Atitit 深度学习原理

把所有的错的平方都加起来，再平均一下，就得到一个“损失函数”。

损失函数当然越小越好，这代表我们的神经网络更厉害。

我们一开始可能在山顶，这意味着“损失”特别大，所以我们要尽快下山止损。

怎么做呢？一种最常用的也很有效的办法就是，永远选择最陡的路线，这样很容易就会到达某个山谷。

当损失函数达到某个最小值，不怎么变化了，就说明训练差不多了。这种方法就被叫做“梯度下降”。

具体而言，其实就是求导数：