This is our old blog. It hasn't been active since 2011. Please see the link above for our current blog or click the logo above to see all of the great data and content on this site.

A very simple ranking system

Posted by Doug on May 8, 2006

My friend Joe Bryant says that the BCS bowl matchups are like getting a shrimp cocktail at Morton's Steakhouse. Sure, it's better than what you normally eat, but at the same time it's frustrating and disappointing because you can see a bunch of far preferable alternatives right there in front of your eyes. I tend to agree. Nonetheless, it is not in any way an exaggeration to say that the BCS revived my interest in college football. Not because of the matchups the system has produced, but because it gave me an excuse to learn some very interesting mathematics.

As you probably know, the participants in the BCS championship game are determined in part by a collection of computer rankings. Those computer rankings are implementing algorithms that "work" because of various mathematical theorems. At some point, I'm going to use this blog to write down everything I know about the topic (which by the way is a drop in the bucket compared to what many other people know; I am not an expert, just a fan) in language that a sufficiently interested and patient non-mathematician can understand.

I'll start that off today by describing one of the most basic ranking algorithms.

The idea is to define a system of 32 equations in 32 unknowns. The solution to that system will be collection of 32 numbers and those numbers will serve as the ratings of the 32 NFL teams. Define R_ind as Indianapolis' rating, R_pit as Pittsburgh's rankings, and so on. Those are the unknowns. The equations are:

```
R_ind = 12.0 + (1/16) (R_bal + R_jax + R_cle + . . . . . + R_ari)

R_pit =  8.2 + (1/16) (R_ten + R_hou + R_nwe + . . . . + R_det)

.

.

.

R_stl = -4.1 + (1/16) (R_sfo + R_ari + R_ten + . . . . + R_dal)
```

One equation for each team. The number just after the equal sign is that team's average point margin. In plain English, the first equation says:

> The Colts' rating should equal their average point margin (which was +12), plus the average of their opponents' ratings

So every team's rating is their average point margin, adjusted up or down depending on the strength of their opponents. Thus an average team would have a rating of zero. Suppose a team plays a schedule that is, overall, exactly average. Then the sum of the terms in parentheses would be zero and the team's rating would be its average point margin. If a team played a tougher-than-average schedule, the sum of the terms in parentheses would be positive and so a team's rating would be bigger than its average point margin.

It would be easy to find the Colts' rating if we knew all their opponents' ratings. But we can't figure those out until we've figured out *their* opponents' ratings, and we can't figure those out until. . ., you get the idea. Everyone's rating essentially depends on everyone else's rating.

So how do you actually find the set of values that solves this system of equations? In high school you probably learned how to solve 2-by-2 and maybe 3-by-3 systems of equations by putting some numbers into a matrix, doing some complicated operations on that matrix, and then reading the solutions off the new matrix. Same thing here, except you've got a 32-by-32 matrix instead of a 2-by-2 matrix. If you wanted college football rankings, it'd be 120-by-120. I recommend using a computer.

It's more instructive, though, to solve it a different way. We'll start by giving everyone an initial rating, which is just their average point margin. I'll use the Colts as an example. Their initial rating is +12.0. Now look at the average of their opponents' intial ratings:

```
Opp   Rating

===========

ari    -4.75

bal    -2.12

cin     4.44

cle    -4.31

hou   -10.69

hou   -10.69

jax     5.75

jax     5.75

nwe     2.56

pit     8.19

ram    -4.12

sdg     6.62

sea    11.31

sfo   -11.81

ten    -7.62

ten    -7.62
```

Those average -1.2, so the Colts' new rating will be 12.0 - 1.2, which is 10.8. So after this calculation the Colts' rating changed from +12 to +10.8. But meanwhile, every other team's rating changed as well, so we have to do the whole thing over again with the new ratings. On the second pass, the Colts schedule looks a bit different:

```
Opp   Rating

===========

ari    -4.76

bal    -1.49

cin     4.09

cle    -3.85

hou    -9.69

hou    -9.69

jax     4.85

jax     4.85

nwe     3.09

pit     8.02

ram    -5.16

sdg     8.62

sea     8.99

sfo   -10.77

ten    -7.30

ten    -7.30
```

The average of these is -1.1, so the Colts' opponents aren't quite as bad as they looked at first. Indy's new rating is 12.0 - 1.1, which is 10.9. Uh oh! Everyone else's ratings just changed again, so we've got to run through the same procedure again. And again. And again. And eventually the numbers stop changing. When that happens, you know you've arrived at the solution. Take a look at the Colts schedule with the final rankings and you'll be able to convince yourself that this method works:

