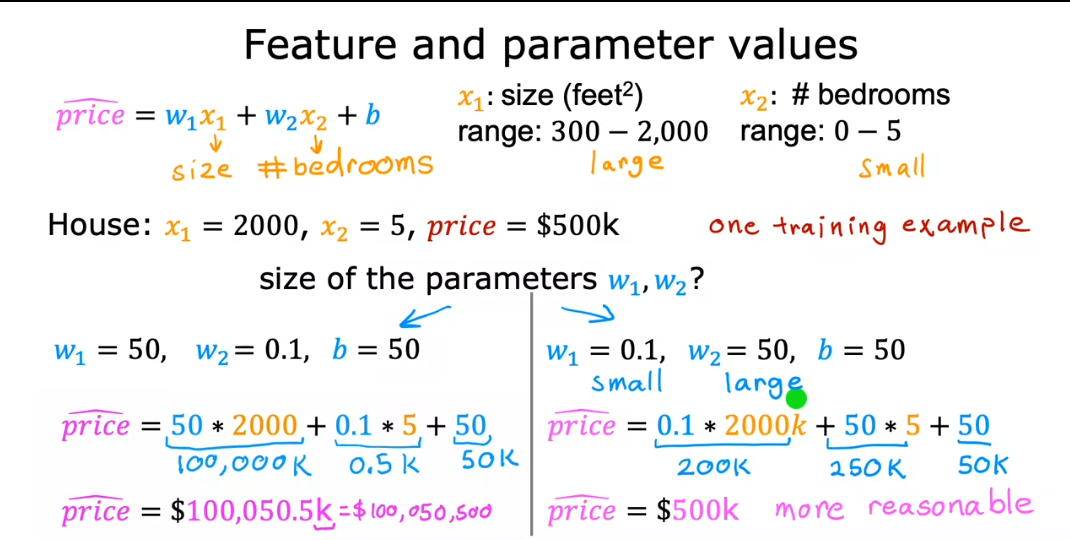
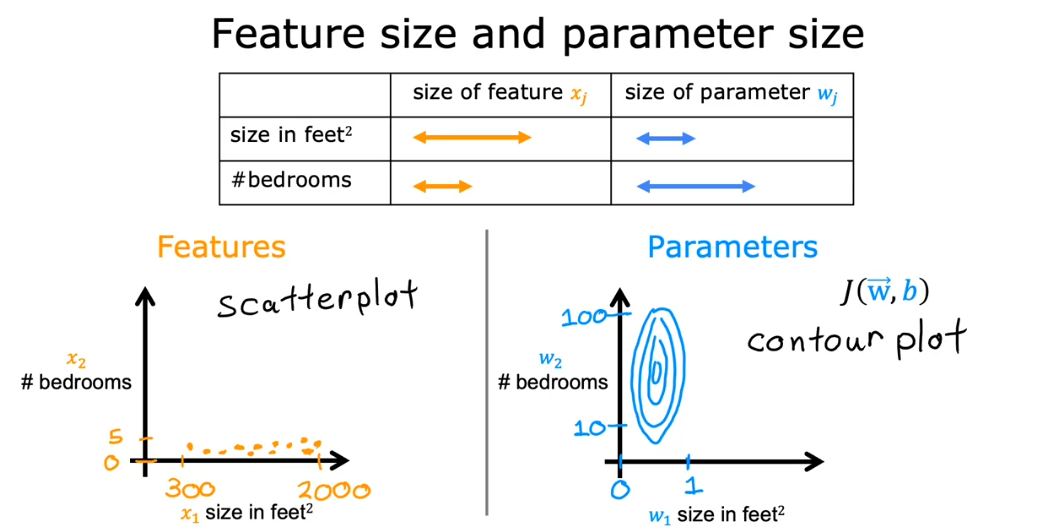
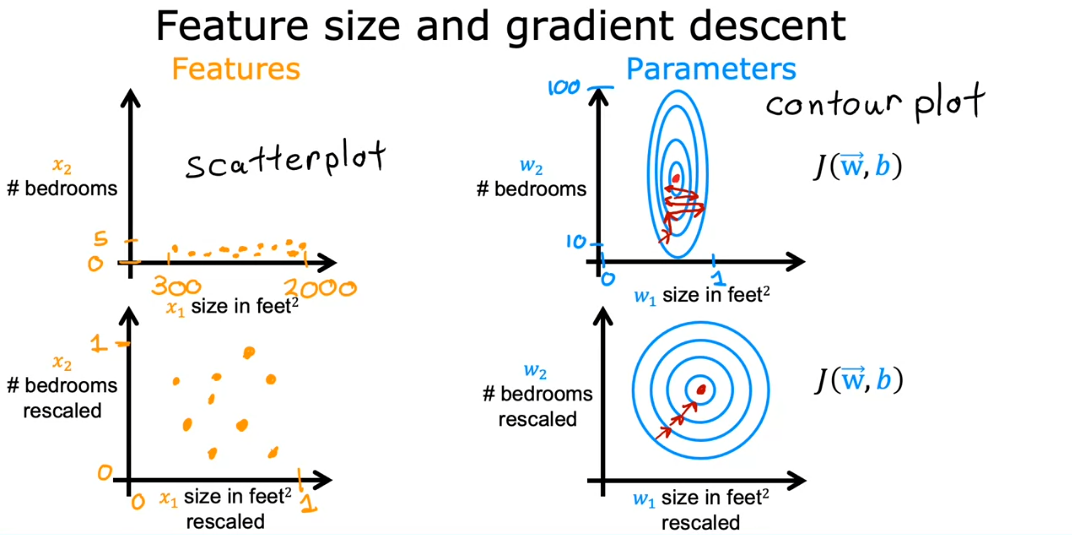
# Feature scaling part 1

