交差熵优点：误差越大，学习速度越快，权重变化量与激活函数的导数无关。

二次代价函数缺点：如下图，权重变化量与激活函数的导数相关，sigmoid激活函数会导致误差很大时，sigmoid导数很小。

