NFL Overtime Markov Models

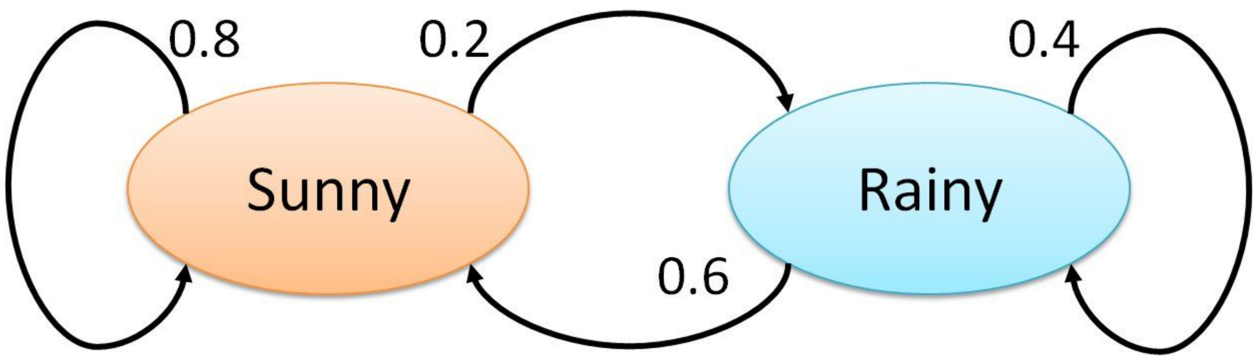
By Austin Grosel

DePaul University
CSC 529 - Advanced Data Mining
Case Study 3

Introduction to Markov Models

A Markov model is a recursive model that is composed of different sequential states and transitions that often are dependent on the prior state. These models are frequently repeated over a certain time interval to predict the chances of events occuring. We can recognize the different patterns of a certain situation and see which events are likely to happen based on a start state.

A classic example of a Markov model is predicting what the weather will be in x days based on what occured today. Our example is shown below, with the model's two weather states sunny and rainy (Primer).



In this example, a day can only be two states: sunny or rainy. If today is sunny, there'd be an 80% chance tomorrow of the day being sunny and a 20% of it being rainy. On the flipside, if today was rainy, there'd be a 40% chance tomorrow will be rainy and a 60% chance it will be sunny. This two state model can be put into a transition matrix (or table), which is displayed below. The rows indicate the weather today and the columns indicate the weather for tomorrow.

	Tomorrow Sunny	Tomorrow Rainy
Today Sunny	0.8	0.2
Today Rainy	0.6	0.4

The matrix above indicates the likelihood of what happens tomorrow based on today, or t = 1. But what if we wanted to find out the probability of the weather the day after tomorrow, or t = 2? For those results, we'd do a matrix multiplication of the transition matrix with itself and get the results below:

	DayAfterTomorrow Sunny	DayAfterTomorrow Rainy
Today Sunny	0.76	0.24
Today Rainy	0.72	0.28

Hypothetically, if today was sunny, there'd be a 24% chance that it'd be rainy in two days. The formula to figure out the weather probabilities n days later is [transition_matrix]^n. In this scenario, the probabilities of a future day being sunny or rainy will converge on 75% and 25%, respectively. This is because the transition probability from Rainy to Sunny is 3 times more than Sunny to Rainy (0.6 to 0.2). Ultimately, identifying the correct states and transition probabilities can allow a user to make accurate forecasts moving forward.

Business Case - NFL Overtimes

Throughout the NFL's history, there has been a discussion on how games should be decided when they are tied at the end of regulation. Because the NFL only has sixteen games in the regular season, the league felt that having ties could play a large factor in deciding which teams will make the postseason. On top of that, fans of either team usually don't want to see the game end without a winner.

The NFL created a sudden death period called overtime in 1974, where an extra timed period after regulation would determine the winner of the game. Since it was sudden death, the team that scored the first points would be victorious. These points can come from a safety (0.3% of all overtimes), touchdown (27% of the time), or a field goal (72.7% of the time). Field goals happen to be the easiest way to score points because a team can be relatively far away from the opposing team's endzone, and still kick it through the other team's field goal posts. It is why field goals are worth less points than touchdowns, 3 to 6, respectively. But since overtimes were sudden death, a field goal was worth just as much as a touchdown. Thus, there were a few problems with this method. In the NFL, the overtime period starts with a coin toss to determine which team receives the ball first. The team that wins the 50/50 coin toss already gets an advantage because if they score on their first possession, they win without the opposing team having a chance at scoring. The team that won the coin flip had a 54.8% chance at winning, and about 27% of overtime games ended with only one team possessing the ball (Sherman).

