Method of analysis:

1. Gradient descent: Multiply parameters with features, use gradient descent to find local minimum after loops
2. Scaling parameters: (Parameters between 0 and 1 could optimize computation)

* Using mean normalization: , where is mean of x in training set, and is the range

1. Ensure the gradient descent is working correctly: Plot the cost function parameter vs the number of iterations. If we have a decreasing plot, gradient descent is working correctly. If not, check:

* Use smaller
* Change the hypothesis function to change the behavior of the curve