STATS 50: Mathematics of Sport **Spring 2019**

Week 5 – Regularized Adjusted Plus Minus

Lecturer: Maxime Cauchois

Warning: *these notes may contain factual errors*

1 Adjusted Plus-Minus

In lots of team sports, data analysts want to evaluate the impact of each player on their team, while going beyond some raw statistics. One sensible number would be to compute for each player the difference of points scored between both teams while the given player was on the floor. For instance in figure 1, one can observe the "impact" of each player during each sequence when he was on the floor.

This first order statistics still has a major flaw: it fails to account for interactions between players on the floor at the same time. For instance, a player who is always on the floor at the same time as weaker players might end up with a lower plus-minus than others, while still having a huge impact on the team's success.

The idea is therefore to try and attribute to each player in the team "part" of its success, that is find a set of coefficients β such that the contribution of the j-th player on the floor is just β_j.

In order to find those β_j, one approach is to record, during each possession, all the players on the floor during as well as the score difference between both teams.

Then, for each possession, if the score difference is y_i, we would like to consider the following model:

$$y_i = \alpha + \beta_{H_1} + \cdots + \beta_{H_k} - (\beta_{A_1} + \cdots + \beta_{A_k}) + \epsilon_i \tag{1}$$

Here, H_1,..., H_k represent the players on the floor for the home team, while A_1,...,A_k are the ones playing for the away team during this possession. α, on the other hand, is meant to account for the home advantage.

This yields a linear model with the following matrix notation:

$$Y = \alpha \mathbf{1} + X\beta + \epsilon \tag{2}$$

where

$$X = \begin{bmatrix} x_{11} & \dots & x_{1k_A} & z_{11} & \dots & z_{1j_B} \\ x_{21} & \dots & x_{2k_A} & z_{21} & \dots & z_{2j_B} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ x_{n1} & \dots & x_{nk_A} & z_{n1} & \dots & z_{nj_B} \end{bmatrix} \text{ where } \begin{cases} x_{ij} = \begin{cases} 1 & \text{if } j \text{ plays for A during } i\text{-th possession} \\ 0 & \text{if } j \text{ plays for A during } i\text{-th possession} \end{cases} \\ z_{ij} = \begin{cases} -1 & \text{if } j \text{ plays for B on } i\text{-th possession} \\ 0 & \text{if } j \text{ plays for B on } i\text{-th possession} \end{cases} \end{cases}$$

Let's consider the following example:

Here, there is no home-away advantage so we can consider $\alpha = 0$. The matrix X formed on our model is as follows:

Notice that when we solve a linear regression, we precisely want to obtain β such that:

$$X^T X\beta = X^T y$$

where y is the vector of score differences during each stint (or each possession). We can then remark that $X^T X$ is none other that a matrix where the diagonal terms represent how many stints were played by each player, whereas each extra-diagonal term $(X^T X)_{ij}$ is the number of stints played by the player i and the player j together. For instance, one can check in figure 4that the 1st player in team A played 8 stints out of 11, 5 with the 3rd player in team A and 6 against the 4th player in team B. On the other hand, $(X^T y)_i$ is simply the plus-minus of player i (or a per hundred possessions version of it).

As a result, we end up solving the following equations for each player i:

$$PM_i = N_i \beta_i + \sum_{j \neq i} N_{ij}^t \beta_j - \sum_k \beta_k N_{ik}^o$$

where N_i is the number of stints played by i, N_{ij}^t is the number of stints played by i alongside with j, and N_{ik}^o is the number of stints played against player k.

Eventually, the adjusted plus-minus differs from the standard plus-minus because it accounts for interactions between players: if you face stronger opponents or play with weaker players, you won't be as much penalized for a low personal plus-minus, as part of it will be explained by other factors.

In our example, the coefficients β we eventually get seem very large for what they are supposed to represent (see figure 5), that is the points difference brought by each each player when compared with an average player over 100 possessions. This naturally leads to the idea of a regularized version of the adjusted plus minus.

