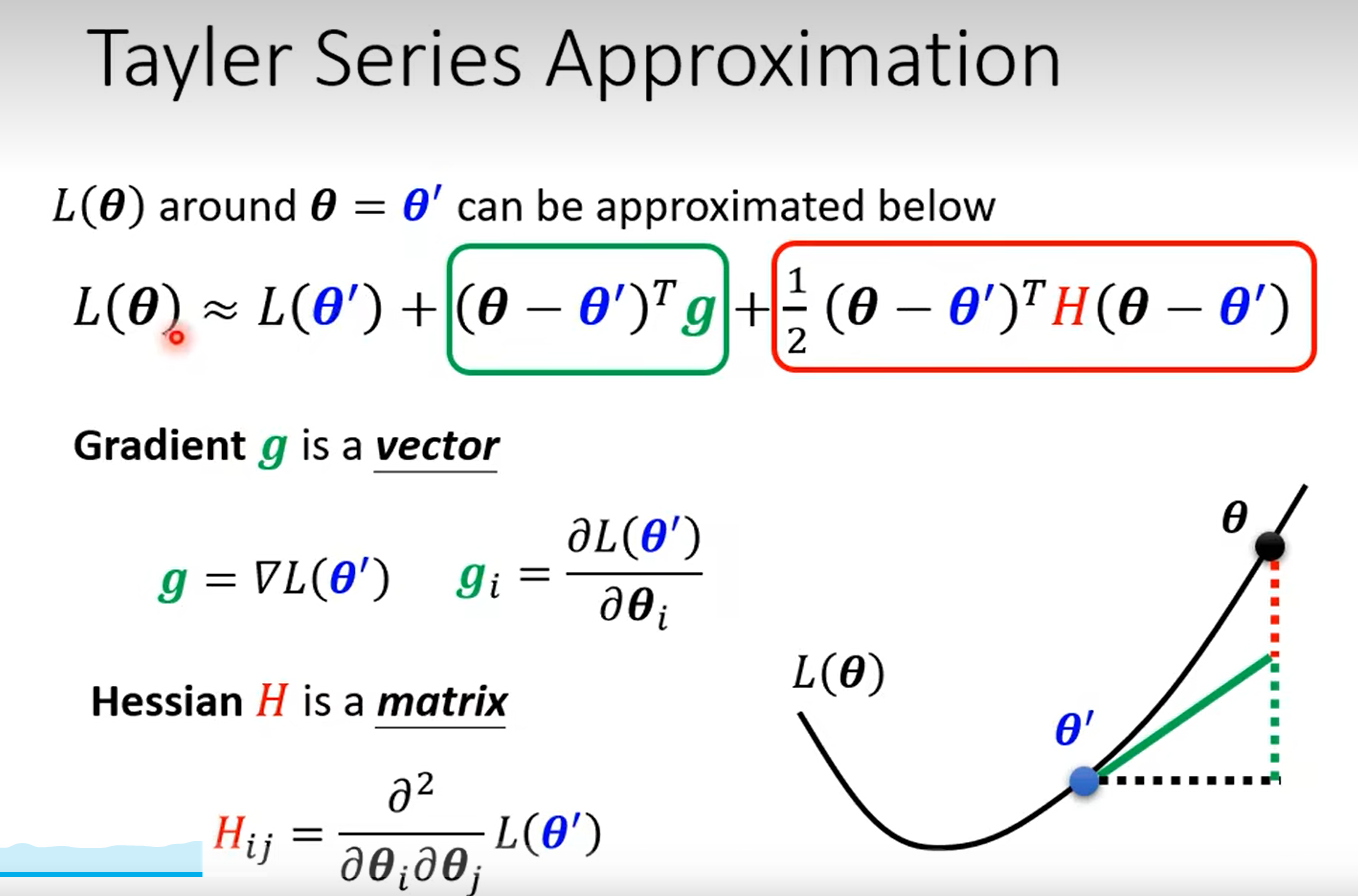
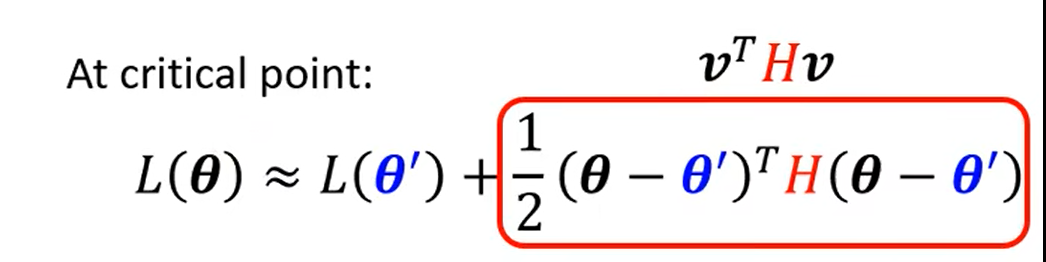