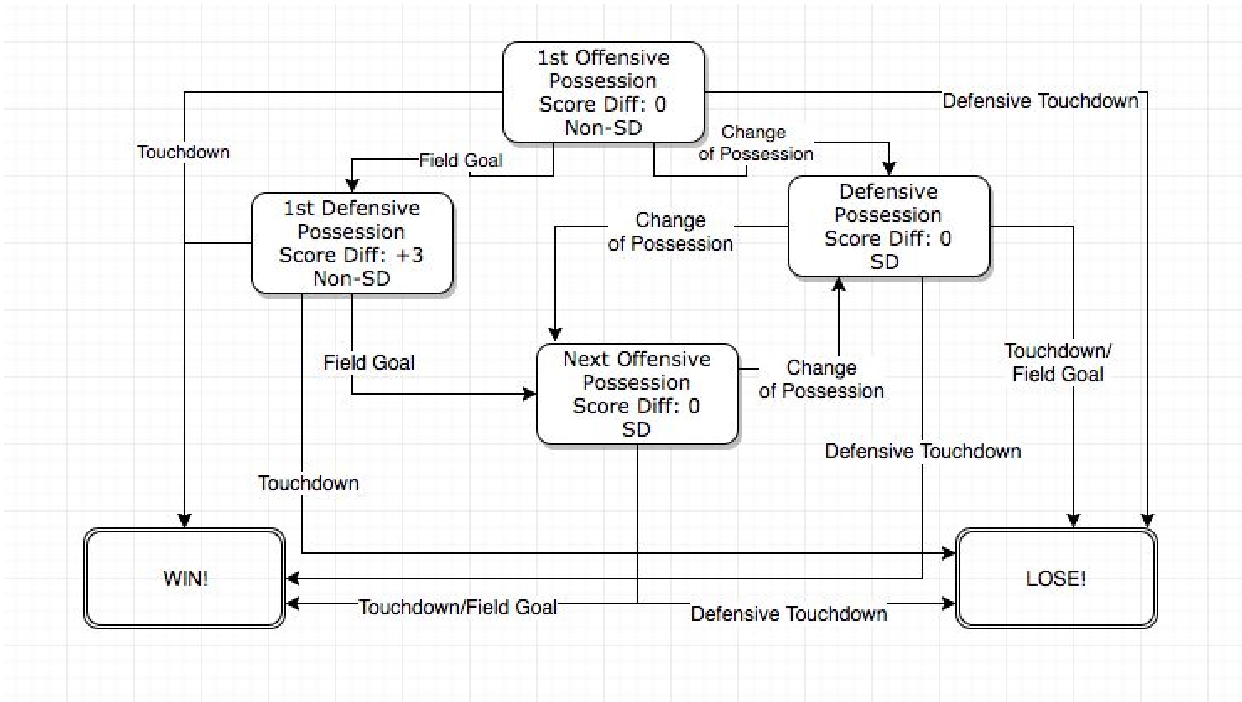
In 2012, the NFL decided that the team that does not receive the ball in overtime will be given a chance to score unless the receiving team scores a touchdown on their first possession. This eliminated the incentive that a field goal was worth the same value as a touchdown, and the receiving team will have to value trying to end the game immediately with a touchdown or giving the opposing team a chance with a field goal.

This created an interesting situation for decision makers on NFL teams. If a team had a stellar defense for example, is it smarter to allow the other team to receive the ball so you can likely force a turnover of possession and kick a field goal to win? Should a team that receives the ball first go for the winning touchdown or settle for an easier success rate event like a field goal and allow the other team a chance at tying or winning the game? Even with the rule changes, I hypothesize that it is still beneficial for the team that wins a coin toss to receive the ball first. Using a Markov Model, we can determine if this hypothesis is correct from looking at the results of overtimes in the past five years. The next section explains how one would set up a Markov chain for a problem like this.

Markov Model Application - NFL Overtimes

As I mentioned before, there have been numerous studies looking at what the best decision a team should do based on prior overtimes (Zauzmer, Jones). However, most of these studies came about a year or two after the overtime rules changed, thus they were working with small sample probabilities or using non-overtime probabilities. Thankfully, with almost five full years of data, I was able to get a better idea of the different probability distributions. In this study, we will assume that this overtime decision is for a playoff game, so there is no possibility of a tie occuring. Teams will play an infinite amount of possessions until one team is victorious.

I first had to identify all of the possible states that happen in overtime. These states would be the different types of possessions a team could possibly have. In this section, I will hypothetically use the team that receives the ball first as my own team. Therefore, at the end of the of Markov model, I can observe the final winning and losing probabilities from obtaining the first possession of overtime. The starting state of the model is our team receives the ball first, hence our first possession on offense. A possible state is a defensive possession with the score being tied, thus making the game sudden death. Another state is a defensive possession while we are leading because we just kicked a field goal. The other possible state would be another offensive possession with the game tied still, but this time we are n sudden death. The final states are a win or a loss. The diagram below shows the flow chart for these different states.



Going from the first offensive possession state, if we score a touchdown, we win. If we give up a defensive touchdown, we lose. If we kick a field goal, we play our first defensive possession winning by three points, giving the other team a chance to score. If we turn the ball over, we are now on defense and the game is officially sudden death (SD) where the next score wins. If we just kicked a field goal on our first possession, the opposing team must score or we win. If the opposing team scores a touchdown, we lose. If they kick a field goal, we come back on offense. This is a different state than the starting state because it is now sudden death rules. In sudden death, we get a different probability distribution because field goals and touchdowns now have the same value. From there, a cycle happens where the teams keep changing possessions unless a team scores.

