Keener Method

Introduction

The Keener method’s purpose is to relate the rating for a given team to the absolute strength of the team, which in turn depends on the relative strength of the team. The relative strength of the team is the strength of the team relative to the strength of the teams that it has played against. Keener bases everything on two stipulations that govern the relationship between a team’s strength and its rating. The first stipulation is that the strength of a team should be gauged by its interactions with opponents together with the strength of these opponents. Another stipulation is that the rating for each team in a given league should be uniformly proportional to the strength of the team.

An Example

Let’s look at an example of five teams from the 2012 Division III college football season: Johns Hopkins, Franklin & Marshall, Gettysburg, Dickinson and McDaniel. Consider following chart of these teams’ (in the rows) wins and losses against teams (in the columns):

<table>
<thead>
<tr>
<th>Teams</th>
<th>Johns Hopkins</th>
<th>F &amp; M</th>
<th>Gettysburg</th>
<th>Dickinson</th>
<th>McDaniel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johns Hopkins</td>
<td>-</td>
<td>Loss 12-14</td>
<td>Win 49-35</td>
<td>Win 49-0</td>
<td>Win 49-7</td>
</tr>
<tr>
<td>F &amp; M</td>
<td>Win 14-12</td>
<td>-</td>
<td>Loss 31-38</td>
<td>Win 36-28</td>
<td>Win 35-10</td>
</tr>
<tr>
<td>Gettysburg</td>
<td>Loss 35-49</td>
<td>Win 38-31</td>
<td>-</td>
<td>Loss 23-13</td>
<td>Win 35-3</td>
</tr>
<tr>
<td>Dickinson</td>
<td>Loss 0-49</td>
<td>Loss 28-36</td>
<td>Win 13-23</td>
<td>-</td>
<td>Win 38-31</td>
</tr>
<tr>
<td>McDaniel</td>
<td>Loss 7-49</td>
<td>Loss 10-35</td>
<td>Loss 3-35</td>
<td>Loss 31-38</td>
<td>-</td>
</tr>
</tbody>
</table>

There are several ways to rank these teams, not using the Keener Method. You can choose any rankings to start so let’s look at the following one:
1. Johns Hopkins
2. Franklin and Marshall
3. Gettysburg
4. Dickinson
5. McDaniel

1 Using this ordering, we set $S_{ij}$ equal to the cumulative number of points scored by team i against team j. In the matrix you will have that number in row i, column j. If you would like to at this point you may apply a skewing function. You do not need to normalize this because all teams have played the same number of games. After this, we can just use the math we discuss below in order to come up with a new set of rankings. This will give us the rankings that Keener has come up with for these 5 teams that have each played each other once. As you can see, the beginning ranking does not matter. We can start off with any order we want for the five teams.

Try this ranking instead:

1. Dickinson
2. Johns Hopkins
3. Franklin and Marshall
4. McDaniel
5. Gettysburg
Doing the same thing we did above, but starting with this ranking system, we will still come out with the same rankings for the Keener Method. The only difference would be that the matrix may have the teams in a different order.

This is just a small 5 team example, but of course Keener’s method could be used with many more teams. It would be a little bit more challenging gathering all the data and putting it into a matrix, but once it is there, it all comes down to the same Math. Once again, we must make sure that the matrix is normalized so that every team plays the same number of games which may be difficult in other scenarios.

The Math

When it comes to applying this method to a much larger example it becomes much more difficult. We must first sign some variables to quantities in question. Let $a_{ij}$ be equal to the value of the statistic produced by team $i$ when competing with team $j$. For ties, assign a value of 1/2 for each time they have tied. Let $W_{ij}$ be the number of times team $i$ has beaten team $j$ and $T_{ij}$ is the number of times that teams $i$ and $j$ have tied. This way you might set $a_{ij}$ equal to $W_{ij}$ plus $T_{ij}$ divided by 2. You could also set $a_{ij}$ equal to $S_{ij}$ if you think that a more relevant attribute that reflects team $i$’s power relative to that of team $j$. If the teams play more than once, then $S_{ij}$ is the cumulative number of points scored by $i$ against $j$.

When comparing teams $i$ and $j$ it is better to take into account the total number of points scored by setting $a_{ij}$ equal to $S_{ij}$ divided by the quantity $S_{ij}$ plus $S_{ji}$. Keener points out in his method that we should really be using something like $a_{ij}$ equals the quantity $S_{ij}$ plus 1 divided by the quantity $S_{ij}$ plus $S_{ji}$ plus 2.

Skewing concerns the issue of how to compensate for the situation in which a stronger team mercilessly runs up the score against a weaker opponent to either enhance their own rating or just ”rub it in.” Keener suggests using a nonlinear skewing function such as $h(x)$ is equal
to $1/2$ plus the quantity $\sin(x-1/2)$ times the square root of the absolute value of $2x-1$ all divided by 2. Skewing is not always needed. Skewing does not really affect the NFL rankings due to the fact that the NFL was well balanced for the years that we are talking about. Things are obviously different for NCAA sports.