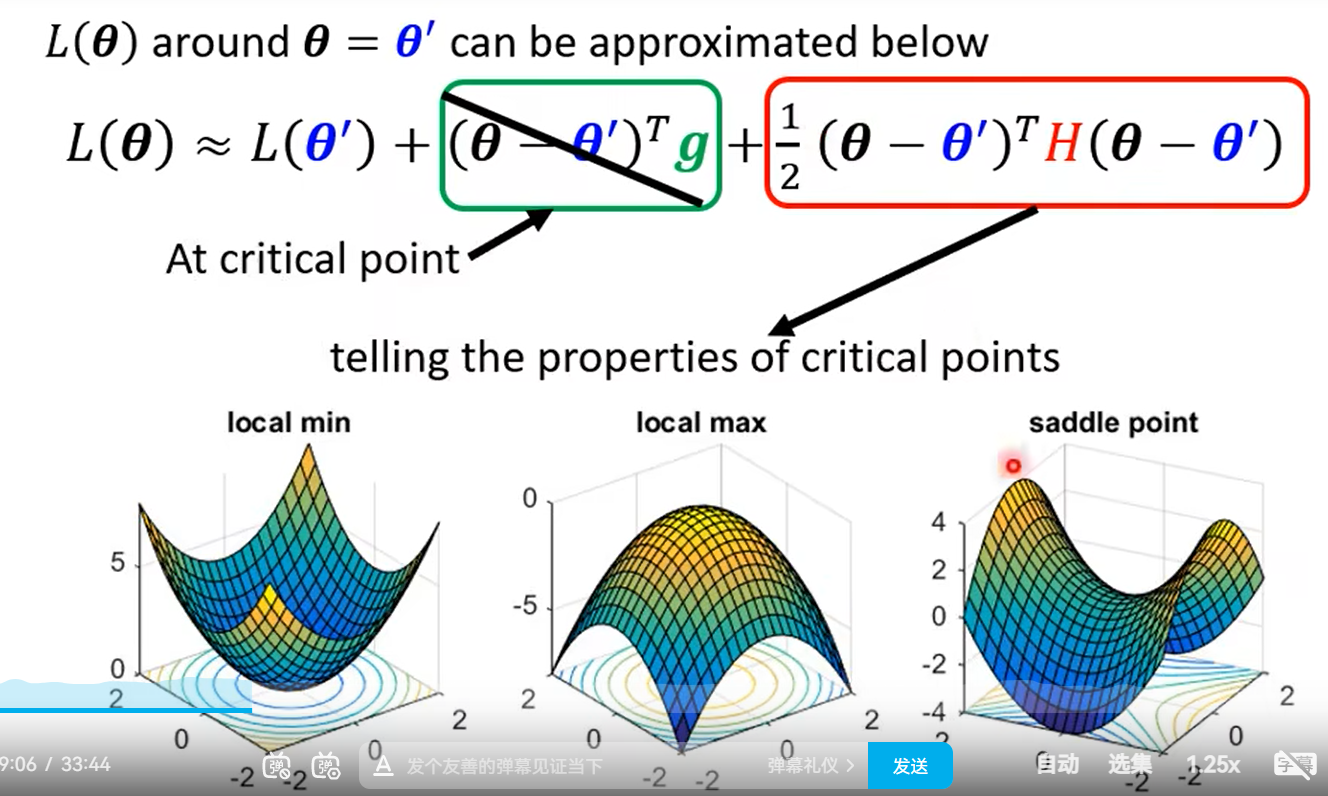
Critical point 包括local minima和saddle point