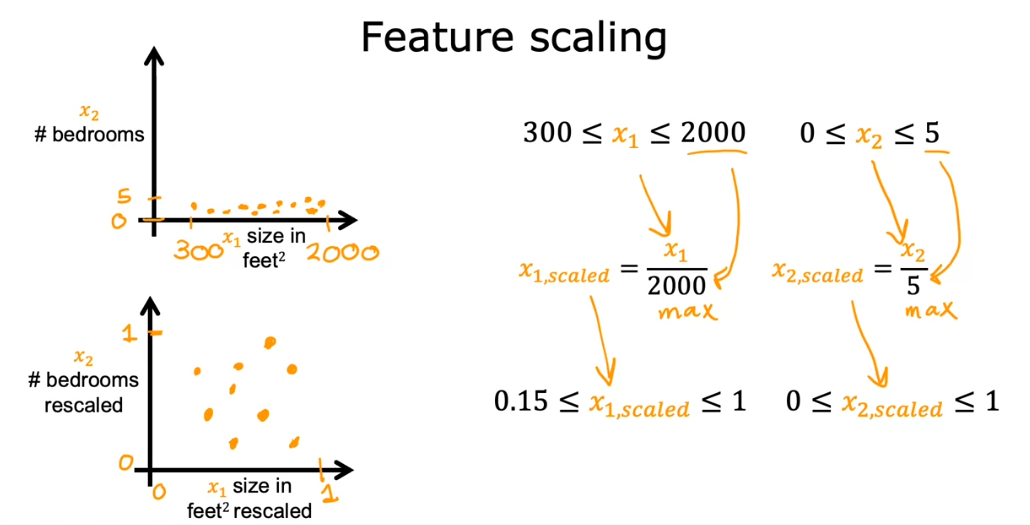




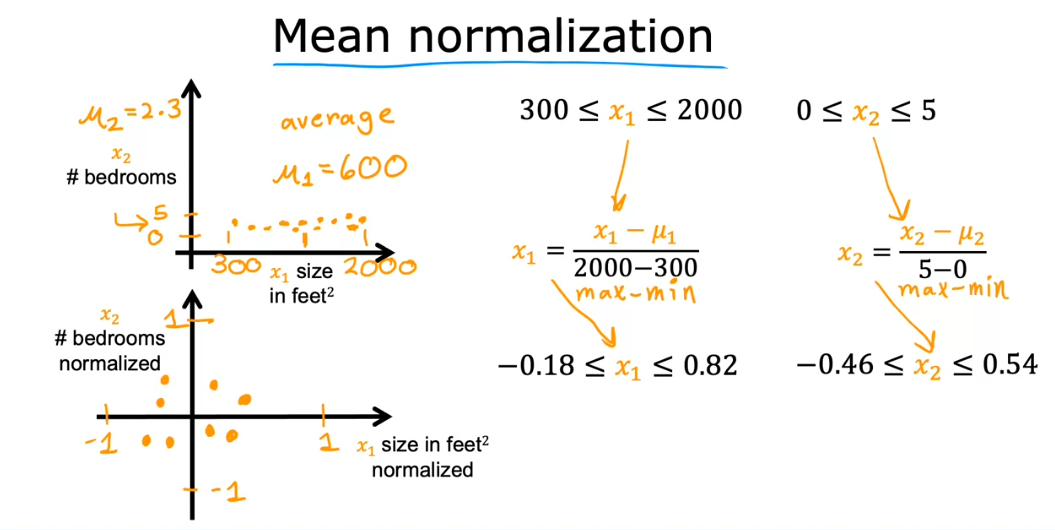


# Feature scaling part 2

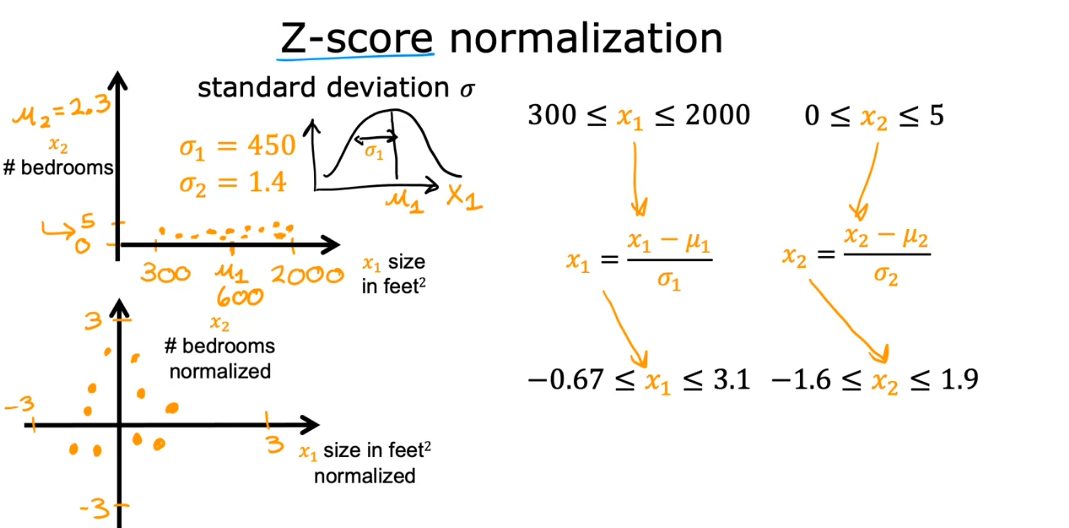
## Dividing by the Maximum

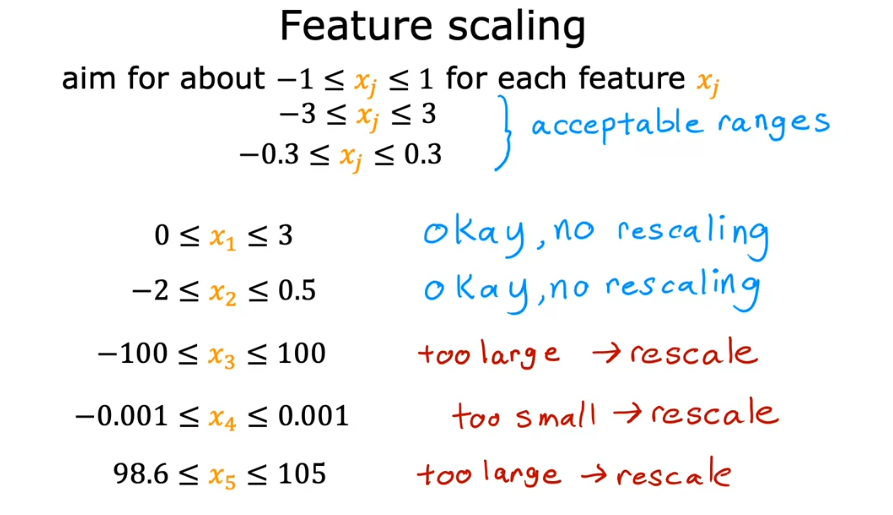


