梯度下降算法:

思想： 1.随机初始化

2.迭代，新的能够使得J()更小

3.如果J()能够继续减少，返回2做



:学习率或步长 J()为目标函数

批处理(批量)梯度下降算法：

拿到所有样本后，迭代直到收敛。

随机梯度下降算法：

拿到一个样本后，迭代，再拿到第二个样本后再次迭代，直到最后一个样本迭代结束。

即loop{

For i=1 to m , {



}

}

Mini-batch梯度下降算法：

每拿到若干样本的平均梯度作为更新方向，假如共有m个样本，每个训练集的子集都有c个样本，那么整个训练集可以分为m/c个mini-batch，c等于1时就是随机梯度下降。