Dominance Matrices: Your Secret Weapon in Footy Tipping

Roger Walter —
Roger is the editor of *Vinculum*. Having spent 30 years teaching secondary mathematics he is now a freelance author/editor of educational publications.

Dominance matrices are a tool which can produce a more accurate ranking of teams when they haven't all played each other.

I must acknowledge Stephen Swift from Queensland for his work on matrices, and thank him for his permission to use his work, which is the inspiration for this article.

The general problem of ranking teams in a competition is that not all teams have played each other. One team high on the ladder may have been fortunate in playing mainly lower ranked teams. Dominance matrices are one tool which can take this into account and produce a more accurate ranking. This can be a very useful tool in predicting the results of matches early in the season, when fewer games have been played.

The simplest dominance matrix is a square matrix which lists the teams twice, giving 1 point for a win, ½ a point for a draw and 0 for a loss. For five teams, A, B, C, D and E a dominance matrix may appear as follows: error : 1/2

$$\mathbf{D} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \end{pmatrix} \end{array}$$

This matrix shows that A has beaten B and D, C and E have drawn, and A has not played C. The leading diagonal will be all zeros as a team does not play itself.

The matrix **1** is defined as a column vector with all 1's, so that **D1** will give the total number of points for each team.

$$\mathbf{D1} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \end{pmatrix} \end{array} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{1}{2} \\ 1 \\ 1 \\ 1\frac{1}{2} \end{pmatrix} \begin{array}{l} \text{1st} \\ \text{5th} \\ \text{not able to rank} \\ \\ \text{2nd} \end{array}$$

1 This is the usual method for most team ladders. Ranking is usually determined by $\mathbf{D1} - \mathbf{D^T1}$, where $\mathbf{D^T}$ is the transpose of **D**. This gives wins – losses, which for some teams will be negative. The final vector is called the first order ranking,

$$\mathbf{r}_1 = \mathbf{D1} - \mathbf{D^T1} = \begin{pmatrix} 2 \\ -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \begin{array}{l} \text{1st} \\ \text{Not able to rank} \\ \\ \\ \text{2nd} \end{array}$$

$\Sigma = 0$

The sum of the elements of first order rankings will always be zero.

If we consider $\mathbf{D^2}$, we get

$$\mathbf{D}^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \end{pmatrix} \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 1\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & 1 & 0 & \frac{1}{4} \end{pmatrix}$$

2nd hand score

* Squaring D1 will derive 2nd hand wining assumption. For example, team A beats team B who beats team C, the 2nd hand win (D^2) will assume that team A will beat team C. This is shown when multiplying the matrix, we get 1 1/2 in D (1,3)

If we examine the process, we will see that if we sum the top row, we get the total of A's individual scores multiplied by the scores of each team that A defeated (B and E).

Row A = (A d. B) × (B total) + (A d. D) × (D total)
= 1 × ½ + 1 × (½ + ½)
= 1½

Clearly, a team will gain a higher score if their opponents have done better, so beating 'quality' opponents will earn a higher score. It is useful for all teams to have played an equal number of games.

Teams can now be ranked by calculating $\mathbf{D^21} - (\mathbf{D^2})^\mathbf{T}\mathbf{1}$.

This will produce a second order ranking vector, in this case

$$\mathbf{r}_2 = \mathbf{D^21} - \left(\mathbf{D^2}\right)^\mathbf{T}\mathbf{1} = \begin{pmatrix} 1.5 \\ -0.5 \\ -2 \\ 0.5 \\ 0.5 \end{pmatrix} \begin{array}{l} \text{1st} \\ \text{4th} \\ \text{5th} \\ \text{not able to rank} \\ \\ \end{array}$$

We can now rank the middle three teams.

2 R =r1 +0.5R2 -> Clearer to rank now compared to just R1 and R2

```
> R
      [,1]
[1,]  2.75   1st  A
[2,] -1.25   4th  B
[3,] -2.00   5th  C
[4,] -0.75   3rd  D
[5,]  1.25   2nd  E
```

Scores are weighted according to the success of the other teams they played, so beating 'quality' opponents will earn a higher score.

* Subtract the games play (because it contribute to the vector) then divide by no of games. Dividing by no of games take into account for teams who have not played as many games as other teams

* We need to subtract the vector with the number of games each team has played because the vector will consists of the number from the diagonal which is the amount the games that the team has played. Which means that the more games a team plays, the more scores they will get if the diagonal wasn't subtracted.

So now, how do we combine the two rankings? 3 Stephen suggests the first order ranking should be weighted more heavily and suggests $\mathbf{r} = \mathbf{r}_1 + 0.5\mathbf{r}_2$. Others may feel that the second order ranking should stand on its own. What do you think? What do your students think? 4

if your opponents did well, and decreases if they did badly. This vector could be used to rank the teams, if you subtracted a column vector consisting of the number of games each team had played, then divided each team's score by the number of games that team had played. * first using D^2

you think? What do your students think? 4

This works well in cases like chess tournaments, where there are no scores, as such. In the case of team sports, it is possible to go further.