2 Regularized Adjusted Plus-Minus

In the regularized version, one adds a penalization for too large β:

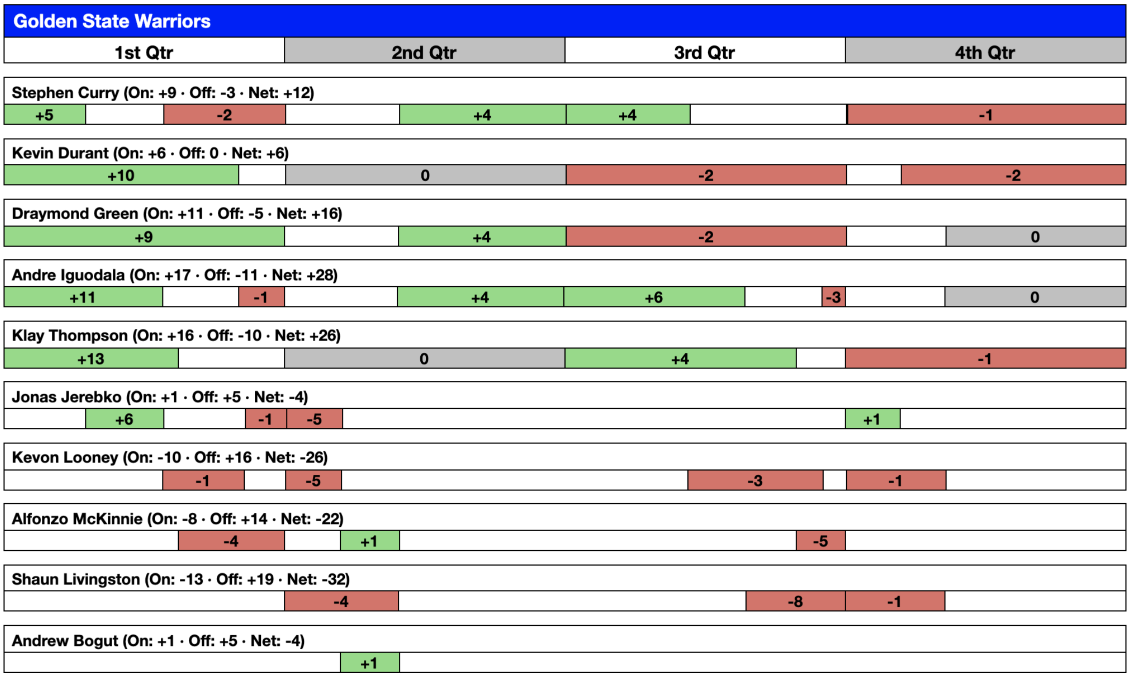
$$\hat{\beta}_\lambda = \underset{\beta}{\operatorname{argmin}} \|Y - \alpha \mathbf{1} - X\beta\|_2^2 + \lambda \|\beta\|_2^2$$

As observed in figure 6, this tends to shrink all β_j towards 0, following the same pattern as in regression to the mean. However, note that the vector $\hat{\beta}_\lambda$ does not represent the average difference of points scored during 100 possessions with respect to an average player, since its value heavily depends on our parameter λ. In short, if one wants to report the values of those coefficients, one also has to mention the value of the regularization term λ, which affects our estimate.

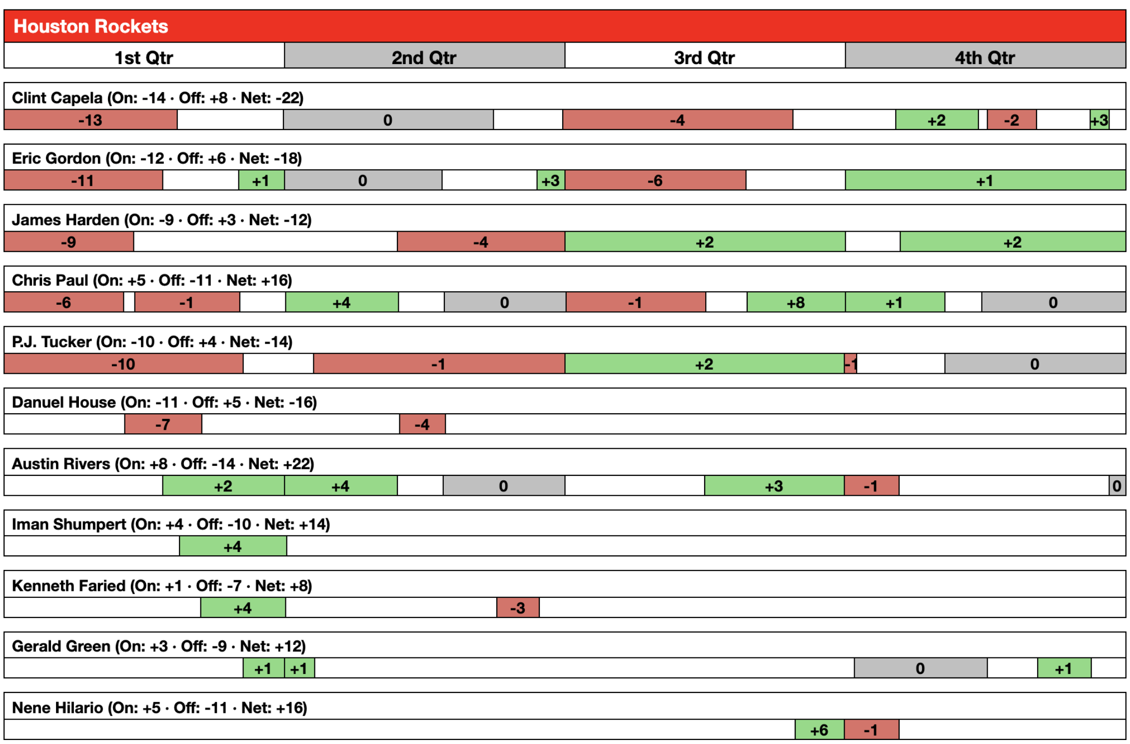
Eventually, suppose that ones wants to deduce a player ranking from the values of $\hat{\beta}_\lambda$. Then, one can observe that the said ranking would not be invariant of the regularization parameter λ! This should drive us to be extra careful when it comes to conclusions such that: "player i is better than player j" only based on the observation of the regularized adjusted plus-minus, because it also assumes an underlying choice of λ.

