学习过程：

过程中遇到的问题和解决方法

最开始拿到训练集的时候，我的选择是将含nan的样本全都删除，模型的最初构建时，我先用的一个主体函数，将损失函数作为迭代停止的条件。但是因为刚拿到题目时，对梯度下降的算法还不是很熟悉，而且线代也没有很好的掌握，所以对于矩阵相乘的维度很难把握，所以写一个整体函数程序总是报错。

所以我选择重新写将损失函数（cost），偏导函数（gradient），梯度函数（descent返回theta值），分开写成不同的函数，在不同的函数中相互调用（在descent函数中调用gradient 函数），通过这样的结构，在训练模型的最后就只选择了调用decent函数，实现对theta的求取。

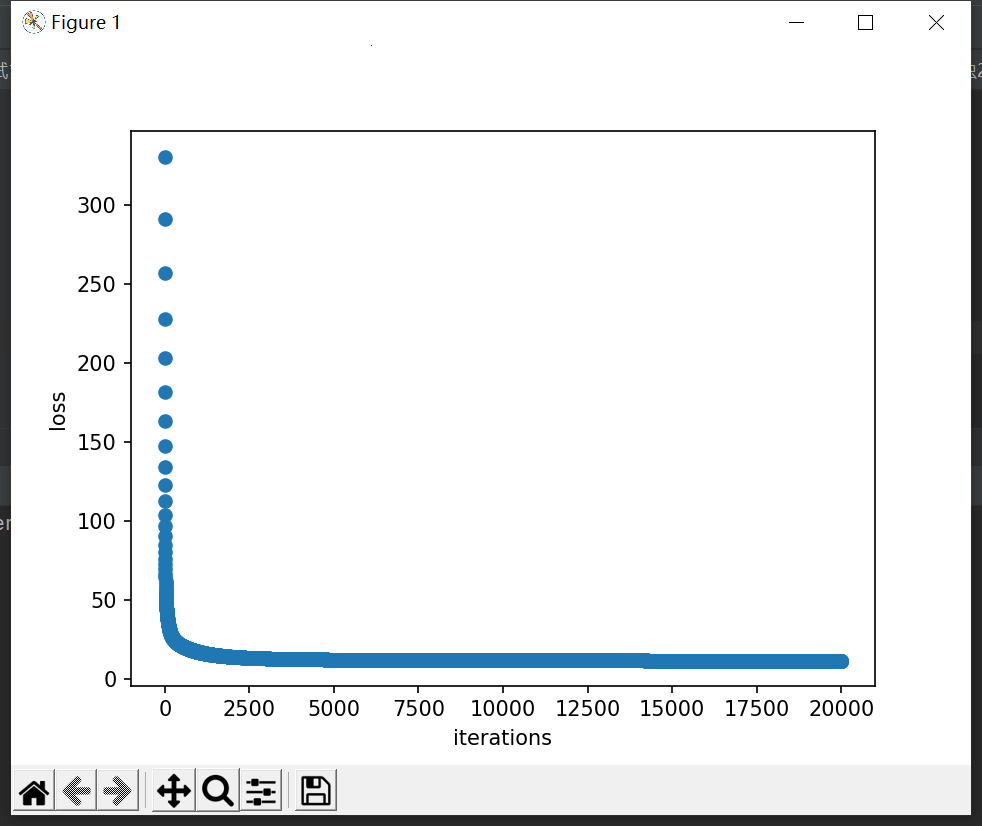
在测试集中也是将theta与X样本特征进行矩阵相乘，得到相应的Y（预测值）。

初步做出来结构后发现，loss值太大。我开始观察训练集，发现不同特征中数值差别很大，所以采用了归一法，将训练集和测试集的特征样本都进行了归一化处理，其实刚用完归一时loss值还是很高，才发现我将训练集的全部数据都进行了归一化处理，但是应该只是将特征x处理就好。 当不同的特征成列在一起时，有些特征的值过大，与其他特征相比对整体函数的影响太大，就将全部数据进行归一处理，保证每个数据都会被平等对待。

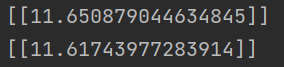
但是经过归一化的处理之后，loss值还是比较高达到了14.6743。我又尝试了3西格玛原则，但是发现这一原则只适用于样本数据符合正态分布。又尝试了四分位距法，同一特征内数值差距太大，异常值太多，对参数影响较大，需要去除异常值，我将高于75%和低于25%都用该特征的均值代替，但是发现结果loss变得更大了，达到了[[41.64436672323648]]，接着尝试缩小75%和25%的范围，发现loss值逐渐变小，又向着[[14.]]靠近，所以四分位距法的使用反而让模型的拟合更差了。

在循环条件，最开始尝试的是用loss低于某个值来作为循环结束的条件，但是这样的模型运行后就不会停止，得不到结果。后面又尝试自己规定循环次数和将偏导约=0作为循环结束的条件，这两者结果差不多。

然后又将之前删掉的涵nan的样本变成了nan用该特征的均值代替，因为nan值删掉就删了大概1/3的样本，误差太大。更改之后，loss降到了[[11.61743977283914]]。再然后我就没有发现更好的方法优化了



、

 这是loss值随迭代次数的图像，不管是改变迭代次数，还是将梯度<1e-5作为循环结束的条件都无法是loss值再降低。同时也使用了正则化直接算出theta，与梯度下降得到的相比较。比较Theta和theta2，因为相信回归有全局最优值，全局最优的参数Thetha可以直接算，但是两种方法得到的theta值差不多

所以现在遇到的问题就是已经收敛了，损失值降不下去，还没有找到解决方法。