|  |
| --- |
| **Generalized Latent Team Ratings Using the Expectation-Maximization Algorithm** |

**Robert Sawyer**

Department of Computer Science

North Carolina State University

Raleigh, NC 27695

[rssawyer@ncsu.edu](mailto:rssawyer@ncsu.edu)

**Abstract**

Recent years have seen an interest in estimating a team or player’s skill level from their observed performance against other teams or players. For example, Elo ratings (popularized in team sports by Nate Silver and ESPN’s Power Indices) are a method for estimating a team’s rating using differences in actual and expected performance. In a mathematical context, these team ratings are latent variables, they are never directly observed, but their values are inferred from observations. In this work, two general predictive models of margin of victory are proposed and used for calculating team latent ratings given game scores. The *Margin* model predicts the margin of victory for the home team using one latent variable, representing team skill, per team in the league. The *Joint* model predicts both the away score and home score using two latent variables, one for offensive skill and one for defensive skill, per team in the league. These models are both trained with the Expectation-Maximization (EM) algorithm. More specifically, the Expectation step (E-step) calculates the optimal latent team ratings given the regression model parameters, and the Maximization step (M-step) calculates the optimal model parameters given the latent team ratings. In order to create a generalizable model for latent team ratings, the Gaussian likelihood of multiple linear regression is used for the regression model. This work demonstrates several theoretical benefits of this general EM algorithm. First, this work demonstrates how extra input features, such as days of rest for each team, are naturally incorporated under the model. Second, this work demonstrates how prior distributions over the latent team ratings can be used to produce maximum a posteriori (MAP) estimates as opposed to maximum likelihood (MLE) estimates, which are especially useful in sports with smaller sample sizes or predictions at the start of a season. Third, this work demonstrates the wide applicability of the model, applying it to both the NBA and Men’s D1 Lacrosse. Accuracy results computed on time-series cross validation data demonstrate comparable accuracy to sportsbooks’ spread baselines and superior accuracy to several Elo-based baselines on predicting the margin of victory for the home team and binary classification of home team victory or defeat in both sports. The final ratings from the model are also produced to illustrate the interpretability of the latent team rating output, including a measure of uncertainty about each team’s latent rating. This model provides an important accuracy baseline due to the easily accessible type of data required, away and home team scores, that allows application to a diverse range of sports and provides flexibility for additional features and prior distributions over team ratings. This model also produces team ratings from only the away and home team scores that can be used to aid ranking systems for all sports, but importantly for sports receiving less data analysis (e.g. Lacrosse) than the big four sports. *Word count: 495*

Table 1. Model accuracies as measured by time-series cross-validation R-squared, averaged on 5% test intervals over the second half of each season. Top portion accuracy reflects accuracy on margin predictions and bottom portion predicts Away or Home Score. Baselines for lacrosse use average goal margin because of lack of spread and over/under data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **2008-NBA** | **2009-NBA** | **2010-NBA** | **2011-NBA** | **2016-Lax** |
| Margin | 0.240 | 0.172 | 0.170 | 0.187 | 0.324 |
| Margin Baseline | 0.277 | 0.200 | 0.194 | 0.220 | 0.120 |
| Joint | 0.241 | 0.170 | 0.172 | 0.133 | 0.318 |
| Baseline (Away) | 0.291 | 0.299 | 0.270 | 0.236 | - |
| Baseline (Home) | 0.348 | 0.261 | 0.204 | 0.210 | - |
| Joint (Away) | 0.289 | 0.231 | 0.153 | 0.0816 | 0.281 |
| Joint (Home) | 0.228 | 0.239 | 0.218 | 0.0942 | 0.090 |

Table 2. Top 10 ranking according to ratings produced from the single-latent variable Margin model. NCAA Rank (from ncaa.com), Average Goal Margin, and output from the two-latent variable Joint model are included for comparison.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Rank** | **Team** | **NCAA Rank** | **Average Goal Margin** | **Latent Rating-Margin\*** | **Latent Offense-Joint** | **Latent Defense-Joint** |
| 1 | Albany | 5 | 5.89 | 7.20 | 6.06 | 1.79 |
| 2 | Maryland | 1 | 3.68 | 7.06 | 3.71 | 3.93 |
| 3 | Denver | 3 | 4.76 | 6.84 | 4.72 | 2.76 |
| 4 | Duke | 7 | 4.67 | 6.78 | 4.06 | 3.46 |
| 5 | Loyola | 12 | 4.00 | 5.25 | 3.02 | 1.14 |
| 6 | Ohio State | 2 | 2.95 | 5.15 | 2.39 | 3.34 |
| 7 | Notre Dame | 8 | 0.73 | 4.77 | 2.10 | 3.17 |
| 8 | Penn State | 10 | 2.63 | 4.33 | 4.10 | 0.96 |
| 9 | UNC | 9 | 1.06 | 4.25 | 3.01 | 1.81 |
| 10 | Yale | 11 | 2.63 | 4.24 | 3.21 | 1.68 |
|  |  |  |  |  | \* sorting variable | |