



人工智能与信息社会

基于神经网络的智能系统I：优化器优化函数

陈斌 北京大学 gischen@pku.edu.cn

如何优化参数

- › 我们已经知道通过损失函数能够定量神经网络预测结果的好坏，那我们如何根据这个评价结果来更改优化我们的参数？
- › 神经网络中的参数是海量的，如果修改的方法不当，就会耗费大量的计算时间。

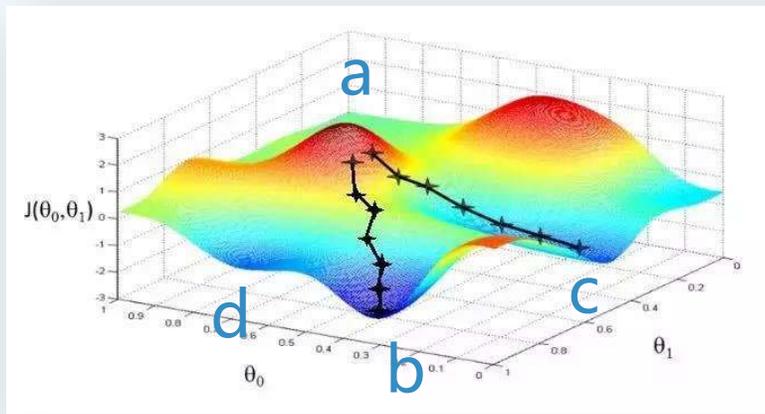


优化器

- › 在损失函数确定情况下，优化网络的任务就交给了优化器。
- › 优化器代表了调整网络参数，使损失函数达到最小的过程。
- › 不同的优化器有不同的算法，改变损失函数的过程和效率也各不相同。