References

[1] Jacobs Justin. https://squared2020.com/2017/09/18/deep-dive-on-regularized-adjusted-plus-minus-i-introductory-example/.



(a) Golden State Warriors



(b) Houston Rockets

Figure 1: Plus-Minus during Game 2 of the 2019 West Conference Semi-Finals

- A1, A2, A3 vs. B1, B2, B3 resulted in 11-5 over 15 possessions
- A1, A2, A3 vs. B1, B2, B4 resulted in 2-4 over 8 possessions; current score 13 – 9
- A1, A4, A5 vs. B1, B4, B5 resulted in 7-1 over 10 possessions; current score 20 – 10
- A1, A4, A5 vs. B3, B4, B5 resulted in 6-6 over 17 possessions; current score 26 – 16
- A2, A3, A4 vs. B2, B3, B4 resulted in 4-7 over 10 possessions; current score 30 – 23
- A2, A3, A5 vs. B2, B3, B5 resulted in 5-7 over 16 possessions; current score 35 – 30
- A1, A2, A3 vs. B2, B4, B5 resulted in 4-5 over 8 possessions; current score 39 – 35
- A1, A2, A3 vs. B1, B4, B5 resulted in 4-7 over 10 possessions; current score 43 – 42
- A3, A4, A5 vs. B1, B4, B5 resulted in 4-0 over 8 possessions; current score 47 – 42
- A1, A3, A4 vs. B1, B2, B3 resulted in 5-2 over 10 possessions; current score 52 – 44
- A1, A4, A5 vs. B2, B3, B4 resulted in 2-9 over 14 possessions; final score 54 – 53

Figure 2: Example of 11 stints

	PlayerA1	PlayerA2	PlayerA3	PlayerA4	PlayerA5	PlayerB1	PlayerB2	PlayerB3	PlayerB4	PlayerB5
	Float64	Float64	Float64	Float64	Float64	Float64	Float64	Float64	Float64	Float64
1	1.0	1.0	1.0	0.0	0.0	-1.0	-1.0	-1.0	0.0	0.0
2	1.0	1.0	1.0	0.0	0.0	-1.0	-1.0	0.0	-1.0	0.0
3	1.0	0.0	0.0	1.0	1.0	-1.0	0.0	0.0	-1.0	-1.0
4	1.0	0.0	0.0	1.0	1.0	0.0	0.0	-1.0	-1.0	-1.0
5	0.0	1.0	1.0	1.0	0.0	0.0	-1.0	-1.0	-1.0	0.0
6	0.0	1.0	1.0	0.0	1.0	0.0	-1.0	-1.0	0.0	-1.0
7	1.0	1.0	1.0	0.0	0.0	0.0	-1.0	0.0	-1.0	-1.0
8	1.0	1.0	1.0	0.0	0.0	-1.0	0.0	0.0	-1.0	-1.0
9	0.0	0.0	1.0	1.0	1.0	-1.0	0.0	0.0	-1.0	-1.0
10	1.0	0.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	0.0	0.0
11	1.0	0.0	0.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	0.0

