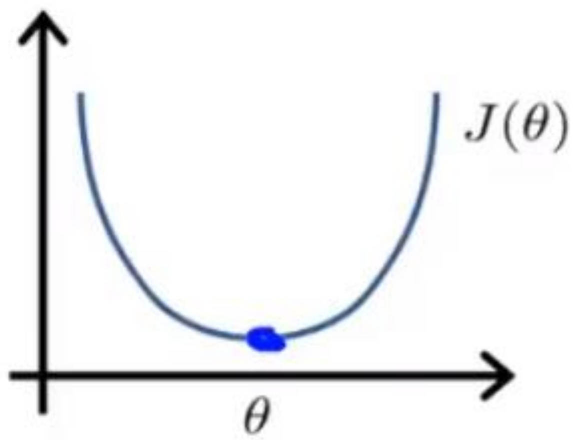
1、回顾

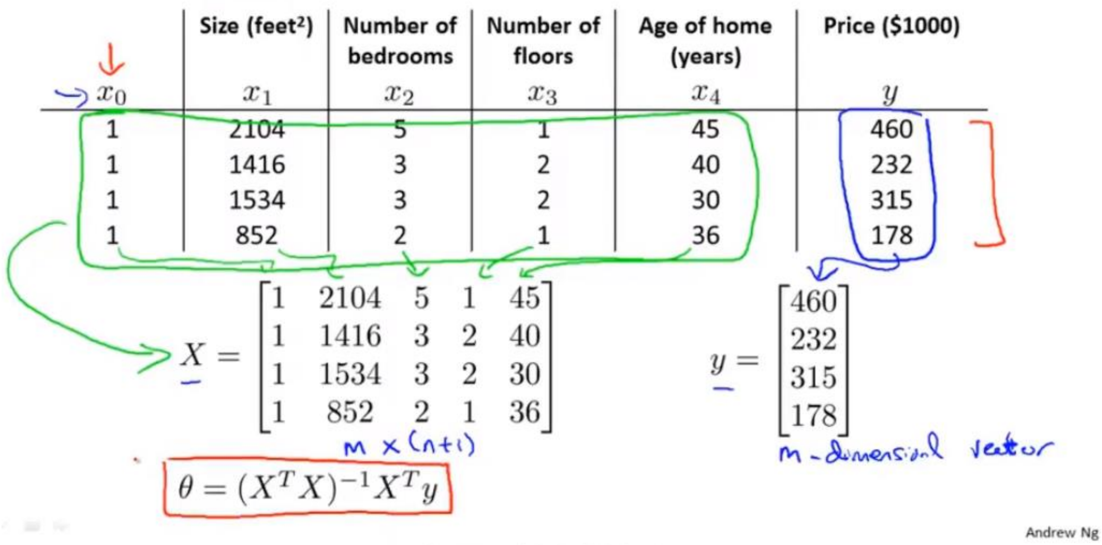
之前在利用梯度下降算法中计算θ值的时候，每次迭代更新时，都要计算该点的偏导数值，更新的速度和次数取决于α学习率和迭代次数，大致思路就是通过每次计算当



前导数值，更新参数值，使得下次代价差趋向于变小，这是之前所说的梯度下降法，既然我们都已经知道代价函数随θ变化规律类似于上图这样，能不能快速找到代价极小值呢？由此引入了正规方程算法。

2、正规方程算法原理

要找到一个极小值θ，令 J 随θ的导数等于零，这样的一个θ值就会使得 J 值为局部最小值，同理，多元参数变量情况类似。需要求得各个偏导数，使其等于零。以下这张图很好的说明了求任意个数参数的方法，只需找到 X 和 y 矩阵即可。



3、比较

梯度下降法	正规方程法
选择α学习率	无α
选择迭代次数	一次计算
需要特征缩放	无需缩放
特征数量很多时，无较大影响	N 个特征，构成 N 阶矩阵，求逆矩阵运算量巨大
适用范围更广	仅适用于特征<1w 的线性回归

6、暗区

你可能会想到如果矩阵不可逆怎么办，首先考虑消除多余特征，其次，pinv 称为伪逆矩阵算法，在正规方程算法中，几乎都可以算出来θ，所以，别过多担心。

5、实战

使用两种方式都可以快速计算得到相应的θ，但仍然有差异。