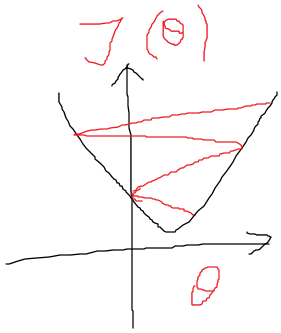
# 梯度方法

1. 公式中，**导数的正**负决定了更新的方向，但是导数的大小会影响步长，但是不能够决定步长，**步长由学习率控制**，
2. 优化算法，优化的**最小**损失函数；

**优化**两个东西，一个是**梯度**，一个是**学习率**；

|  |
| --- |
| 使用梯度更新，  学习率太小：不容易收敛  学习率太大：容易弥散 |



## 正规方程