Figure 3: Matrix X of players on the floor

8.0	4.0	5.0	4.0	3.0	-5.0	-5.0	-4.0	-6.0	-4.0
4.0	6.0	6.0	1.0	1.0	-3.0	-5.0	-3.0	-4.0	-3.0
5.0	6.0	8.0	3.0	2.0	-5.0	-6.0	-4.0	-5.0	-4.0
4.0	1.0	3.0	6.0	4.0	-3.0	-3.0	-4.0	-5.0	-3.0
3.0	1.0	2.0	4.0	5.0	-2.0	-2.0	-3.0	-4.0	-4.0
-5.0	-3.0	-5.0	-3.0	-2.0	6.0	3.0	2.0	4.0	3.0
-5.0	-5.0	-6.0	-3.0	-2.0	3.0	7.0	5.0	4.0	2.0
-4.0	-3.0	-4.0	-4.0	-3.0	2.0	5.0	6.0	3.0	2.0
-6.0	-4.0	-5.0	-5.0	-4.0	4.0	4.0	3.0	8.0	5.0
-4.0	-3.0	-4.0	-3.0	-4.0	3.0	2.0	2.0	5.0	6.0

Figure 4: Matrix $X^T X$ of interactions

```
10-element Array{Float64,1}:
  -7.8927278116679656
  24.758944753644197
 -39.89921002061191
  55.83771154947786
 -15.30404792675653
 -49.40926009168065
 -10.576706771118376
  20.426554810534228
  60.84759824384196
 -38.797402339031706
```

Figure 5: Adjusted Plus-Minus of each player

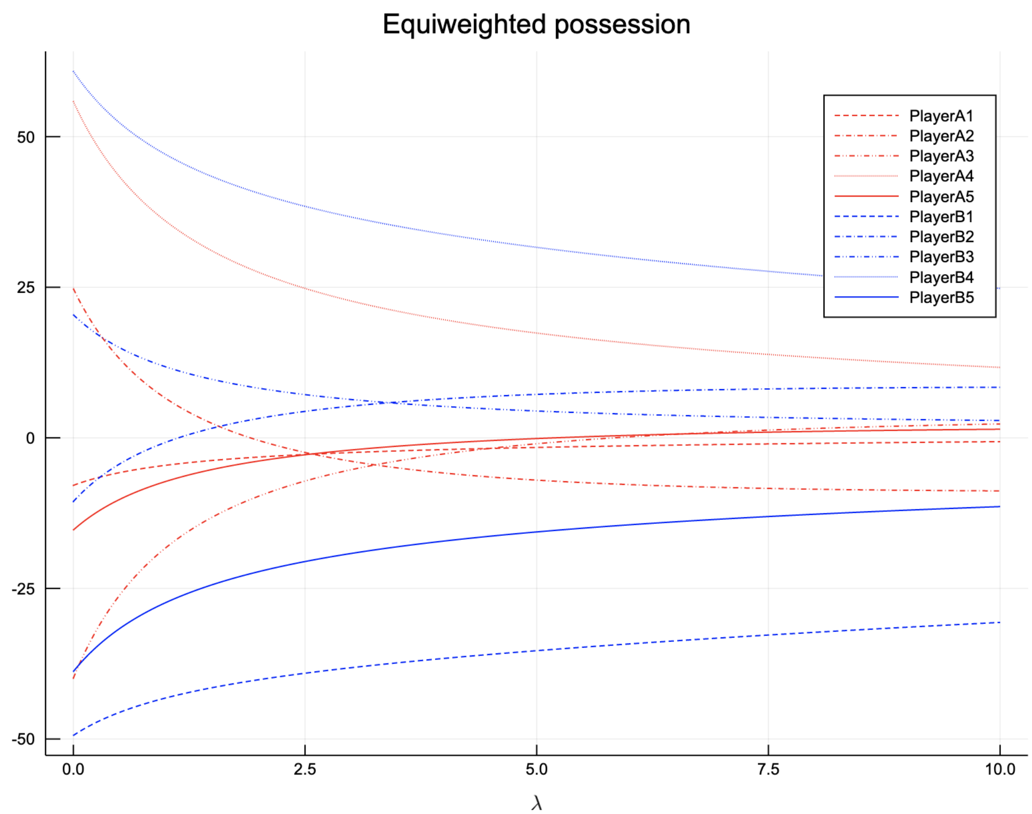


Figure 6: Regularized Adjusted Plus-Minus in function of λ