One last thing after skewing is normalization of the $a_{ij}$ in which not all teams play the same number of games. In this case, you must make the replacement $a_{ij} = a_{ij} / n_i$ where $n_i$ is the number of games played by team $i$. To understand why this is necessary, teams playing more games than other teams have the possibility of producing more points, and thus inducing larger values of $a_{ij}$. Once the $a_{ij}$'s are defined, skewed, and normalized, you organize them into a square $m$ by $m$ matrix where $m$ is the number of teams in the league. Then, you construct a rating value for each of $m$ teams based on their perceived "strength" at some time $t$ during the current playing season. You let $r(t)$ denote the column containing each of these $m$ ratings.

Recall that Keener’s first stipulation is that the strength of a team should be measured by how well it performs against opponents but tempered by the strength of those opponents. The relative strength of team $i$ compared to team $j$ is defined as $s_{ij} = a_{ij} r_j$. The absolute strength of team $i$ is defined to be the sum of the relative strength.

Keener’s second stipulation concerning the relationship between strength and rating requires that the strength of each team be uniformly proportional to the team’s rating in the sense that there is a proportionally constant lamda such that $s_i = \lambda r_i$. In terms of the ratings vectors $r$ and $s$ this says that $s = \lambda r$. We know that $s = Ar$ so the conclusion is that the ratings vector $r$ is related to the statistics $a_{ij}$ in matrix $A$ by the equation $Ar = \lambda r$. This says that the ratings vector $r$ must be an eigenvector, and the constant lamda must be an associated eigenvalue for matrix $A$.

Keener imposes three mild constraints on the amount of interaction between teams and on the resulting statistics $a_{ij}$ in $A$. One is nonnegativity. This means that whatever attribute you use to determine the statistics $a_{ij}$, in the end you must ensure that each statistic $a_{ij}$
is a nonnegative number. Another constraint is Irreducibility. There must be enough past competition within the league to ensure that it is possible to compare any pair of teams, even if they hadn’t played each other. One last constraint is primitivity. This is just a more stringent version of the irreducibility constraint. We now require each pair of teams be connected by a uniform number of games.

By imposing the first 2 constraints above, the Perron-Frobenius theory can be brought to bear to extract a unique ratings vector from Keener’s equation $Ar = \lambda r$. The value $\lambda$ is called the Perron value and the vector $r$ is called the Perron vector.

Finally, here is a quick summary of Keener’s method. First, begin by choosing one attribute of the sport under consideration that you think will be a good basis for making comparisons of each item’s strength relative to that of the other teams in the competition. Next, whatever attribute is chose, compile statistics from past competitions, and set $a_{ij} =$ the value of the statistic produced by team $i$ when competing against team $j$. After that, massage the raw statistics $a_{ij}$ in step 2 to account for anomalies. Then, if after massaging the data you feel that there is an imbalance in the sense that some $a_{ij}$ ’s are very much larger or smaller than they should be, then restore a balance by constructing a skewing function. Next, if not all teams have played the same number of games, account for this by normalizing $a_{ij}$ by making the replacement $a_{ij}$ divided by $n_i$ where $n_i$ is the number of games played by team $i$. Also, organize these teams into a nonnegative matrix $A$. Next, you must make sure that matrix $A$ satisfies the irreducibility and primitivity conditions or else, perturb $A$ by adding some artificial games with negligible game statistics. Lastly, compute the rating vector $r$ by using the power method. If you have forced irreducibility or primitivity, then slightly modify the power iteration.
Advantages and Disadvantages

The advantages of this method, compared to some others is that it is a good way to compare teams directly to other teams. It is an easy way to find a rating for each team by comparing them directly to their success against another team.

A disadvantage to this method is the constraints explained above. At different points in a season, there are teams that play more than other teams. I know that Keener has ways around this, but it still is a problem because the teams do not play the same amount of games in everything we may want to look at.

Possible Adjustments

There are some small adjustments we could make to this method to make it a little better. One main thing that I could think of is that it should still be able to work if the teams don’t play the same number of games. This way you would not have to fix the matrix as to include some unnecessary ”imaginary” games. Also, another adjustment could be to make it worth more if you beat a team by more points. For example, you should maybe receive a little more reward for beating a team by 50 points as to beating a team by 7 points. Those are just a few possible adjustments that could possibly make this method a little better.

Conclusion

Overall, this is a very good ranking method. It allows you to start with any rankings that you would like, and use the results you have with some math to come up with a logical ranking for the teams that you have in mind. It may be frustrating at times if the teams do not play the same amount of games to get your matrix ready. All in all, it still is a very
good way to come up with a logical ranking for whatever you are looking at.

**Additional Reading**

If you would like to learn more about the specifics of this method, here are some links to papers and websites that explore the topic more thoroughly than we do here.

This first link is to an article that explains more of the math behind this method.

[https://umdrive.memphis.edu/ccrousse/public/MATH%207375/PERRON.pdf](https://umdrive.memphis.edu/ccrousse/public/MATH%207375/PERRON.pdf)

This second link shows how you can use Keener’s method to rank NFL teams.