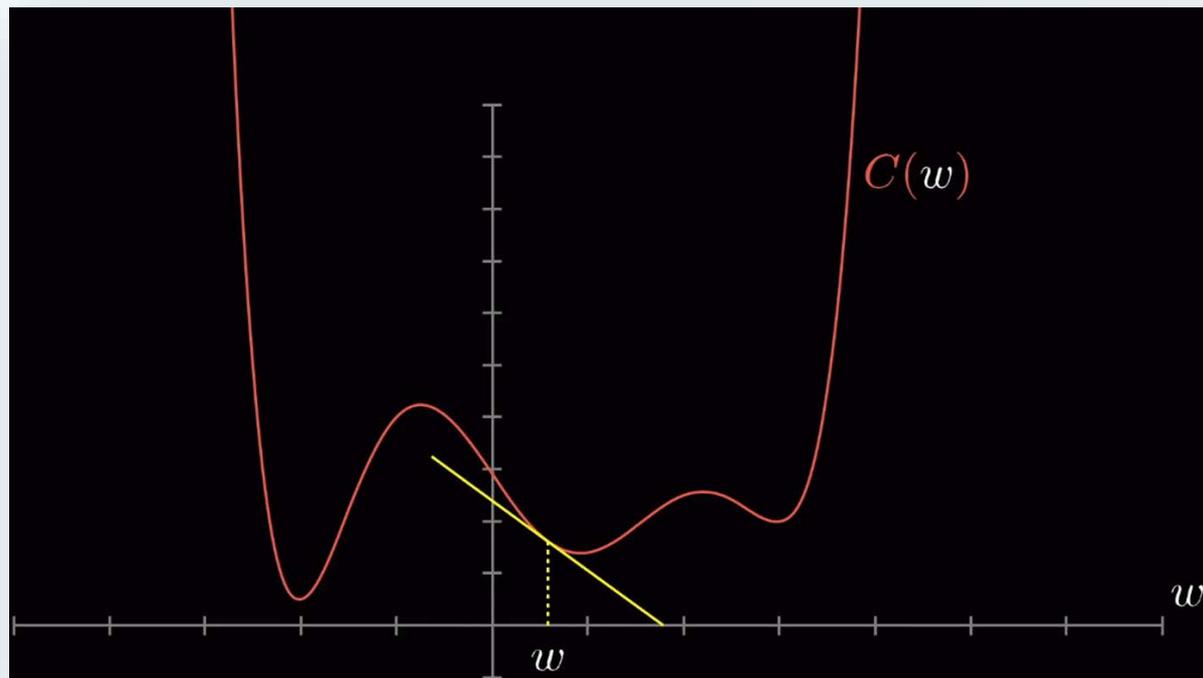
函数优化例子

- › 下图曲面表示具有两个参数的损失函数，可以把它想像成一个起起伏伏的山谷。
- › 假设小明现在处于图中最高点a，要去到最低点b，那该如何选择最优路径？



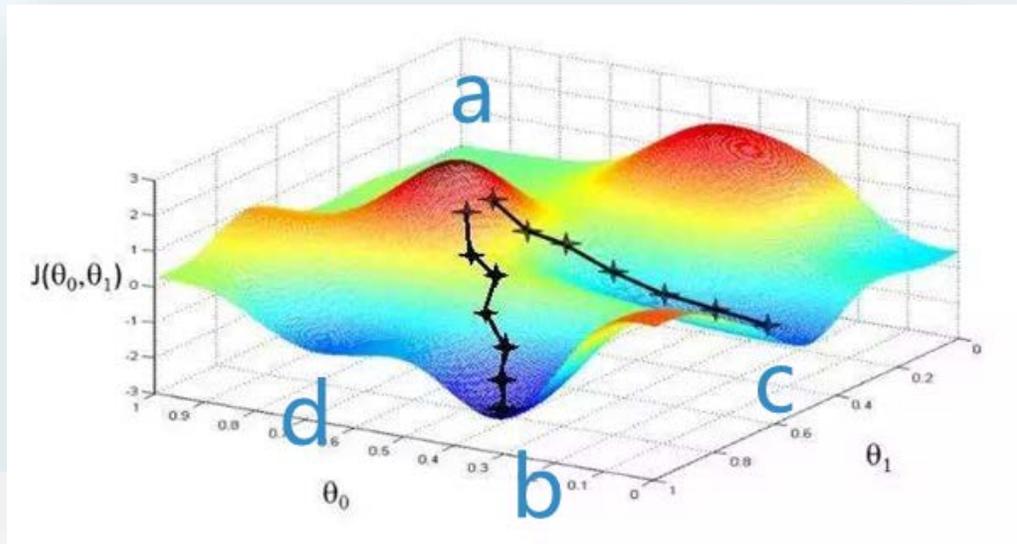
梯度

- › 表示函数在某一点变化最快的方向
- › 图中 w 点的梯度方向就是黄色的切线



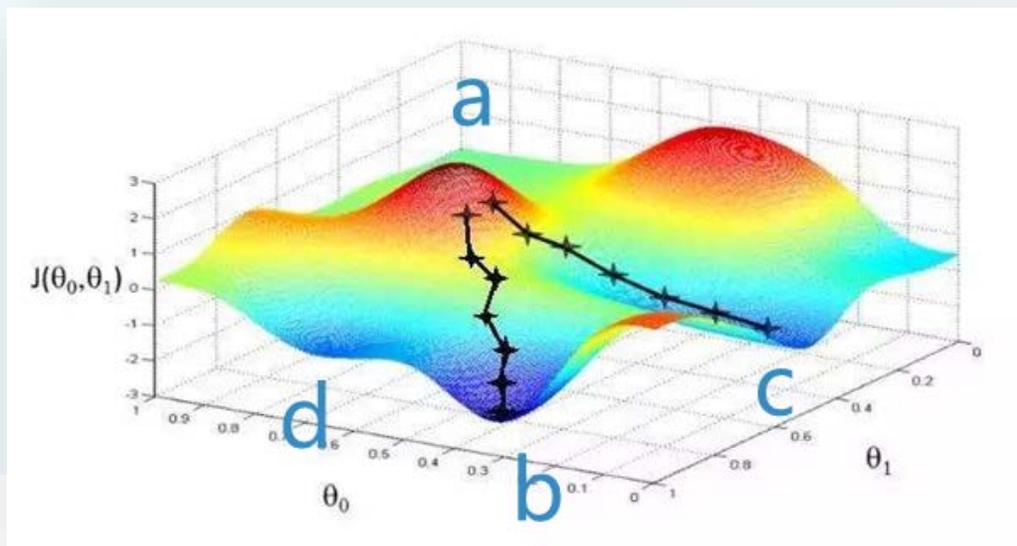
梯度下降算法

- › 梯度下降法的基本思想可以类比为下山的过程。
- › 假设小明被困在山上，浓雾很大可见度低，下山的路径就无法确定，他必须利用自己周围的信息去找到下山的路径。



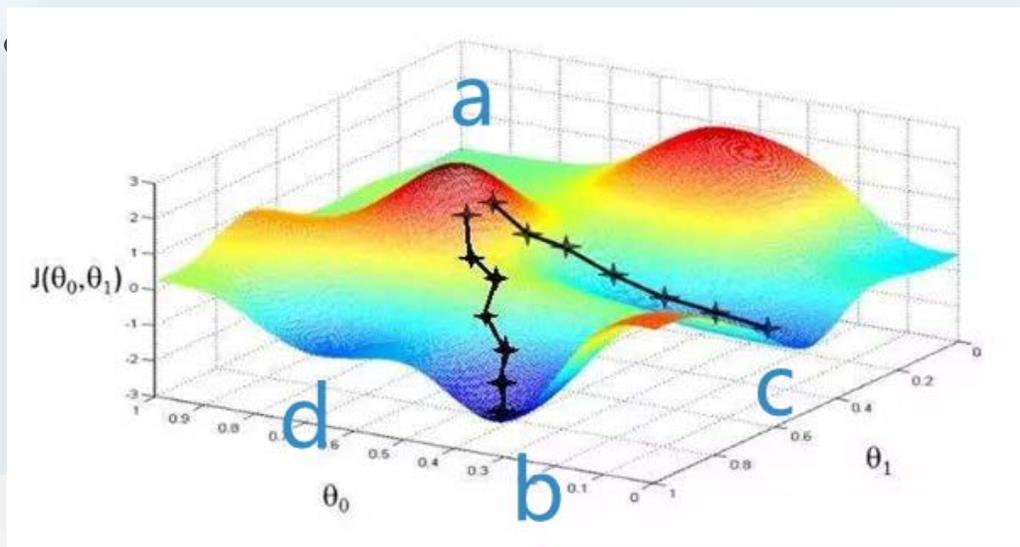
梯度下降算法

- › 这个时候，他就可以利用梯度下降算法来帮助自己下山。
- › 具体来说就是，以他当前的所处的位置为基准，寻找这个位置最陡峭的地方，然后朝着山的高度下降的地方走。



梯度下降算法

- › 小明每次沿着当前位置下山一小步，这样就能保证越来越接近山谷，即得到更小的损失函数，直到到达b，得到损失函数最小值。
- › 如果不小心进入c、d点，那么梯度下降的算法就无法再回到b，这时就需要其他优化算法。



总结

- › **优化器代表了调整网络参数，使损失函数达到最小的过程。**
- › **不同的优化算法的计算过程和效率也各不相同。**
- › **梯度下降算法是最常用也是最有效的优化。**