17.4.14（线性回归的梯度下降算法和正规方程）

1. 梯度下降算法：

   a. theta = theta - (alpha/m)*[sum((X*theta)-y),sum((X*theta)-y)*X(:,2)]'

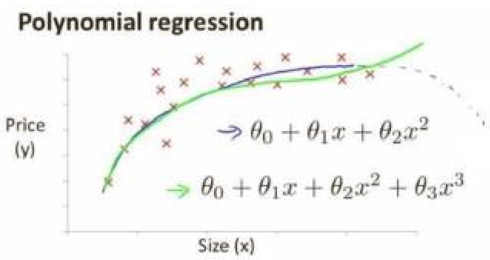
   b. 由于[sum((X*theta)-y),sum((X*theta)-y)*X(:,2)]'=X'*(X*theta-y) 假设函数求出来的Y矩阵与X矩阵的转置矩阵相乘，矩阵乘法，行与列对应相乘并相加。

2. 特征和多项式回归（Features and Polynomial Regression）：有的时候某些有关系的特征可以联系起来用一个特征计算。实际的拟合函数需要用到二次或三次，乃至更高次时，就会用到多项式回归，让你的算法更贴合实际。这种情况下特征缩放是必不可少的。

线性回归并不适用于所有数据，有时我们需要曲线来适应我们的数据，比如一个二次方

模型：$h_\theta(x)=\theta_0+\theta_1x_1+\theta_2x_2^2$

或者三次方模型：$h_\theta(x)=\theta_0+\theta_1x_1+\theta_2x_2^2+\theta_3x_3^3$



通常我们需要先观察数据然后再决定准备尝试怎样的模型。 另外，我们可以令：

$x_2=x_2^2$

$x_3=x_3^3$

从而将模型转化为线性回归模型。

3.正规方程（Normal Equation）：

正规方程是解析式计算，MATLAB中方程写为：theta = pinv（X'*X）*X ‘*y。但是对于那些不可逆的矩阵，正规方程是不可用的。

矩阵不可逆一般有两种情况：

1.冗余的特征，即两种特征之间存在某种线性关系

2.特征的量太大导致正规方程的时间复杂度太大

总之当你发现的矩阵 X'X 的结果是奇异矩阵，或者找到的其它矩阵是不可逆的，我会建议你这么做。

首先，看特征值里是否有一些多余的特征，像这些 x_1 和 x_2 是线性相关的，互为线性函数。同时，当有一些多余的特征时，可以删除这两个重复特征里的其中一个，无须两个特征同时保留，将解决不可逆性的问题。因此，首先应该通过观察所有特征检查是否有多余的特征，如果有多余的就删除掉，直到他们不再是多余的为止，如果特征数量实在太多，我会删除些 用较少的特征来反映尽可能多内容，否则我会考虑使用正规化方法。

如果矩阵 X'X 是不可逆的，（通常来说，不会出现这种情况），如果在 Octave 里，可以用伪逆函数 pinv ( ) 来实现。这种使用不同的线性代数库的方法被称为伪逆。即使 X'X 的结果是不可逆的，但算法执行的流程是正确的。总之，出现不可逆矩阵的情况极少发生，所以在大多数实现线性回归中，出现不可逆的问题不应该过多的关注 X^TX 是不可逆的。