## 如何判断是局部极小点还是鞍点？

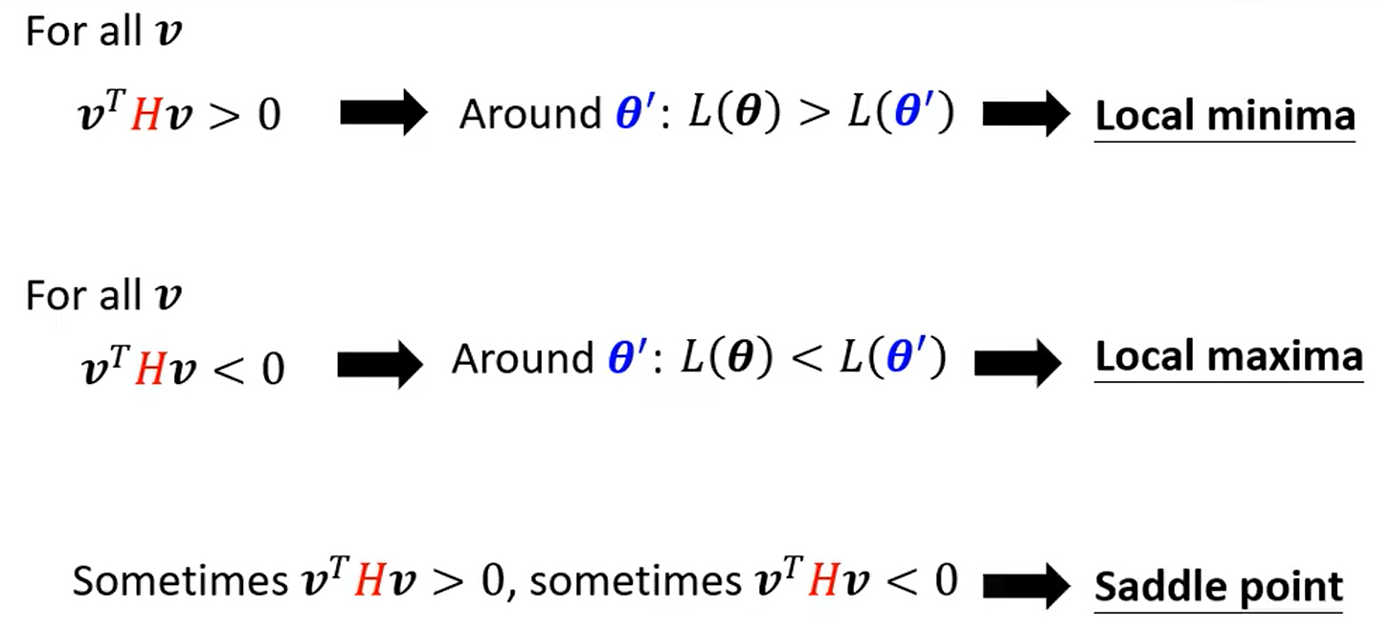


Critical point的位置，就是梯度为0，此时损失函数可以写诶陈：

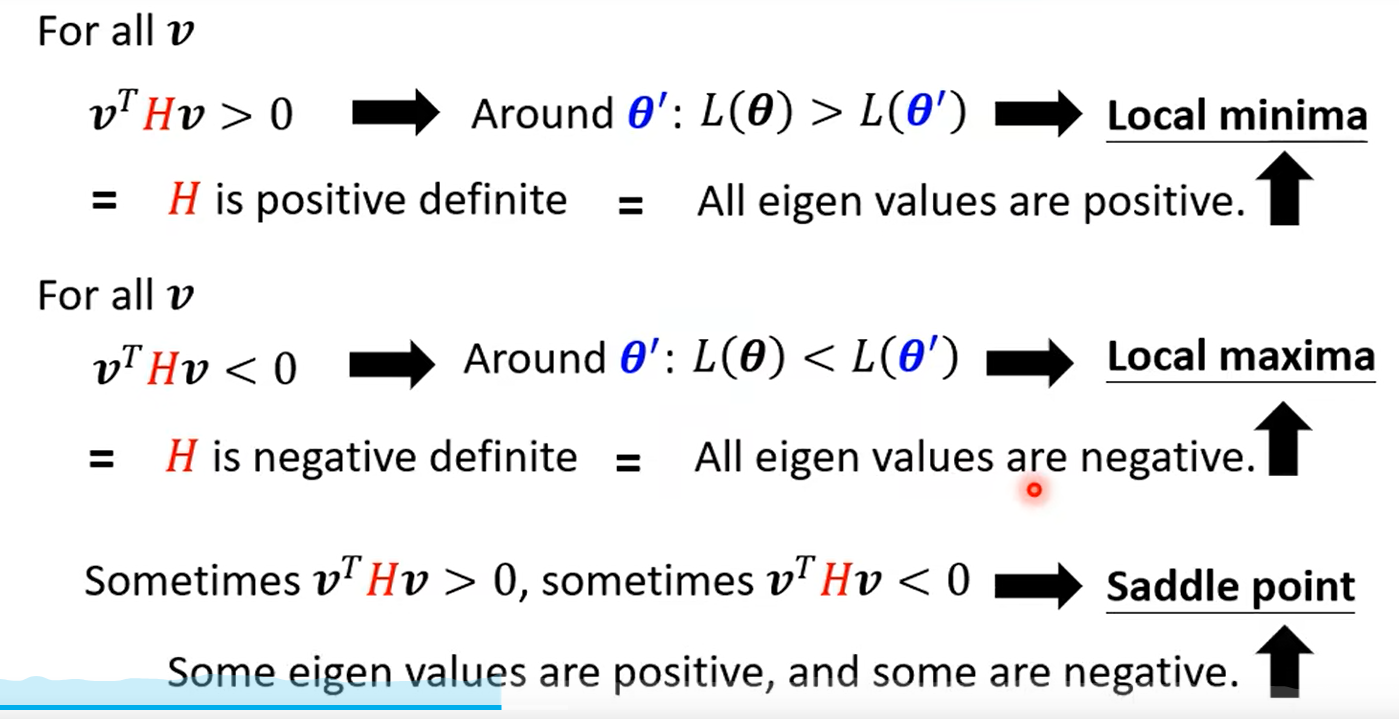




