1、单变量线性回归

**整体思路：**

（1）假设模型为

（2）对学习速率 、迭代次数iteration 进行取值，一般取

（3）记得给训矩阵X的左边添加一列1的元素，即：X=[ ones(m,1) X ]

**当特征数多时：（>10000）**

（单变量线性回归中，特征数为1个，建议使用正规方程法求解 ）

法一：梯度下降法，计算出 的值，（在该过程中，也可顺带计算出 ，绘制的图像，用于观察曲线是否单调递减，用于判断代码是否有错）

**当特征数少时：（<10000）**

法二：正规方程法， ，无需使用迭代、学习速率，直接计算出 的值。

（4）最后，直接将算出的，回代，得到测试数据的期望值。

**详细内容：**

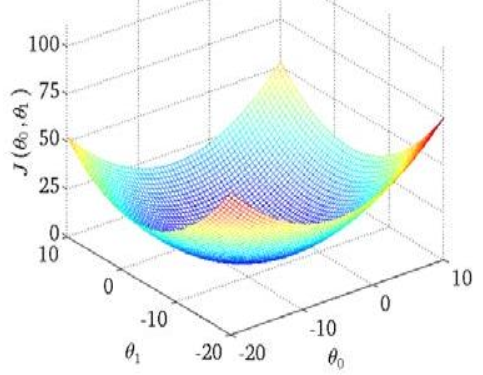
1）方程：

只含有一个特征/输入变量，称之为单变量线性回归。

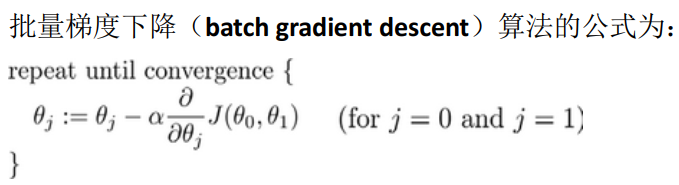


2）代价函数

令代价函数的倒数为零，得到对应的 值，使得建模误差的平方和为最小。如下图，为代价函数的空间图形。



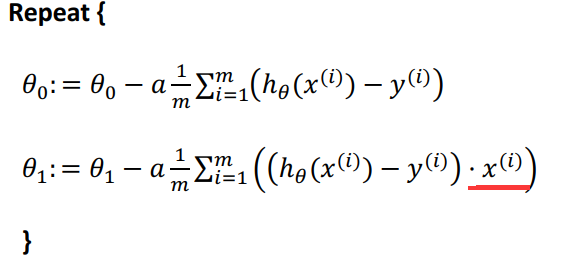
3）梯度下降法来求解 的最小值



理论上，更新参数，需要保证 同步更新。具体做法如下：



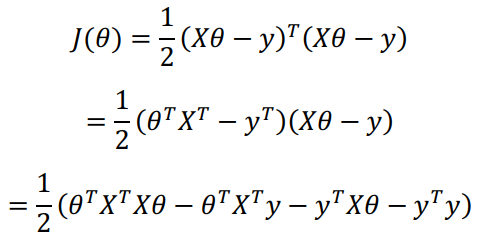
在程序中的实现方式：

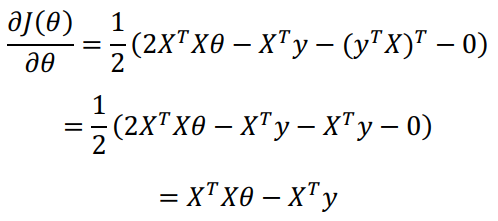


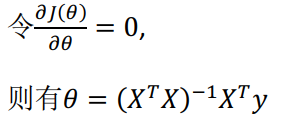
4）正规方程法来求解 的最小值

，一步到位，无需使用迭代、学习速率。

具体推导：







**2、多变量线性回归**

**整体思路：**

与单变量线性回归一样。

**详细内容：**

1）方程：

模型中，含有多个变量，模型中的特征为

2）求解 ，也可使用梯度下降法、正规方程法。

