Question 2a Recall the optimal value of θ should minimize our loss function. One way we've approached solving for θ is by taking the derivative of our loss function with respect to θ, like we did in HW5.

Write/derive the expressions for following values and write them with LaTeX in the space below.

- $R(\mathbf{x}, \mathbf{y}, \theta_1, \theta_2)$: our loss function, the empirical risk/mean squared error
- $\frac{\partial R}{\partial \theta_1}$: the partial derivative of R with respect to θ_1
- $\frac{\partial R}{\partial \theta_2}$: the partial derivative of R with respect to θ_2

Recall that $R(\mathbf{x}, \mathbf{y}, \theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2$

(, , 1, 2) = 1 (− (1 + (2)))2

/ 1 =1 =−2 (−(1 + (2)))

/ 2 =1 = − 2 (− (1 + (2))) ((2))

In 1-2 sentences, describe what you notice about the path that theta takes with a static learning rate vs. a decaying learning rate. In your answer, refer to either pair of plots above (the 3d plot or the contour plot).

The path that theta takes with a static learning rate stops at more points and seems to consider more points than the decaying learning rate's path does. The decaying learning rate's path considers points more towards the beginning of the path. They do share similar points such as the (0,0) point.

0.0.1 Question 4b

Is this model reasonable? Why or why not?

No, it seems to always have a greater than 50% chance of winning.

0.0.2 Question 4c

Try playing around with other theta values. You should observe that the models are all pretty bad, no matter what θ you pick. Explain why below.

No matter what value of theta is used, the win probability is always at least 0.5.

0.0.3 Question 5b

Using the plot above, try adjusting θ_2 (only). Describe how changing θ_2 affects the prediction curve. Provide your description in the cell below.

Increasing 2 shifts the curve to the left and upward.

0.0.4 Question 7c

Look at the coefficients in `theta_19_hat` and identify which of the parameters have the biggest effect on the prediction. For this, you might find `useful_numeric_fields.columns` useful. Which attributes have the biggest positive effect on a team's success? The biggest negative effects? Do the results surprise you?

In [ ]:
```
theta_19_hat
```

In [ ]:
```
useful_numeric_fields.columns
```

It appears 'FGM', 'FG3_PCT', 'FT_PCT', and 'PTS' have the biggeest positive effect on a teams success. This makes sense, as the first three have to do with how well a team is shooting, and the last is how many points they score. The better a team shoots, the more points they have and the more points they have means they are more likely to outscore their opponents and win the game. The biggest negative effect is 'FG_PCT', which again measures how well a team is shooting, so the worse a team shoots the more likely they are to lose. These results are surprisingly because shooting metrics are both the biggest postive and negative effects on a teams success.

To double-check your work, the cell below will rerun all of the autograder tests.

In [ ]:
```
grader.check_all()
```

0.1 Submission

Make sure you have run all cells in your notebook in order before running the cell below, so that all images/graphs appear in the output. The cell below will generate a zip file for you to submit. **Please save before exporting!**

In [ ]:
```
# Save your notebook first, then run this cell to export your submission.
grader.export("hw7.ipynb")
```