} r_j)^2 \qquad \text{subject to} \qquad \sum r_i^2 = 1$$

Method 3:Finding r

Use Lagrange multipliers and minimize

$$\sum_{i,j}(S_{ji}r_i - S_{ij}r_j)^2 - \mu(\sum_i r_i^2 - 1)$$

equivalently, find r where

$$Br = \mu r$$

where

$$B = (b_{ij}), \qquad b_{ii} = \sum_k S_{ik}^2, \qquad b_{ij} = -S_{ij}S_{ji} \text{ for i} \neq \text{j}$$

Method 3:Finding r

Pick λ_0 so that $B+\lambda_0 I$ is diagonally dominant. Then, $(B+\lambda_0 I)^{-1}$ is a positive matrix.

From the Perron Frobenius theorem,

$$\lim_{n\to\infty} \frac{(B+\lambda_0 I)^{-n} r^0}{|(B+\lambda_0 I)^{-n} r^0|} = r$$

"Linear Models"

Remark: This model is one of a class of models called linear models for which

$$\pi_{ij} = h(r_i - r_j)$$

For example, h could be a Gaussian (bell-shaped) or an exponential function.

The choice $\pi_{ij} = \frac{r_i}{r_i + r_j}$ is exponential, since with $r = e^v$

$$\pi_{ij} = \frac{e^{v_i}}{e^{v_i} + e^{v_j}} = \frac{e^{v_i - v_j}}{1 + e^{v_i - v_j}}.$$

This is the model (purportedly) used for national tennis, squash, etc. rankings.

Method 4:Maximum Likelihood

Suppose the outcome of a game is a Bernoulli trail with π_{ij} the probability that team i beats team j. If a_{ij} represents the outcome of games, then the probability of observing that outcome is

$$P = \Pi_{i<j} \begin{pmatrix} a_{ij} + a_{ji} \\ a_{ij} \end{pmatrix} \pi_{ij}^{a_{ij}} \pi_{ji}^{a_{ji}}$$

With $\pi_{ij} = \frac{r_i}{r_i + r_j}$, this is

$$P(r) = \prod_{i<j} \begin{pmatrix} a_{ij} + a_{ji} \\ a_{ij} \end{pmatrix} \left(\frac{r_i}{r_i + r_j} \right)^{a_{ij}} \left(\frac{r_j}{r_i + r_j} \right)^{a_{ji}}$$

Method 4:Maximum Likelihood

Equivalently, maximize

$$\begin{aligned}\ln F_A(r) &= \ln \Pi_{i<j} \left(\frac{r_i}{r_i + r_j}\right)^{a_{ij}} \left(\frac{r_j}{r_i + r_j}\right)^{a_{ji}} \\ &= \sum_{i<j} \left(a_{ij} \left(\ln r_i - \ln(r_i + r_j)\right) + a_{ji} \left(\ln r_j - \ln(r_i + r_j)\right)\right)\end{aligned}$$

Consequently, require

$$\nabla_r(\ln F_A(r)) = \frac{\alpha_k}{r_k} - \sum_j \frac{A_{jk}}{r_j + r_k} = 0$$

where $\alpha_k = \sum_j a_{jk}$, $A_{jk} = a_{jk} + a_{kj}$.

Method 4:How to find r

Integrate the differential equation

$$\frac{dr_k}{dt} = \frac{\alpha_k}{r_k} - \sum_j \frac{A_{jk}}{r_j + r_k}$$

This always converges because

- It is a gradient system
- $\ln F_A(r)$ is nondecreasing along trajectories
- The Hessian is negative definite (proof uses the Perron Frobenius Theorem)

Method 4:Remarks

- This method is usually referred to as the Bradley-Terry model.

Method 4:Remarks

- This method is usually referred to as the Bradley-Terry model.
- There is no reason to keep $a_{ij} = 0$ or 1. A better choice is

$$a_{ij} = \frac{S_{ij}}{S_{ij} + S_{ji}}$$

Results

Week 9: 2003 Football Season

Team	W-L	M1	M2	M3	M4	AP	ESPN
Oklahoma	7-0	2	1	1	2	1	1
USC	6-1	1	2	9	6	5	4
Georgia	6-1	9	3	3	1	4	5
Miami	7-1	16	4	4	5	2	2
Virginia Tech	6-0	33	5	2	12	3	3
Florida State	6-1	27	6	12	4	6	7
Washington State	6-1	8	7	19	14	6	6
Nebraska	6-1	10	8	6	7	14	11
Kansas State	5-2	6	9	8	10	-	-
Purdue	5-1	31	10	10	25	10	10
Minnesota	6-2	46	11	15	16	-	24
Ohio State	6-1	11	12	24	15	8	8

Results

Team	W-L	M1	M2	M3	M4	AP	ESPN
Michigan	5-3	15	13	26	9	13	15
LSU	6-1	22	14	5	3	9	9
Michigan State	6-1	34	15	13	21	11	12
Oklahoma State	6-1	24	16	17	18	18	19
Miami(Oh)	6-1	19	17	61	27	-	-
Utah	6-1	3	18	20	23	24	23
TCU	7-0	42	19	29	19	15	13
Iowa	5-2	13	20	31	8	16	16
Texas Tech	5-2	28	21	25	31	-	-
Florida	5-3	21	22	16	11	25	25
Boise State	6-1	39	23	34	44	-	-
Alabama	3-5	43	24	7	13	-	-
Texas	5-2	25	25	23	17	19	18