"Compughter Ratings" Theory

Steven M. Pugh

In this paper, we devise a mathematical model for measuring the past performance of distinct teams (or individuals) competing against each other in a series of games. The resulting ratings can be used to assign each competitor an ordinal ranking from highest performer to lowest performer. This ordinal ranking is useful for seeding teams into playoff positions and/or determining the champion of a season in which no playoff is performed.

We consider a single season consisting of p teams competing in a total of n games overall. We assume that all p teams participate in at least one game and that there is no subset of the teams which competes independently from the remainder of the field.

Let X' denote the $(n \times p)$ design matrix representing the wins and losses in the competition. With the columns of X' representing the p teams, we construct the rows of X' such that the i^{th} row represents the i^{th} game of the season. Wins are denoted with a 1 and losses are denoted with a -1. If the q^{th} team ($1 \le q \le p$) did not compete in the i^{th} game, then we use 0 to denote the absence of the q^{th} team from the i^{th} game.

To promote clarity around the presentation of the design matrix, we express X' in the following way:

$$X' = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}_{(n \times p)}$$

Proposition 1.0. *Let X represent the $(n+1) \times p$ matrix with the first n rows of X identical to those of X' (as defined above) and the $(n+1)^{th}$ row consisting of the scalar value 1 in every column:*

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \\ 1 & 1 & 1 & 1 \end{bmatrix}_{(n+1) \times p}$$

Then $rank(X) = p$.

Proof: We know from elementary linear algebra that since X is an $(n+1) \times p$ matrix, $rank(X)$ is at most

$\min(n+1, p)$. We also observe that $n \geq p/2$ since we assumed that each team played at least one game. Furthermore, since we assumed that no subset of the teams played independently from the remainder of the field, we know that $n \geq p/2 + (p/2 - 1) = p - 1$. Hence, $n + 1 \geq p$ and so $\min(n+1, p) = p$. Therefore, $rank(X)$ is at most p. So to prove that $rank(X)$ is *exactly* p, all we have to do is verify that $rank(X)$ cannot be less than p. We can prove this by contradiction. Suppose that $rank(X) < p$. Then there must exist at least two columns of X which are linearly dependent. That is, by definition, there must exist two distinct column vectors of X, $\bar{x}_i$ and $\bar{x}_j$, and scalar values a_i and a_j, not both zero, such that $a_i\bar{x}_i + a_j\bar{x}_j = 0$. Since the $(n+1)^{th}$ element of both vectors is known to be 1, it is not possible for a_i or a_j to be equal to zero. So there must exist some non-zero scalar $c = \dfrac{a_j}{a_i}$ such that $\bar{x}_i = c\bar{x}_j$. Again, since the $(n+1)^{th}$ element of both vectors is known to be 1, the only possible value for c is 1. So $\bar{x}_i$ and $\bar{x}_j$ must be identical. Now, the only vector which could possibly satisfy this condition is $(0,0,0,...,0,1)^T$, because if any two row elements of $\bar{x}_i$ and $\bar{x}_j$ were equal and non-zero, this would contradict the very design of X. To see this, just recall that each row of X has exactly two non-zero values and those values are 1 (indicating a win) and -1 (indicating a loss). We have thus concluded that $\bar{x}_i = \bar{x}_j = (0,0,0,...,0,1)^T$. But this implies that teams p_i and p_j did not participate in any games, since the first n rows of each vector do not contains any 1s (wins) or -1s (losses). This contradicts our assumption that all p teams participated in at least one game. We therefore conclude that $rank(X) = p$. That is, X has full rank. ∎

Proposition 1.0 is an important result, because we can now use it to make a very powerful statement about $X^T X$. From elementary linear algebra, we know that $X^T X$ has the same rank as X. So since X has full rank, $X^T X$ must also have full rank. In this case, since $X^T X$ is $p \times p$ and $rank(X^T X) = p$, we know that $X^T X$ is invertible. We will make a note of this conclusion for later use in this paper.

Returning to our model, recall that we have represented our newly formed design matrix X in the following way:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \\ 1 & 1 & 1 & 1 \end{bmatrix}_{(n+1)\times p}$$

Now let y denote the $(n+1)\times 1$ column vector which represents the dependent variable. If we assume that all games played are equal in terms of outcome (i.e., that no factors impact the weight of a win), then we can populate y with any arbitrary positive value as long as $y_1 = y_2 = \cdots = y_{(n+1)}$. For simplicity, we choose the value of 1 such that $y_i = 1$ for all $1 \le i \le (n+1)$. We then represent the dependent variable as:

$$y = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(n+1)\times 1}$$

If we think of the columns of X as a set of p independent variables, then essentially we have a system of $(n+1)$ equations in p unknowns. We can represent this general linear regression model as:

$$y_{(n+1)\times 1} = X_{(n+1)\times p}\beta_{p\times 1} + \varepsilon_{(n+1)\times 1}$$

where:

- y is the response vector;
- β is a vector of parameters;
- X is a matrix of constants;
- ε is a vector of independent normally distributed random variables with expectation $E\{\varepsilon\} = 0$

Let b represent the column of coefficients (parameters $\beta_1, \beta_2, ..., \beta_p$) which minimizes the general linear regression model. The least squares normal equations for the general linear regression model are known to be:

$$X^T Xb = X^T y$$

Since $X^T X$ is invertible (from our results above), we can multiply both sides of the equation by $(X^T X)^{-1}$ to yield the vector of least squares estimators (i.e., the unique solution to the equation):

$$b = (X^T X)^{-1} X^T y$$

The independent and dependent variable values on the right hand side of the equation are known constants. Hence, the solution b can be computed easily with any mathematical software package.

Now, let X_i denote the i^{th} row of X. Then the dot product of $X_i b$ is a scalar value. So for $1 \le i \le (n+1)$, we can compute each $\hat{\pi}_i$ as:

$$\hat{\pi}_i = \frac{e^{X_i b}}{1 + e^{X_i b}}$$

We now construct the $(n+1)\times(n+1)$ diagonal matrix V such that the diagonal values are given by:

$$V_{ii} = \hat{\pi}_i(1-\hat{\pi}_i)$$

for $1 \le i \le (n+1)$.

Then V can be represented as:

$$V = \begin{bmatrix} \hat{\pi}_1(1-\hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2(1-\hat{\pi}_2) & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \hat{\pi}_{n+1}(1-\hat{\pi}_{n+1}) \end{bmatrix}$$

Let r represent the $(n+1)\times 1$ column vector of residual values:

$$r = (y - \hat{\pi})$$

Since V is a square diagonal matrix with non-zero diagonal entries, it is known to be invertible. Therefore, we can compute the $(n+1)\times 1$ column vector z as:

$$z = Xb + V^{-1}r$$

Note that since V is invertible and X is of full rank, X^TVX must also be of full rank and, thus, invertible. We can now compute our $(p \times 1)$ ratings vector as:

$$g = (X^TVX)^{-1}X^TVz$$

where g_j represents the rating of the j^{th} team, $1 \le j \le p$.

REFERENCES

[1] S. H. Friedberg, A. J. Insel, and L. E. Spence, Linear Algebra (Second Edition), Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[2] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, Applied Linear Statistical Models (Fifth Edition), McGraw-Hill/Irwin, New York, New York, 2005, 222-223.

[3] D. Hosmer and S. Lemeshow, Applied Logistic Regression (Second Edition), John Wiley and Sons, New York, New York, 2000, 128-130.