## Mean normalization



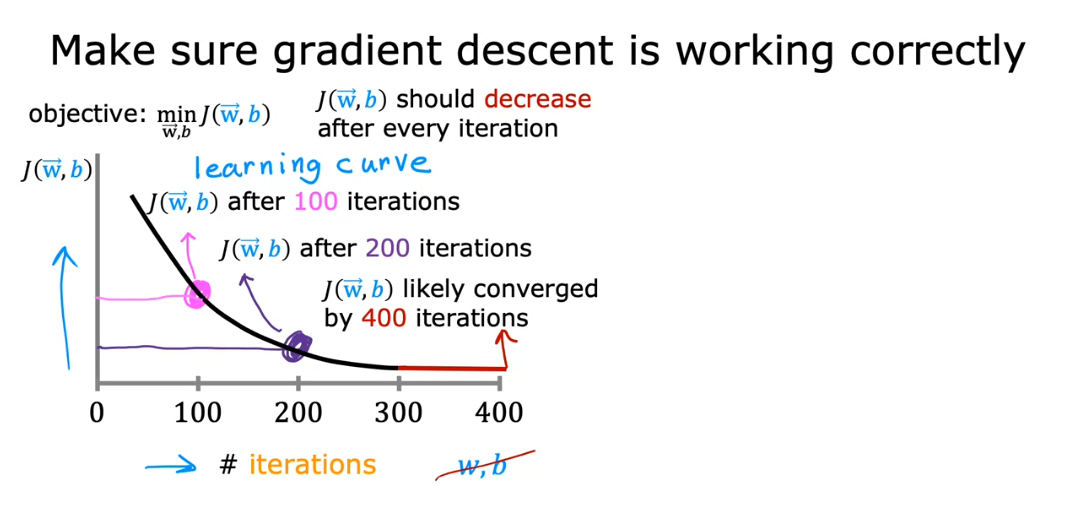
## Z-score normalization



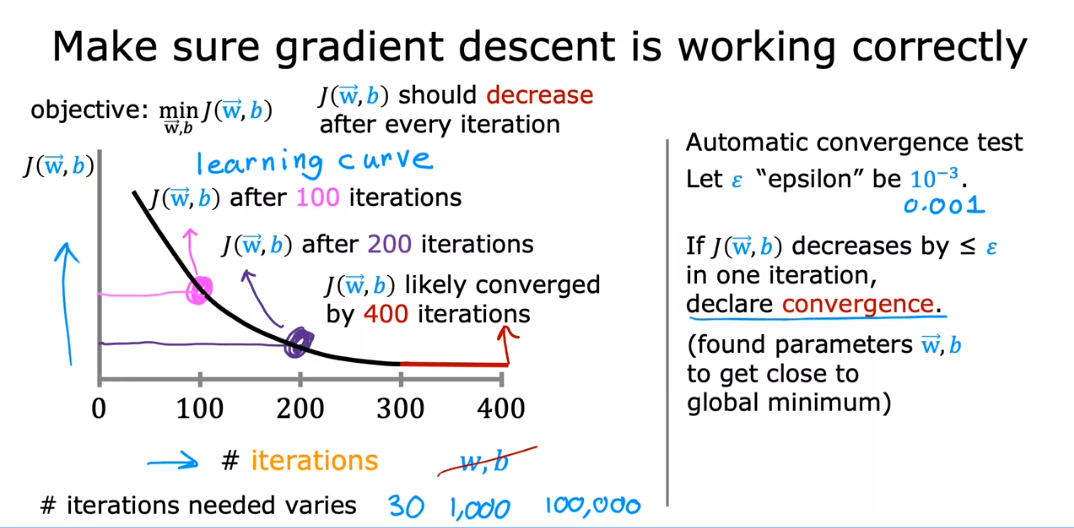


* When in doubt, just rescale! No harm in any way

# Checking gradient descent for convergence

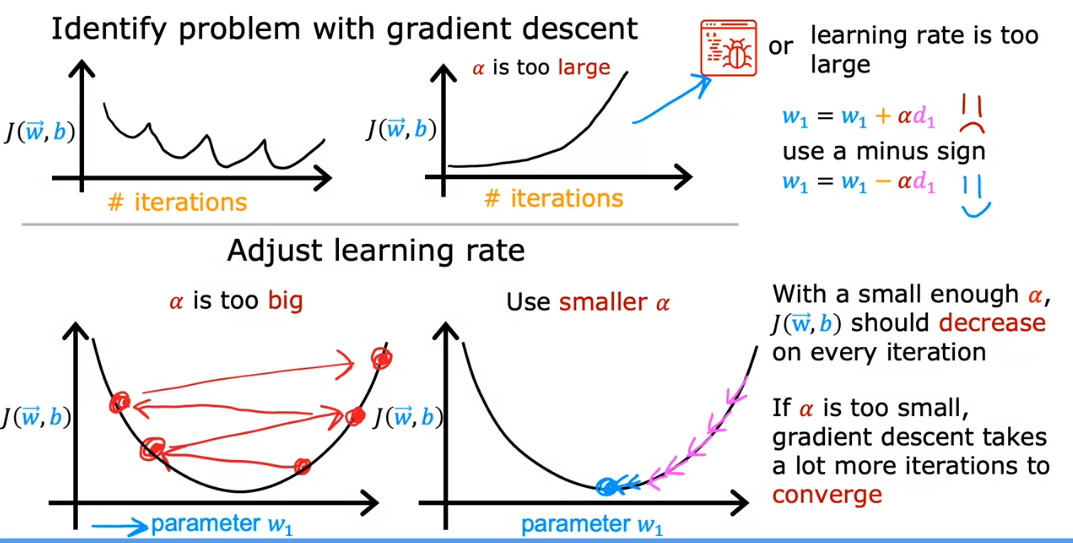


