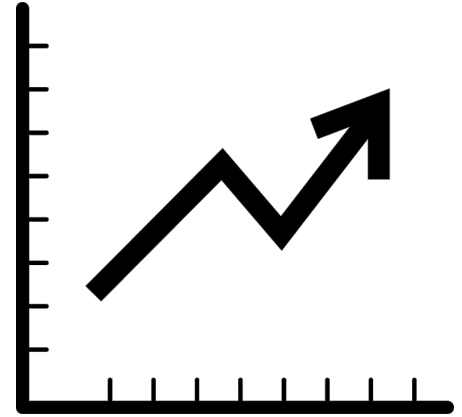


Introduction to Data Analysis



How do we analyze data?



- Two technologies you are allowed to use:
 - **Python**
 - Open source
 - Easier to use for CS majors
 - MATLAB
 - Great for hefty math operations
 - Not the most friendly for those familiar with CS

Why we are using Python

- Plenty of libraries in Python support brain data analysis and machine learning (Numpy, Scipy, etc.)
- Python notebooks make it easy to share code and graphs with fellow researchers (important when collaborating)



Intro to linear regression

Linear regression formulas

$$y = ax + b$$

$$a = \frac{\sum_{i=1}^n ((x_i - \text{mean}(x)) * (y_i - \text{mean}(y)))}{\sum_{i=1}^n ((x_i - \text{mean}(x))^2)}$$

$$b = \text{mean}(y) - a * \text{mean}(x)$$

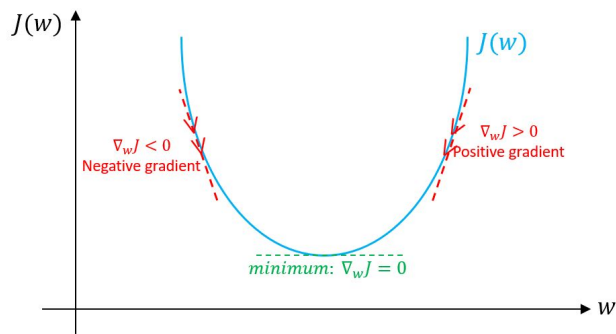
Quick in-class exercise

Find the absolute minima of the following function:

$$f(x) = x^2 - 5$$

Gradient Descent

- When testing a model on a dataset, we derive an **error function** for the model
- To improve our model, we try to find the minima of the error function
- The **gradient** gives us the vector with which to adjust our initial parameters to approach the minima



Applying gradient descent to Linear Regression

Logistic Regression

Recap

- Python has libraries that aid in data analysis
-

Questions?