We can compare two teams playing each other by choosing the ratio of their scores. If A scored 10 points and B scored 3 points, we could say A is $\frac{10}{3}$, or about 3.3 times as good as B. We can rate B as $\frac{3}{10}$ as good as A, or 0.3 times A. The ratio rather than the difference should be chosen, to allow for high and low scoring games. This is a little like going by percentage instead of numbers of wins and losses. It would be an interesting class discussion as to which is most reliable when ranking abilities of teams. An example of six teams playing each other is given below. The leading diagonal will still be all zeros, and elements reflected in the leading diagonal will be reciprocals of each other.

$$\mathbf{D} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{pmatrix} 0 & \frac{10}{3} & \frac{5}{4} & 0 & \frac{2}{1} & 0 \\ \frac{3}{10} & 0 & 0 & \frac{5}{3} & 0 & 1 \\ \frac{4}{5} & 0 & 0 & \frac{2}{3} & \frac{4}{3} & 0 \\ 0 & \frac{3}{5} & \frac{3}{2} & 0 & 0 & \frac{6}{5} \\ \frac{1}{2} & 0 & \frac{3}{4} & 0 & 0 & \frac{10}{9} \\ 0 & 1 & 0 & \frac{5}{6} & \frac{9}{10} & 0 \end{pmatrix} \end{array}$$

The first order ranking would be $\mathbf{r}_1 = \mathbf{D1} - \mathbf{D^T1}$

Gives you points for winning and also punishes you for losing. The number of points will matter a lot more. For example, when transpose is taken into consideration, for example R (1,2) = 10/3-3/10

Space does not permit, but the reader may verify by calculating $\mathbf{D^2 1}$ that the score for A, the first element in the vector, will be as follows:

$$A = \tfrac{10}{3}\times\tfrac{3}{10}+\tfrac{5}{4}\times\tfrac{4}{5}+\tfrac{2}{1}\times\tfrac{1}{2}+\tfrac{10}{3}\times(\tfrac{5}{3}+1)+\tfrac{5}{4}\times(\tfrac{2}{3}+\tfrac{4}{3}) +\tfrac{2}{1}\times(\tfrac{3}{4}+\tfrac{10}{9})$$
$$= 3 + (A/B) \times (B \text{ total}) + (A/C) \times (C \text{ total}) + (A/E) \times (E \text{ total})$$

This is almost the same, except that for each team the leading diagonal will be the number of games played. The rest of the score consists of their score multiplied by the total of the scores of the teams they have played. Once again, your score improves

ie. $AB \times BA = \frac{10}{3} \times \frac{3}{10} = 1$

number of games that team had played.

If you calculate $\mathbf{r}_2 = \mathbf{D^2 1} - (\mathbf{D^2})^\mathbf{T}\mathbf{1}$ the number of teams played will cancel. I believe this gives the 5 best ranking of teams, if they've all played the same number of games. Where teams have played different numbers of games, you should divide each score by the number of games played.

vs just using D^2

The diagonal will be the number of games. Since the diagonal stays the same when doing the transpose, it will become 0 after subtracting

~~As before, the relative weighting of the first and second order rankings is an interesting point for discussion, as well as the meaning and usefulness of third and higher order rankings.~~

~~An interesting application of the dominance matrix method has been to rank students in Queensland for tertiary entrance. Students usually do 4–6 subjects where they receive an overall position, or OP, which is a ranking from 1 to 25. Some students do seven OP subjects. Students choose from about 25 different subjects, and as in Victoria some are easier than others. A computer is used, and though the calculations are the same, matrices as such are not used, as the numbers (over 30 000) are too high.~~

~~The first year, they did a successful trial run using the previous year's students. Unfortunately a small increase in numbers in that year meant that the program crashed because the data was too large for the computer. This delayed the rankings by about 3 weeks while they waited to use a larger government computer!~~

~~Dominance matrices are an interesting application of matrices. If anyone is interested, I am preparing an excel spreadsheet which will use 20 × 20 matrices to rank AFL teams. I am happy to provide this free of charge by email. You may request copies be emailing me at **vinculum@mav.vic.edu.au**. All I ask is that you include in your email at least one thing you can share from your teaching experience.~~

~~**If anyone is interested, I am preparing an excel spread sheet which will use 20 × 20 matrices to rank AFL teams.**~~

~~**vinculum@mav.vic.edu.au**~~

Reference

Brodie, R. & Swift, S. (2009). *NewQMaths 11c* (3rd Ed.). (pp. 186–189). Qld. Nelson Cengage Learning.