Now that the states and actions are mapped out, we can put together the probability distributions. The tables below show the distributions of results given a certain state. The comment in the parentheses is the very next state after the result. All of these probabilities were found by querying the Drive Finder application on the Pro Football Reference website (Pro).

For the first overtime possessions (n = 91):

Result	**Probability**
Touchdown (WIN):	0.165
Field Goal (1st Defensive Poss, +3)	0.198
Defensive TD (LOSE):	0.02

Change of Possession (Defensive Poss, +0)	0.617

For the first defensive possession after a made field goal (n = 18):

Result	Probability
Touchdown (LOSE):	0.111
Field Goal (Offensive Possession, +0)	0.333
Change of Possession, Safety, Defensive TD (WIN):	0.556

For defensive possessions after a change of possession with the score tied (n = 125):

Result	Probability
Touchdown or Field Goal (LOSE):	0.451
Defensive TD or Safety (WIN):	0.028
Change of Possession (Offensive Poss, +0)	0.521

The above table will have the same probability distributions during an offensive possession as well, just flipped resulting states.

Now that we have our probabilities, we can create our transition matrix, which is displayed below.

	OFF_1P	DEF_1P_plus3	DEF_Poss_0	OFF_Poss_0	WIN	LOSE
OFF_1P	0	0.617	0.198	0.000	0.165	0.020
DEF_1P_plus3	0	0.000	0.000	0.521	0.028	0.451
DEF_Poss_0	0	0.000	0.000	0.333	0.556	0.111
OFF_Poss_0	0	0.521	0.000	0.000	0.451	0.028
WIN	0	0.000	0.000	0.000	1.000	0.000
LOSE	0	0.000	0.000	0.000	0.000	1.000

To figure out if receiving the ball first is the correct decision, we must look at the values of the WIN and LOSE column in the OFF_1P row. The OFF_1P row signifies the first possession of the game, so if WIN has a higher value than LOSE in that row, we can assume that receiving the ball first leads to a higher chance of winning. When we raise the matrix to the 100000th power, the following values appear:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
OFF_1P	0	0	0	0	0.5399275	0.4600725
DEF_1P_plus3	0	0	0	0	0.3609467	0.6390533
DEF_Poss_0	0	0	0	0	0.7688047	0.2311953
OFF_Poss_0	0	0	0	0	0.6390533	0.3609467
WIN	0	0	0	0	1.0000000	0.0000000
LOSE	0	0	0	0	0.0000000	1.0000000

The 5th column is WIN and the 6th column is LOSE. We see that there is about a 54% chance of winning when a team receives the ball first. This lowers the win probability of the receiving team from 54.8% under the previous rules, but not by much. Ultimately, there is still a big advantage to receiving the ball to start overtime.

Conclusion

While the NFL was able to decrease the advantage of winning the coin flip to start overtime, they probably aren't particularly satisfied with the results. Over an infinite amount of trials, we are able to show that teams win 54% of the time when they receive the ball first.

There are some issues that need to be addressed in this analysis, however. First, we do not adjust for team performance. It's possible that in our 91 game sample, the better teams have received the ball first more than worse teams. Another issue is that there are still not enough possessions to account for after the receiving team obtains the ball again. In our dataset, there are only 6 drives that were like this, and none of the drives resulted in scores. One of the likely scenarios is that these drives ended when teams hit the time limit, resulting in a tie. Because our situation accounts for teams to play infinite possessions until there's a winner, these possessions that end in regulation are pointless for this study. Thus, we had to throw out this entire situation and assume the probabilities were the same as a sudden death condition. I'd imagine we'd start getting more accurate data in these situations as the sample size grows. Overall, I'm satisfied with these results and surprised that the news rules don't have much of an effect in decreasing the receiving team's chances of winning.

Works Cited

"Primer on Hidden Markov Models, Part 1". *sgmustadio,* https://sgmustadio.wordpress.com/2011/02/17/primer-on-hidden-markov-models-part-1/

Sherman, Rodger. "The NFL's Overtime Rules Aren't Fair-But Neither are the Alternatives." *The Ringer,* https://www.theringer.com/2017/2/6/16042116/nfl-overtime-rules-super-bowl-li-patriots-falcons-62316a6f8e3c

Zauzmer, Ben. "Modeling NFL Overtimes as a Markov Chain." *Harvard Sports Analytics Collective*, http://harvardsportsanalysis.org/2014/01/modeling-nfl-overtime-as-a-markov-chain/

Jones, Chris. "A look at Overtime in the NFL." *Math Aware*, http://www.mathaware.org/mam/2010/essays/JonesOvertime.pdf

Pro Football Reference. "Drive Finder" https://www.pro-football-reference.com/play-index/drive_finder.cgi