特征数量（>10000）时，使用梯度下降法，会提高计算速度。

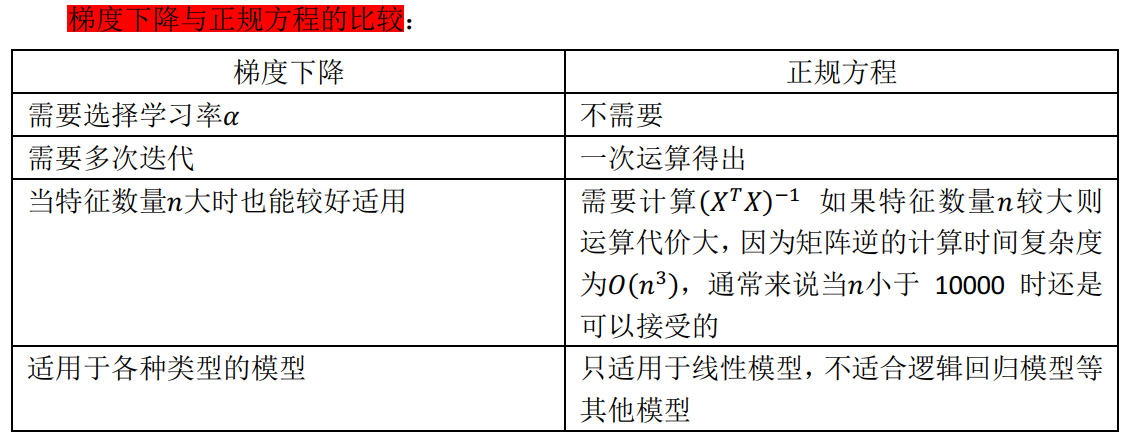
特征数量（<10000）时，使用正规方程法，较为快捷。

3）**备注：**

 在**使用梯度下降法**时，若**特征之间的数量级相差较大**，需要先使用**特征缩放**，再进行的求解。

特征缩放，将其化为标准正态分布，方程

其中。 在得到后，使用来测试数据时，必须对测试的数据也进行相应的标准化处理（否则得到的结果会出错）。



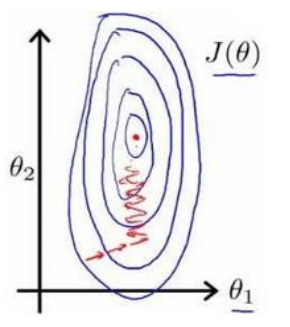
**3、特征缩放**

在我们面对多维特征问题的时候，我们要保证这些特征都具有相近的尺度，这将帮助梯度下降算法更快地收敛。

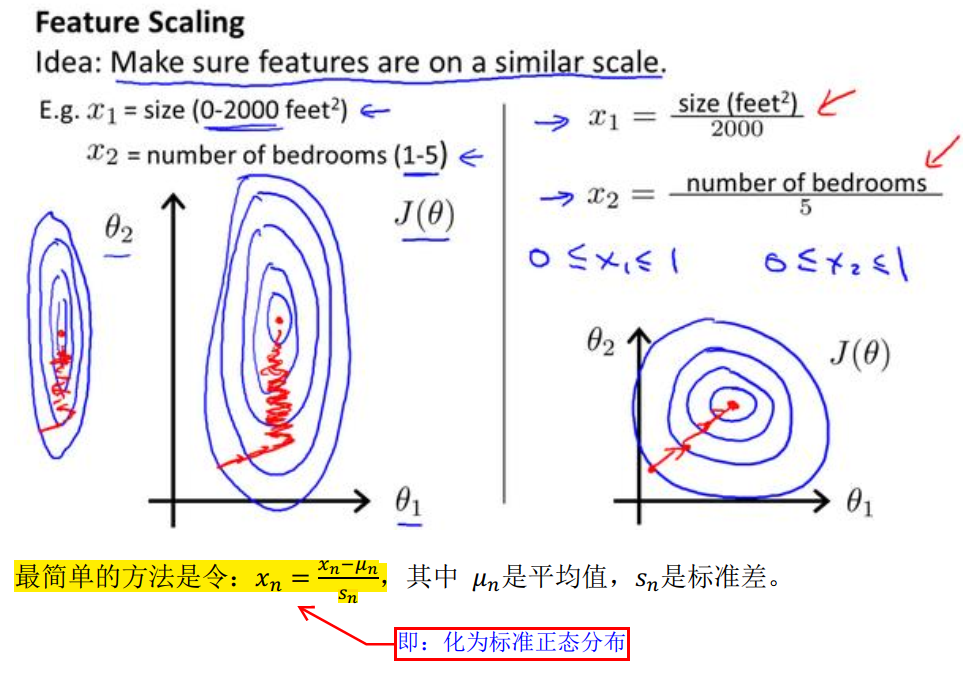
以房价问题为例，假设我们使用两个特征，房屋的尺寸和房间的数量，尺寸的值为 0-2000 平方英尺，而房间数量的值则是 0-5，以两个参数分别为横纵坐标，绘制代价函数的等高线图能，看出图像会显得很扁，梯度下降算法需要非常多次的迭代才能收敛。

**（注意：要对所有的特征变量进行缩放）**

等高线图如下所示：



解决的方法是尝试将所有特征的尺度都尽量缩放到-1 到 1 之间。



示例代码：

**% 特征缩放**

% 求解平均值：X\_norm：两列，是各个参数的平均值

mu = mean(X\_norm);

% 求解标准差

sigma = std(X\_norm);

% 标准化步骤

X\_norm(:,1) = (X\_norm(:,1)-mu(1))/sigma(1);

X\_norm(:,2) = (X\_norm(:,2)-mu(2))/sigma(2);

**% 测试数据时，必须对测试数据进行标准化**

price = [1 (([1650 3]-mu) ./ sigma)] \* theta ;