			OPP	Adj
WK	OPP	Margin	Rating	Margin
===	===	=======	======	=====
1	bal	17	-1.83	15.17
2	jax	7	4.76	11.76
3	cle	7	-4.22	2.78
4	ten	21	-7.57	13.43
5	sfo	25	-11.15	13.85

```
 6   ram     17    -5.15    11.85

 7   hou     18   -10.03     7.97

 9   nwe     19     3.14    22.14

10   hou     14   -10.03     3.97

11   cin      8     3.82    11.82

12   pit     19     7.81    26.81

13   ten     32    -7.57    24.43

14   jax      8     4.76    12.76

15   sdg     -9     9.94     0.94

16   sea    -15     9.11    -5.89

17   ari      4    -4.98    -0.98

============================

AVERAGE  12.0    -1.20    10.80

============================
```

How to read this table: in week 1, the Colts beat the Ravens by 17. The Ravens were, all things considered, 1.83 points worse than average, so the Colts got a "score" of 17 - 1.83, or 15.17 for that game. In week 2, the Colts beat the Jaguars by 7. Jacksonville was 4.76 points better than average, so the Colts get an 11.76 for that game. Average their scores for each game and you've got their rating. The bottom line says:

> The Colts' won their games by an average of 12 points each. Their opponents were, on average, 1.2 points worse than average. Thus the Colts were 10.8 points better than average.

Let's examine some of the features of this system:

- **The numbers it spits out are easy to interpret** - if Team A's rating is 3 bigger than Team B's, this means that the system thinks Team A is 3 points better than Team B. With most ranking algorithms, the numbers that come out have no real meaning that can be translated into an English sentence. With this system, the units are easy to understand.

- **It is a *predictive* system rather than a *retrodictive* system** - this is a very important distinction. You can use these ratings to answer the question: which team is stronger? I.e. which team is more likely to win a game tomorrow? Or you can use them to answer the question: which of these teams accomplished more in the past? Some systems answer the first questions more accurately; they are called predictive systems. Others answer the latter question more accurately; they are called retrodictive systems. As it turns out, this is a pretty good predictive system. For the reasons described below, it is not a good retrodictive system.

- **It weights all games equally** - every football fan knows that the Colts' week 17 game against Arizona was a meaningless exhibition, but the algorithm gives it the same weight as all the rest of the games.

- **It weights all points equally, and therefore ignores wins and losses** - take a look at the Colts season chart above. If you take away 10 points in week 3 and give them back 10 points in week 4, you've just changed their record, but you haven't changed their rating at all. If you take away 10 points in week 3 and give back 20 points in week 4, you have made their record worse but their rating better. Most football fans put a high premium on the few points that move you from a 3-point loss to a 3-point win and almost no weight on the many points that move you from a 20-point win to a 50-point win.

- **It is easily imressed by blowout victories** - this system thinks a 50-point win and a 10-point loss is preferable to two 14-point wins. Most fans would disagree with that assessment.

- **It is slightly biased toward offensive-minded teams** - because it considers point margins instead of point ratios, it treats a 50-30 win as more impressive than a 17-0 win. Again, this is an assessment that most fans would disagree with.

- **This should go without saying, but** - I'll say it anyway. The system does not take into account injuries, weather conditions, yardage gained, the importance of the game, whether it was a Monday Night game or not, whether the quarterback's grandomother was sick, or anything else besides points scored and points allowed.

This system, like all systems, has some drawbacks, but it has the virtue of simplicity. It is easy to understand and it produces numbers that are easy to interpret. That is not to be sneezed at.

Furthermore, most of its drawbacks have easy fixes. For example, when computing a team's initial rating --- i.e. their average point margin --- you can tweak the individual game margins to make the initial rating "smarter." One way to do that is to cap the margin of victory at 21 points, or 14 points or whatever you want. You can explcitly incorporate wins and losses by giving the winning team a bonus of 3 points or 10 points or however many you want. To take it to the extreme, you could simply define all wins to be one-point wins and all losses to be one-point losses. This removes margin of victory from the scene completely. As usual, when you tweak the method to stengthen its weaknesses, you also weaken its strengths. In particular, if you use a modified margin of victory, the numbers don't have as nice an interpretation.

