

Write your answers below. You do not need to share your code for this workshop.

1. In class we guessed that a simple logistic regression model for predicting gender based on height (in inches) might have the form:

$$\log\text{-odds}_{\text{male}} = 0.5(\text{height}) - 33.$$

Now use gradient descent to find the best possible coefficients for a logistic regression model with the following data. In your model, you can count male as +1 and female as -1.

Height	Gender
72	M
69	M
64	F

You will need to use this formula for the gradient of each term of the logistic loss function:

$$\nabla L_i(\mathbf{w}) = \begin{cases} -(1 - p_i)X_i & \text{if } y_i \text{ is male,} \\ p_i X_i & \text{if } y_i \text{ is female} \end{cases}$$

where the probability $p_i = \frac{e^{X_i \cdot \mathbf{w}}}{e^{X_i \cdot \mathbf{w}} + 1}$.

2. What does the logistic regression model say about the gender of someone who is 67 inches tall? What are the odds that they are male? What is the probability that they are male?
3. What is the total logistic loss for your parameters \mathbf{w} ? Recall that the logistic loss function is

$$L_i(\mathbf{w}) = \begin{cases} -\log(p_i) & \text{if } y_i = 1 \\ -\log(1 - p_i) & \text{otherwise.} \end{cases}$$

4. Now use gradient descent to find the best coefficients for predicting gender based on height using hinge loss instead of logistic loss. The formula for the gradient of the hinge loss function is:

$$\nabla L_i(\mathbf{w}) = \begin{cases} -y_i X_i & \text{if } y_i(X_i \cdot \mathbf{w}) < 1 \\ 0 & \text{otherwise.} \end{cases}$$

What are the coefficients \mathbf{w} for a linear classifier with the lowest hinge loss? Can you get the hinge loss all the way down to zero? Recall that the terms of the hinge loss are

$$L_i(\mathbf{w}) = \begin{cases} 1 - y_i(X_i \cdot \mathbf{w}) & \text{if } y_i(X_i \cdot \mathbf{w}) < 1 \\ 0 & \text{otherwise.} \end{cases}$$