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.

In the space below, use LaTeX to write/compute the following values: * $R(\mathbf{x}, \mathbf{y}, \theta_1, \theta_2)$: 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 $R(\mathbf{x}, \mathbf{y}, \theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2$

$R(x, Y, \theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^{n} (y_i - (\theta_1 x_i + \sin(\theta_2 x_i)))^2$

$\frac{\partial R}{\partial \theta_1} = \frac{-2}{n} \sum_{i=1}^{n} x_i (y_i - (\theta_1 x_i + \sin(\theta_2 x_i)))$

$\frac{\partial R}{\partial \theta_2} = \frac{-2}{n} \sum_{i=1}^{n} x_i \cos(\theta_2 x_i)(y_i - (\theta_1 x_i + \sin(\theta_2 x_i)))$

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 main difference between the staic learning rate and decaying learning rate as seen in both the 3d and contour plot is that in the static learning rate case, the distance of each step is much larger and theta estimate jumps back and forth from being above and below the optimal value until it converges to the optimal value. In the decaying learning rate case, the process takes one big jump and then smooths out, taking smaller steps as the number of iterations increase until it converges to the optimium.

0.0.1 Question 4b

Is this model reasonable? Why or why not?

No. This model predicts that every team in the data set will win. Looking at the plot above, the model predicts a win at all scores even though half of the datapoint represent losses. This model has bad accuracy.

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.

Since our model does not have an intercept term, we can shift the sigmoid (which is centered at zero) towards the center of our data which approximately 100. For this reason, changing theta on changes the steepness of the curve, not its center. This means that assuming all scores > 0 points, the model will predict all of those scores as winning games if $\theta > 0$ *or all losing if* $\theta < 0$.

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.

Changing θ_2 shifts the entire curve left and right. For instance changing to -6 shifts the curve right, while changing it to -4 shifts the curve left.

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?

The attributes that had the biggest positive effect on a teams success were FGM, FG3_PCT, FT_PCT, and BIAS. The attribute with largest negative effect was FG_PCT. The results for the positive effect categories seem reasonable, but the fact tha field goal percentage (FG_PCT) has a negative effect is surprising.

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

In [168]:
```
grader.check_all()
```

Out[168]:
```
q1:

    All tests passed!


q2b:


    0 of 1 tests passed


    Tests failed:

        ./tests/q2b.py

Test result:
Trying:
    abs(sin_MSE([0, np.pi]) - 19.49000412080223) <= 1e-5
Expecting:
    True
**********************************************************************
Line 1, in ./tests/q2b.py 0
Failed example:
    abs(sin_MSE([0, np.pi]) - 19.49000412080223) <= 1e-5
Expected:
    True
Got:
    False
```

```
q3a:

    All tests passed!

q3b:

    All tests passed!

q4a:

    All tests passed!

q5a:

    All tests passed!

q5c:

    All tests passed!

q6a:

    All tests passed!

q6b:

    All tests passed!

q6c:

    All tests passed!

q6d:

    All tests passed!

q7a:

    All tests passed!

q7b:
```

```
All tests passed!
```

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 zipfile for you to submit. **Please save before exporting!**

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