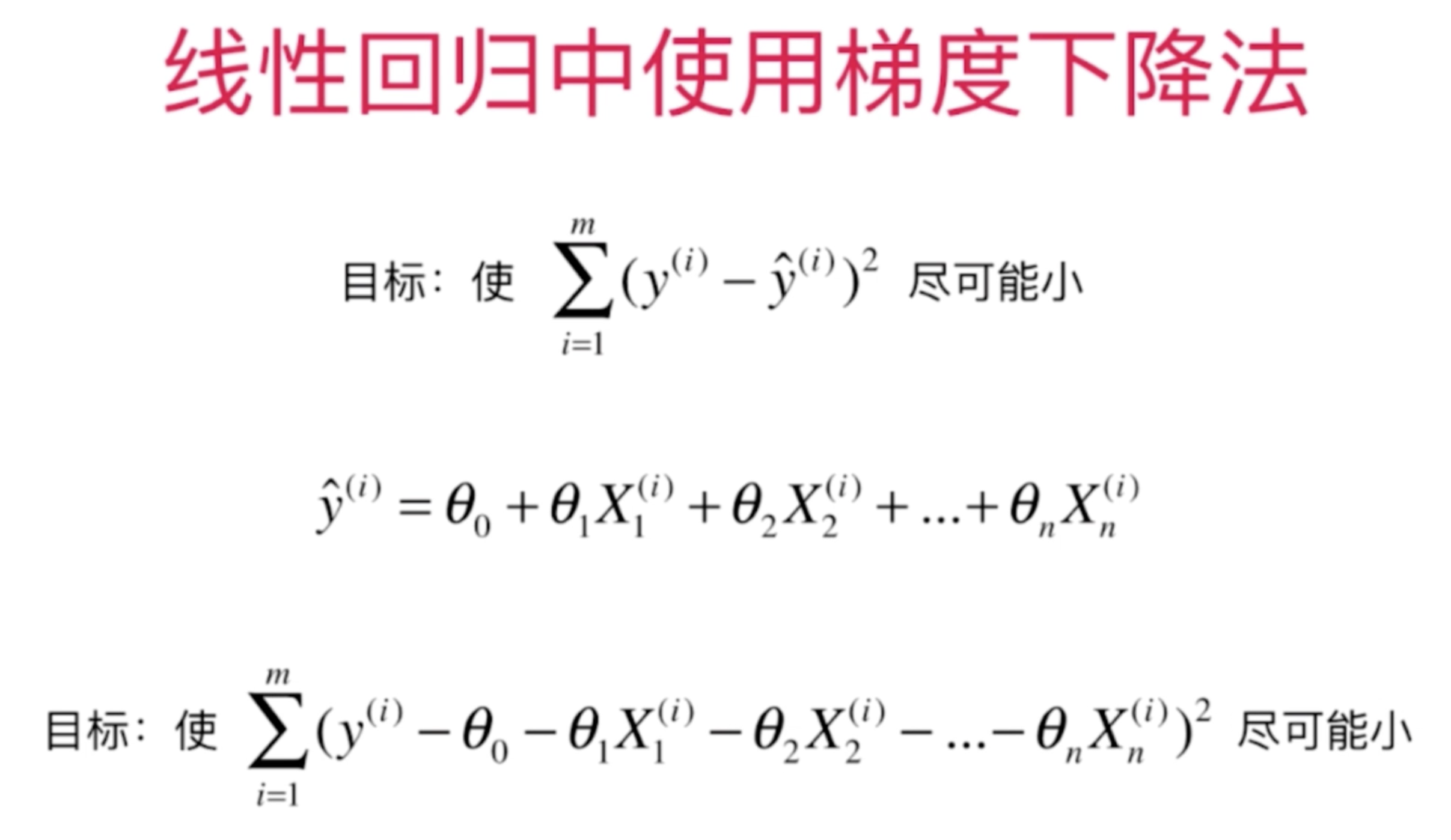
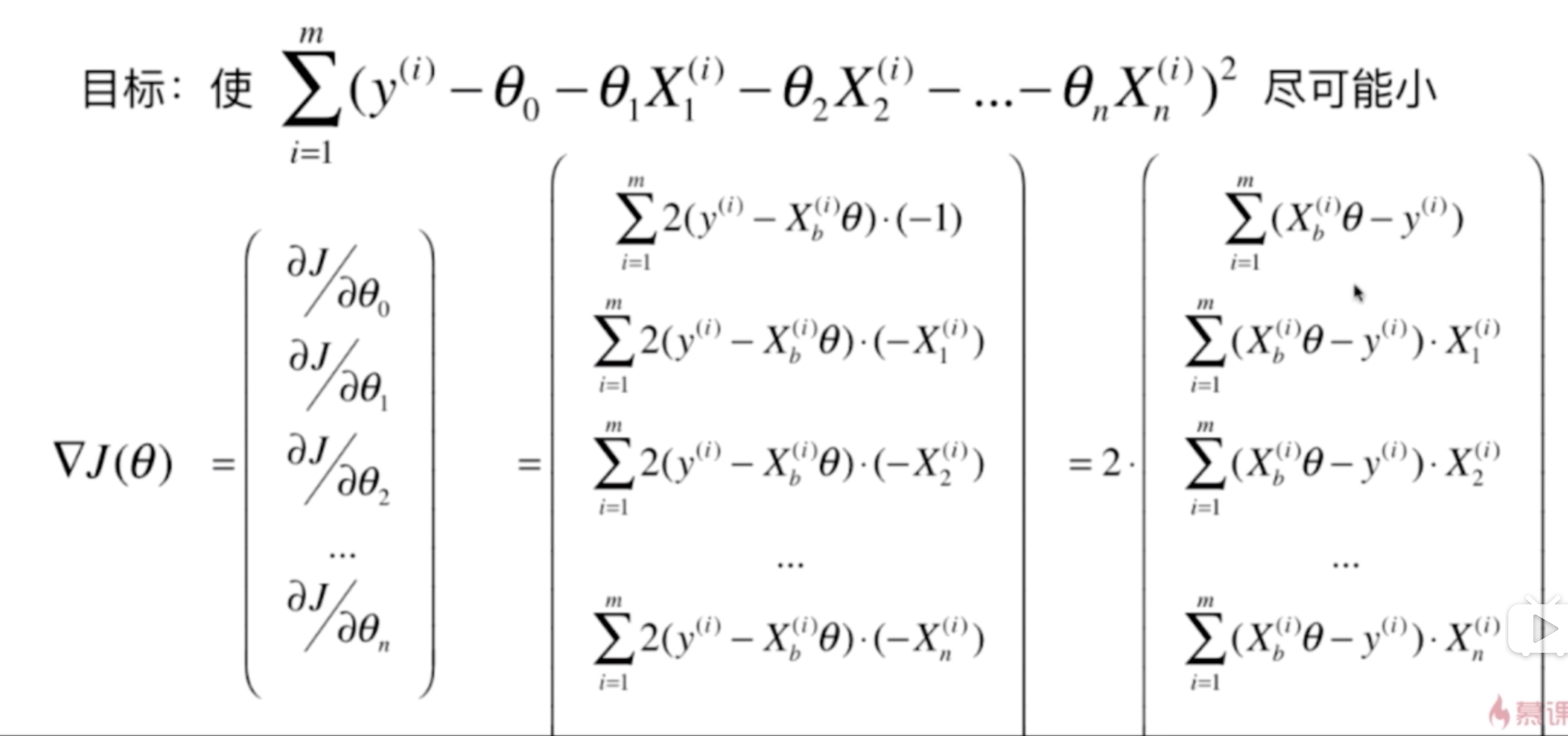


