根据我们所需要处理的数据类型，我们确定了三种主要的算法进行数据处理。

1. 局部加权回归算法

局部加权回归算法是对线性回归的扩展，当目标假设是线性模型时，使用线性回归自然能拟合的很好，但如果目标假设不是线性模型，比如一个忽上忽下的的函数，这时用线性模型就拟合的很差。

因为我们处理的ofo轨迹路线数据显然不是简单的线性数据，仅仅采用线性回归会造成很大的误差。为了解决这个问题，当我们在预测一个点的值时，我们选择和这个点相近的点而不是全部的点做线性回归。

目标函数: 寻找合适的 θ使得J(θ)最小

当

2. 逻辑回归算法

回归就是一个线性分类模型，它与线性回归的不同点在于：为了将线性回归输出的很大范围的数，例如从负无穷到正无穷，压缩到0和1之间，这样的输出值表达为“可能性”才能说服广大民众。当然了，把大值压缩到这个范围还有个很好的好处，就是可以消除特别冒尖的变量的影响（不知道理解的是否正确）。通过在输出加一个objective函数，就可以将原本的线性回归标准化。在我们的ofo数据集中，这种方法也提供给我们更新的思路。

选择一个合适的 去做objective函数:

根据选择的, 我们可以找到每个数据所对应的classification值:

将 代入, 我们会找到新的θ表达式:

3. 牛顿方法

是一种可以快速的得出对应解的办法，对算法的效率有数量级上的提高。

迭代过程中的更新规则:

## 遇到的问题：

当我们运用这三个办法去训练数据集时，出现了以下问题：

1、局部加权回归算法比较适合数据量较小的数据集。所以当我们采集的数据量变大时，这个算法的效率下降的非常快。

2、逻辑回归算法中选择一个合适的是非常关键的，所以面对不同的数据集我们都要重新寻找合适的。这个算法在前期准备工作中是比较耗时耗力的。

3、牛顿法的速度是目前为止最能够达到我们预期期望的。但是使用这个算法有很多条件的限制，不是每一个数据集都可以满足他的条件。