其实就是按照多元微分的知识，可以知道：



按照线性代数的理论，其实就是看黑塞矩阵是否正定，也就是看其特征值是正的还是负的：

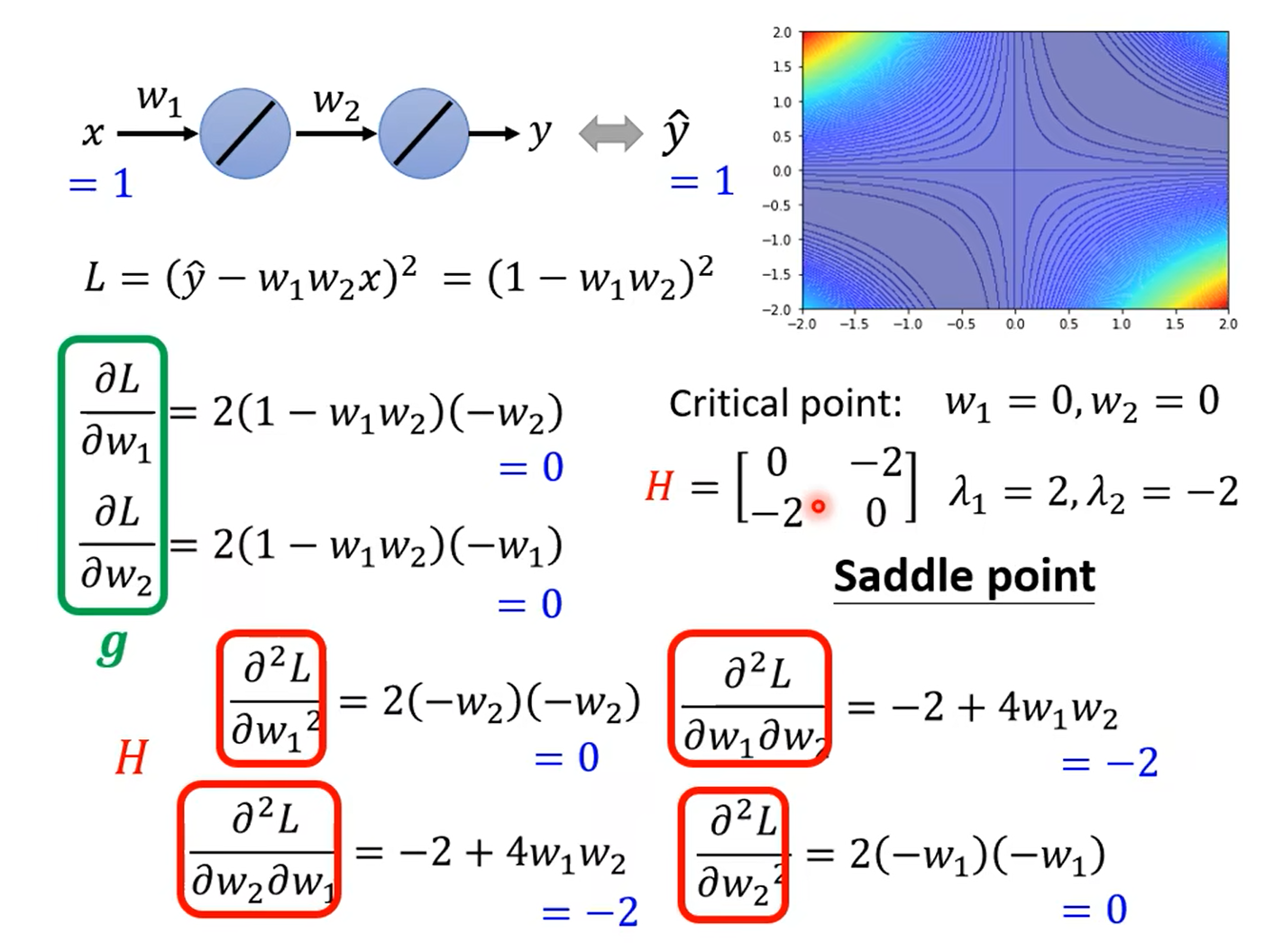


## 2.一个例子

### 2.1 根据error surface可视化