I'll close with some rankings. Here are the NFL's 2005 regular season rankings according to the original method:

```
    Team    Rating  StrOfSched

============================

 1. ind     10.8      -1.2

 2. den     10.8       2.2

 3. sdg      9.9       3.3

 4. sea      9.1      -2.2

 5. pit      7.8      -0.4

 6. nyg      7.5       0.7

 7. kan      7.0       2.1

 8. was      6.0       1.9

 9. car      5.1      -3.2

10. jax      4.8      -1.0

11. cin      3.8      -0.6

12. dal      3.2       2.1

13. nwe      3.1       0.6

14. chi      1.4      -2.2

15. mia     -0.8      -0.8
```

```
16. tam       -1.0      -2.6

17. atl       -1.2      -1.9

18. bal       -1.8       0.3

19. phi       -2.3       2.6

20. oak       -2.8       3.0

21. min       -3.5      -1.1

22. gnb       -3.7      -0.8

23. cle       -4.2       0.1

24. ari       -5.0      -0.2

25. ram       -5.1      -1.0

26. buf       -5.8       0.2

27. nyj       -6.4       0.8

28. det       -6.7      -1.0

29. ten       -7.6       0.1

30. hou      -10.0       0.7

31. nor      -11.1      -0.9

32. sfo      -11.1       0.7
```

Here they are if every win of less than 7 points is counted as a 7-point win and if the margin of victory is capped at 21.

```
    Team     Rating   StrOfSched

==============================

 1. den       10.1       1.6

 2. ind        9.9      -1.4

 3. sea        7.1      -1.9

 4. sdg        6.9       2.9

 5. nyg        6.3       0.7

 6. pit        6.1      -0.6

 7. was        5.5       1.6

 8. kan        5.4       1.7

 9. car        4.8      -2.3

10. jax        4.8      -1.1

11. cin        3.8      -0.9

12. dal        3.6       1.6

13. nwe        2.8       0.7
```

```
14. chi      1.5      -1.8

15. tam      0.9      -1.9

16. mia      0.6      -0.7

17. atl     -0.4      -1.3

18. min     -1.8      -1.1

19. phi     -1.9       2.1

20. cle     -3.2      -0.2

21. bal     -3.4       0.3

22. oak     -3.6       2.6

23. gnb     -4.9      -0.5

24. buf     -5.1       0.3

25. ram     -5.1      -0.7

26. ari     -5.1      -0.1

27. nyj     -5.8       0.8

28. det     -6.0      -0.8

29. ten     -6.7      -0.1

30. sfo     -8.1       0.5

31. nor     -9.2      -0.4

32. hou     -9.8       0.5
```

Here they are with margin of victory removed altogether:

```
   Team    Rating   StrOfSched

=============================

 1. den     0.69       0.07

 2. ind     0.66      -0.09

 3. sea     0.50      -0.12

 4. jax     0.42      -0.08

 5. nyg     0.42       0.04

 6. was     0.37       0.12

 7. pit     0.34      -0.03

 8. kan     0.33       0.08

 9. cin     0.31      -0.07

10. sdg     0.29       0.17

11. nwe     0.26       0.01
```

```
12. chi     0.26     -0.11

13. tam     0.25     -0.13

14. car     0.24     -0.13

15. dal     0.22      0.09

16. mia     0.06     -0.07

17. min     0.05     -0.07

18. atl    -0.06     -0.06

19. phi    -0.14      0.11

20. bal    -0.23      0.02

21. cle    -0.26     -0.01

22. ram    -0.28     -0.03

23. oak    -0.36      0.14

24. ari    -0.37      0.01

25. buf    -0.37      0.01

26. det    -0.41     -0.03

27. sfo    -0.44      0.06

28. nyj    -0.45      0.05

29. gnb    -0.49      0.01

30. ten    -0.50      0.00

31. nor    -0.63     -0.01

32. hou    -0.71      0.04
```

ADDENDUM: I need to clarify one thing about the simple rating system: it's not my system. I didn't invent it. In fact, it's one of those systems that has been around for so long that no one in particular is credited with having developed it (as far as I know anyway). People were almost certainly using it before I was born. I like the system and use it a lot because it's fairly easy to interpret and understand, and because the math behind it is nifty. But I just realized that I had never been clear enough about the fact that it's not my system. I just use it.

Related posts:

[2007 Standings: Simple Ranking System](#)

[Another ranking system](#)

[The Simple Rating System: 2007 Offenses and Defenses](#)

[Trivial observations from the simple rating system](#)

[Another rating system: maximum likelihood](#)

This entry was posted on Monday, May 8th, 2006 at 4:06 am and is filed under [BCS](#), [Statgeekery](#). You can follow any responses to this entry through the [RSS 2.0](#) feed. Both comments and pings are currently closed.