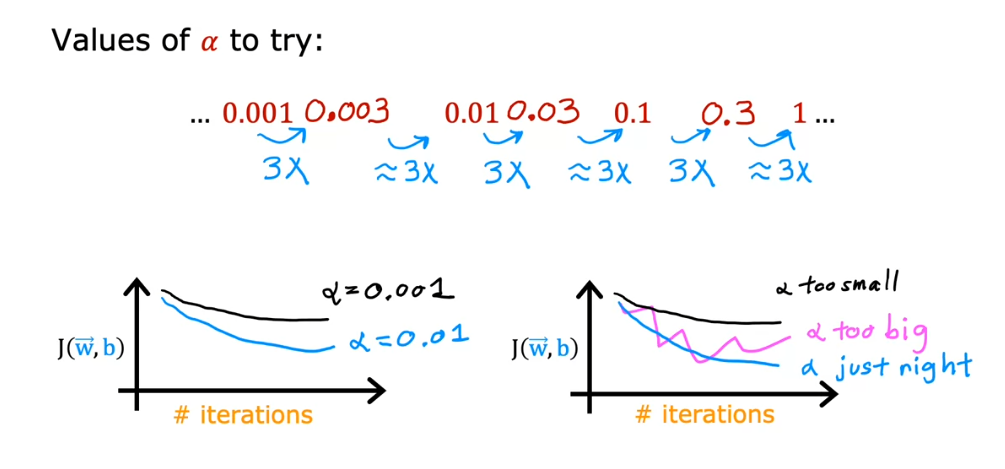
* Even only in one increase of cost J:
  + A bug in the code
  + Poorly chosen alpha
  + Gradient descent not working correctly already
* Must always decrease no matter what!