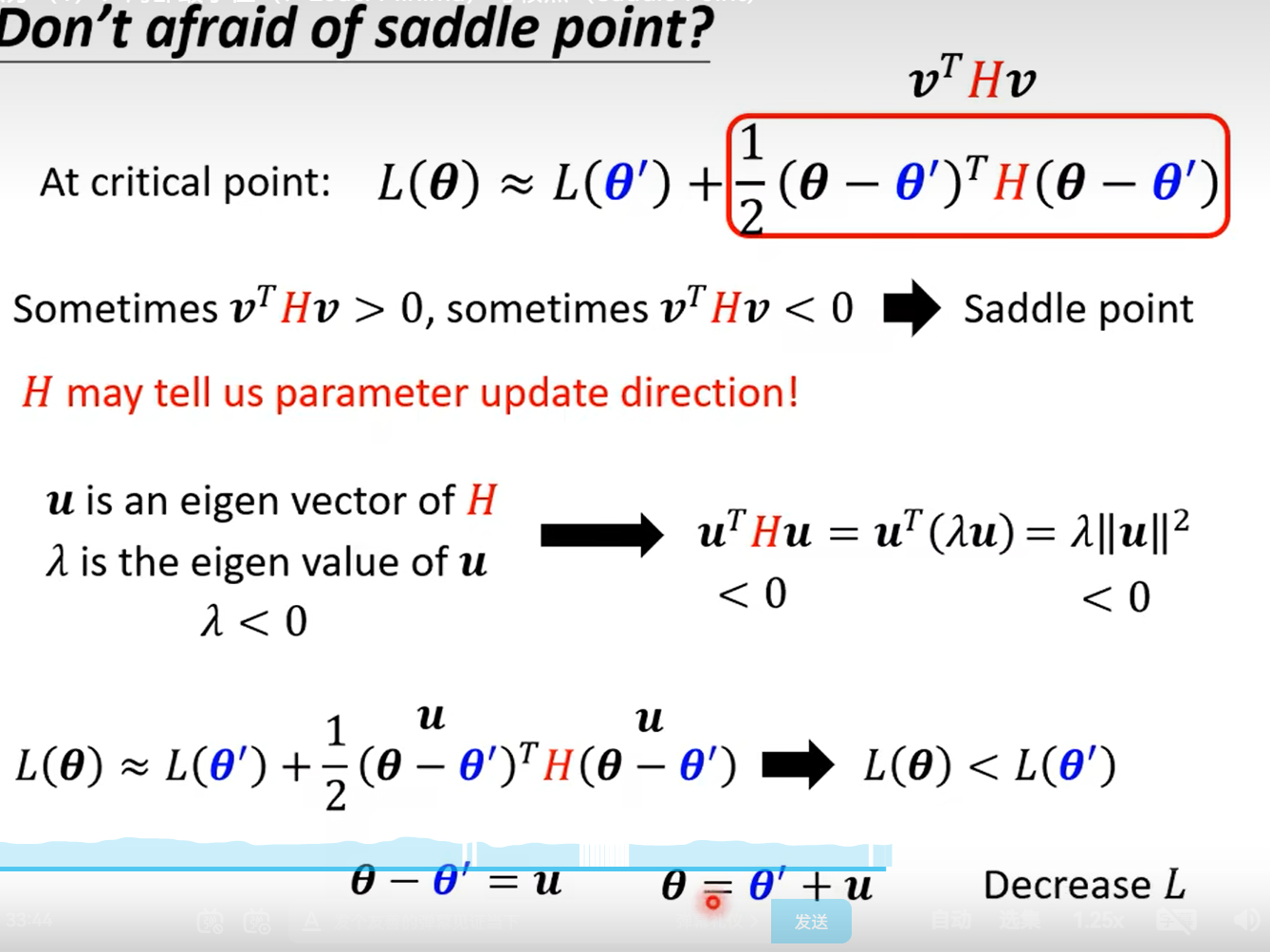


### 2.2 理论计算黑塞矩阵H



## 3.不用担心saddle point

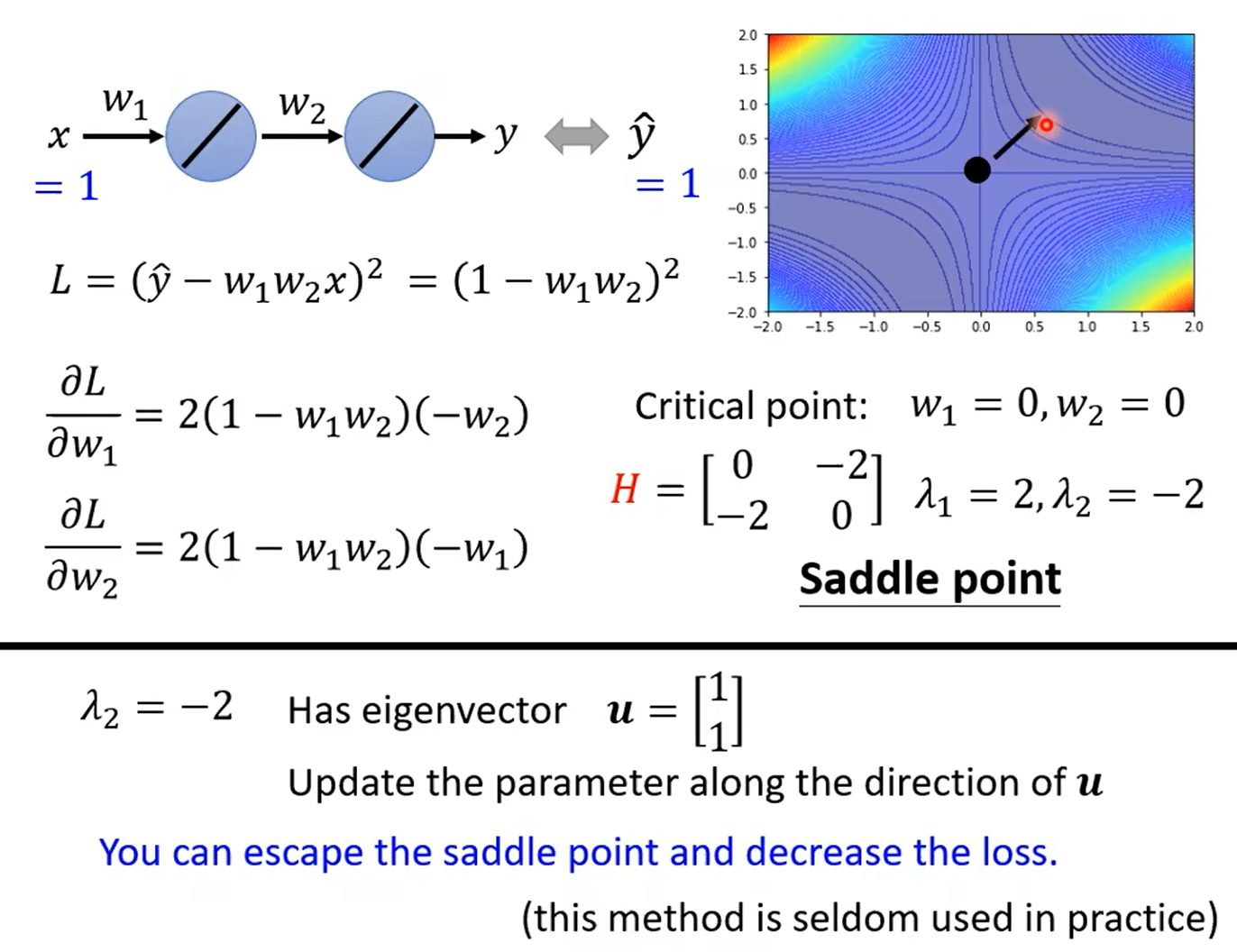
