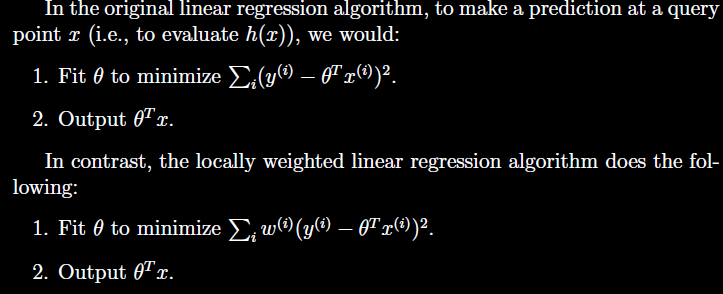
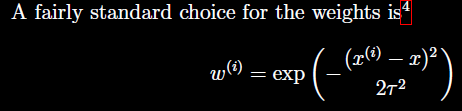
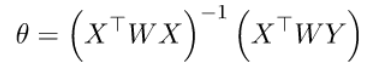
* Sources:
  + CS229 notes (17-19)
  + <https://towardsdatascience.com/locally-weighted-linear-regression-in-python-3d324108efbf>
  + <https://aman.ai/cs229/locally-weighted-linear-regression/>
* Used when data is non-linear, since it can fit non-linear data
* Assuming there is sufficient training data, makes the choice of features less critical







* w (i) are not random variables
* tau (bandwidth parameter) controls how quickly the weight the weight of a training example falls off with distance of its x (i) from the query point x
* Non-parametric learning algorithm - because we need to keep the entire training set to make predictions. Non-parametric refers to the fact that the amount of stuff we need to keep in order to represent the hypothesis (h) grows linearly with the size of the training set.
  + Linear regression is a parametric learning algorithm because it has a fixed, finite number of parameters (thetas), which are fit to the data. Once we have the thetas, we no longer need the training data to make predictions.
* **Assumptions**
  + w (i) are non-negative valued weights
* **Advantages**
  + Don’t need to think about which features to use
  + Useful when number of data points is small
  + Useful when number of dimensions are less
* **Disadvantages**
  + Not ideal if dataset is massive since high memory and computation cost
  + Can get affected by outliers (applies to all least square methods in ML)
  + In high dimensions, possible that not many points are near a particular query