* The graph is more reliable than the automatic convergence tests

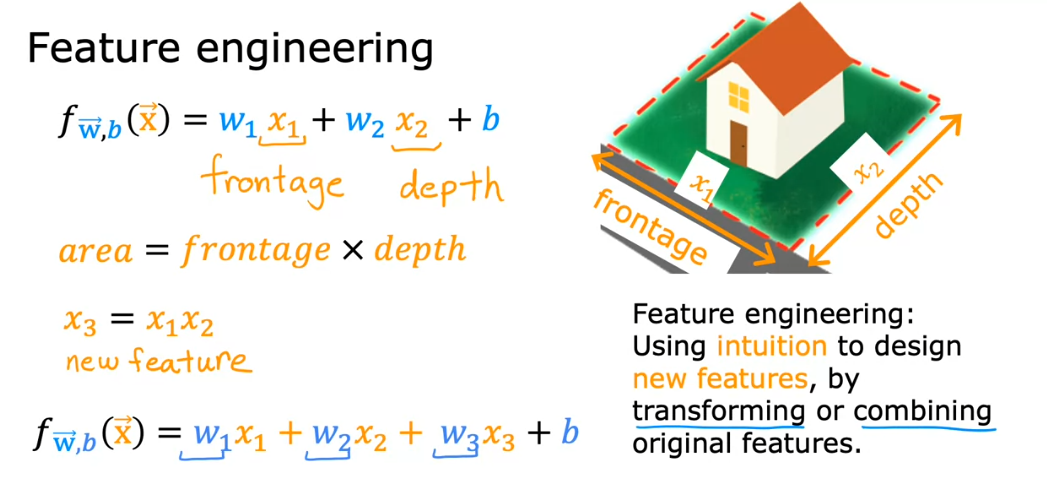
# Choosing the Learning Rate





* Technique:
  + Find too small,
  + Then find too big
  + Then go from that range

# Feature Engineering



# Polynomial Regression