### 3.1 基本原理



特征值小于0的时候，只要沿着特征向量u的方向对参数进行更新，就能够使得损失函数变小

也就是说在saddle的地方卡住了，找特征值小于0的特征向量，对参数就能下降，就能使得损失函数下降。

### 3.2 在2的例子上说明



### 3.3 只是理论，不是实际做法

因为需要计算二次微分，还需要计算特征值和特征向量

## 4.鞍点和极小点哪个更常见？

### 4.1 直观理解

一维参数来看，可能是极值点

但是更高维度来看，更可能是鞍点

因为要在所有维度上，都是极值点，概率还是很小的，这要求黑塞矩阵的特征值全是正数或者负数，是很不常见的。

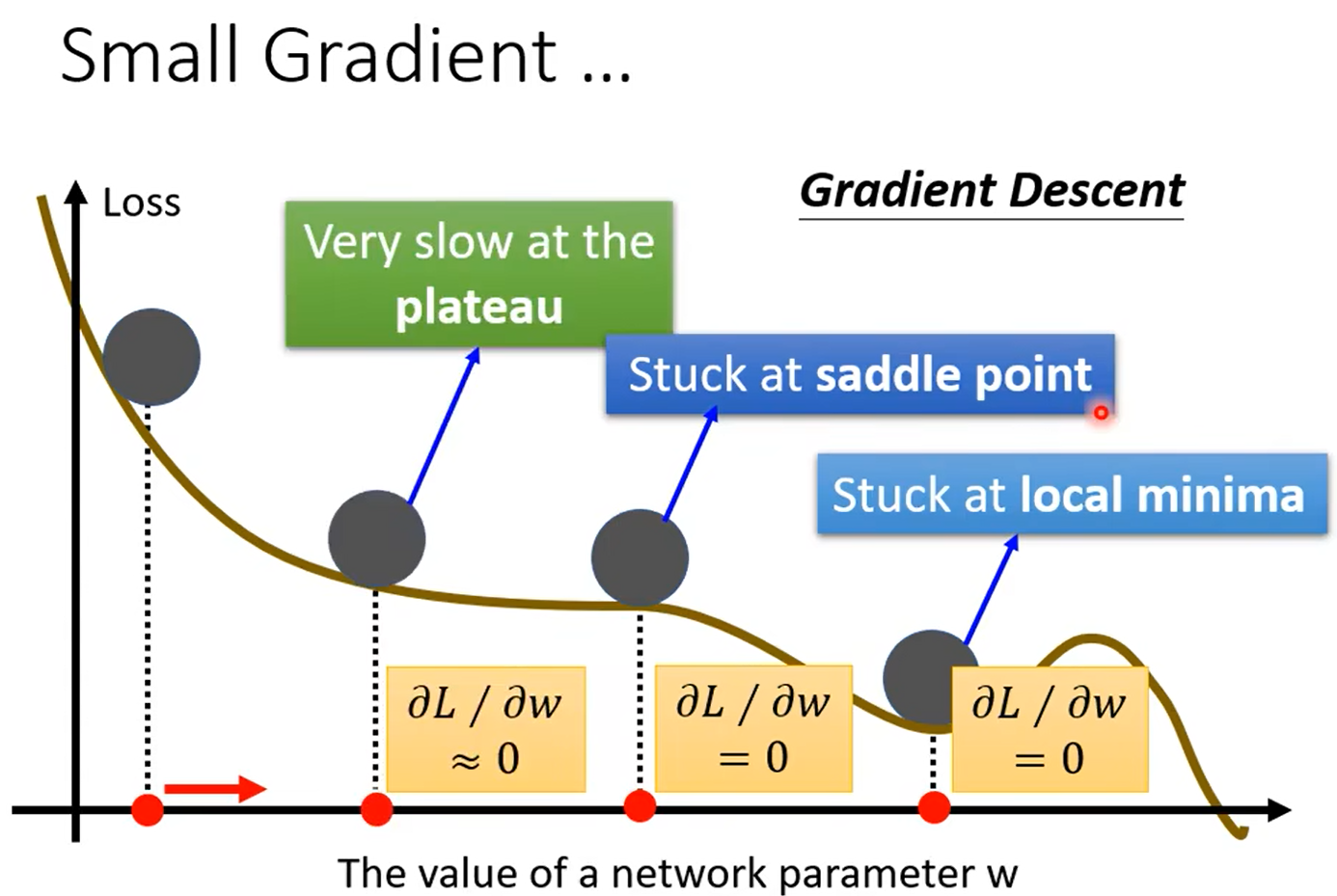
结论就是极值点很少见

### 4.2 经验分析



横轴的意思是x越大，则越有可能是极小值点，但是经验上看，极小值点不是很常见

## 5.小的梯度可能造成的几种问题



可能是平坦区域，可能是鞍点，很少可能是极小值点