Game Theory in the NFL: Optimal Last Minute Football Strategy

Andrew deJong, Class of 2016

In American Football, outcomes of games are sometimes directly determined by the results of the final few plays. Offensive play calling is particularly crucial in these situations and it is rarely clear what the strategically optimal call is. The debate surrounding the controversial decisions by both coaches on the final few plays of the Super Bowl this year illustrates the lack of consensus with respect to play calling in critical "must score" moments. The goal of our ongoing project is to carefully characterize and provide a systematic analysis of optimal strategies at the end of close football games.

The scenarios analyzed so far include the down (1st, 2nd, 3rd, or 4th) and distance (from one to 15 yards) needed for a first down, the field position (from one to 30 yards away from scoring), and the time remaining in the game (from one to ten seconds). I have written a program that calculates the probability of winning for different play calls; run or pass, for each possible scenario within the constraints, while assuming that the call yielding the higher probability of scoring is optimal. Instead of using team specific archived statistics as data relevant for only specific teams, we used universal vectors with open parameters so that our code would be user friendly and flexible. If a coaching staff were to use this program, they could enter in the probability distributions of gained yards they felt were most accurate against their opponent and have optimal calls scripted for the game.

The game-theoretic tool of backward induction is primarily used for our analysis along with some probability concepts. Using Matlab as our coding platform and time as our state variable, we started by simply coding the base case where the offensive team has only one second left in the game to score. After analyzing the decision for the final play of the game and finding the optimal play selection (either run or pass), we continued to work backwards second by second. After repeating this process and continuing to work backwards, we constructed a matrix of dimensions 120 by 150 with every element populated with a probability of winning, taking into account gaining first downs to keep the drive alive. The code requires the user to input his run and pass distributions, along with the average amount of time needed for each run and pass play.

The distributions entered have a probability assigned to every possible "loss" ranging from no gain to losing 15 yards, as well as each "gain" from 1 yard up to 30. With that information, our model will automatically calculate the entire 120 by 150 matrix which gives the user the optimal strategy for every possible scenario within the bounds of our matrix, as well as the probability of scoring assigned to the optimal strategy. The (i,t) element is interpreted as the probability of scoring from the "i" yard line with "t" seconds left in the game. The down and distance are the last two dimensions which help categorize the exact scenario. Our calculations for running and passing probabilities are based on the following formula: probability(winning)=prob(win now)+prob(no win now)*prob(win later).

Based on our current assumptions, running ends the game so you would have to score on the first play. Passing the ball, if time permits, allows for future plays. The probability of winning later (subsequent plays) is a summation of each probability of all non-scoring results on the first play multiplied by its corresponding probability of winning from the new scenario (down, distance, field position, time) after the initial play. Our program selects the maximum among the run and pass probabilities and assigns the corresponding play as the optimal strategy. Therefore, hypothetically a coach using our program knows exactly whether to run or pass in order to optimize his team's chance of scoring in any possible end game scenario.

Below is a small portion of an example we've input into our program that provides an illustration of the output. With further extension of our program, we will get richer and more meaningful results.

						3rd down and 2					
						Time Remaining (seconds)					
		1	2	3	4	5	6	7	8	9	10
	1	0.75	0.75	0.75	0.75	0.817688	0.817688	0.817688	0.817688	0.817688	0.817688
	2	0.6	0.6	0.6	0.6	0.740775	0.740775	0.740775	0.740775	0.740775	0.740775
	3	0.48	0.48	0.48	0.48	0.668226	0.668226	0.668226	0.668226	0.848218	0.848218
Yard Line	4	0.4	0.4	0.4	0.4	0.614605	0.614605	0.614605	0.614605	0.809970	0.809970
	5	0.372	0.372	0.372	0.372	0.584256	0.584256	0.584256	0.584256	0.790519	0.790519
	6	0.346	0.346	0.346	0.346	0.554797	0.554797	0.554797	0.554797	0.771248	0.771248
	7	0.321	0.321	0.321	0.321	0.525743	0.525743	0.525743	0.525743	0.751339	0.751339
	8	0.297	0.297	0.297	0.297	0.497145	0.497145	0.497145	0.497145	0.730498	0.730498
	9	0.274	0.274	0.274	0.274	0.469056	0.469056	0.469056	0.469056	0.708694	0.708694
	10	0.252	0.252	0.252	0.252	0.441528	0.441528	0.441528	0.441528	0.686037	0.686037
	11	0.231	0.231	0.231	0.231	0.414612	0.414612	0.414612	0.414612	0.662649	0.662649
	12	0.211	0.211	0.211	0.211	0.38836	0.38836	0.38836	0.38836	0.638667	0.638667
	13	0.192	0.192	0.192	0.192	0.362819	0.362819	0.362819	0.362819	0.614195	0.614195
	14	0.174	0.174	0.174	0.174	0.338035	0.338035	0.338035	0.338035	0.589337	0.589337
	15	0.157	0.157	0.157	0.157	0.314053	0.314053	0.314053	0.314053	0.564195	0.564195

The image above shows the probability of the optimal play call from field position 1 up to 15 and time 1 up to 10 seconds remaining in the game. This example shows elements from the "3rd and two" matrix; the scenario when it is third down and two yards are needed to gain a first down. The two clear trends are the cell values (probability of scoring) decrease going down and increase going right. This is to be expected because having more time gives more opportunities to score, and starting farther away decreases the likelihood of reaching the endzone. The chart below has a corresponding value (1 or 2) for every element in the above chart, where "1" denotes a run play and "2" shows that passing is more favorable.

							3rd down and 2				
							Time Remaining (seconds)				
		1	2	3	4	5	6	7	8	9	10
	1	1	1	1	1	2	2	2	2	2	2
	2	1	1	1	1	2	2	2	2	2	2
	3	1	1	1	1	2	2	2	2	2	2
	4	1	1	1	1	2	2	2	2	2	2
Yard Line	5	2	2	2	2	2	2	2	2	2	2
	6	2	2	2	2	2	2	2	2	2	2
	7	2	2	2	2	2	2	2	2	2	2
	8	2	2	2	2	2	2	2	2	2	2
	9	2	2	2	2	2	2	2	2	2	2
	10	2	2	2	2	2	2	2	2	2	2
	11	2	2	2	2	2	2	2	2	2	2
	12	2	2	2	2	2	2	2	2	2	2
	13	2	2	2	2	2	2	2	2	2	2
	14	2	2	2	2	2	2	2	2	2	2
	15	2	2	2	2	2	2	2	2	2	2

Faculty Mentor: Daniel Stone

Funded by the Surdna Foundation Undergraduate Research Fellowship