针对于损失函数中有多元theta的情况，就不是针对一个变量的导数，而是针对每个变量求偏导后组成的梯度。



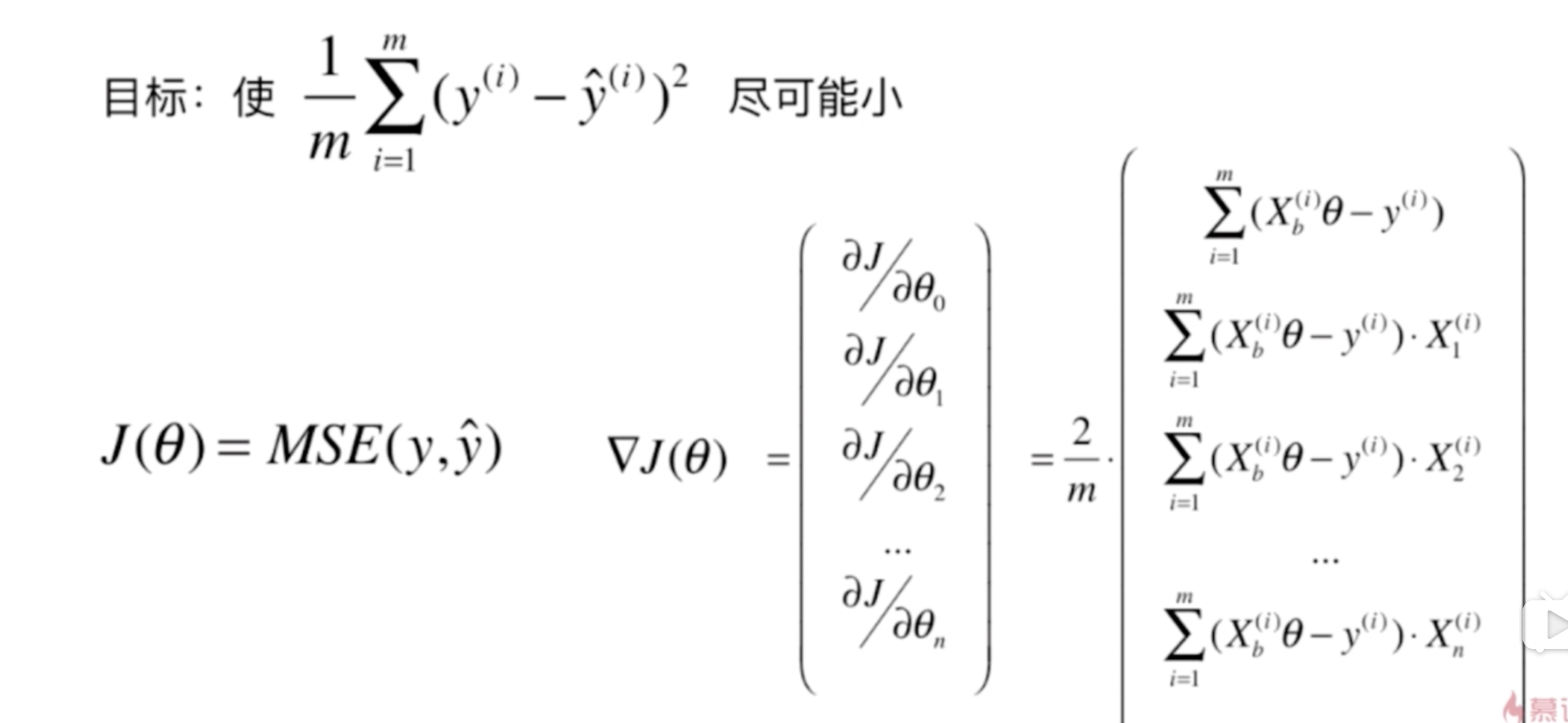
进一步转换：

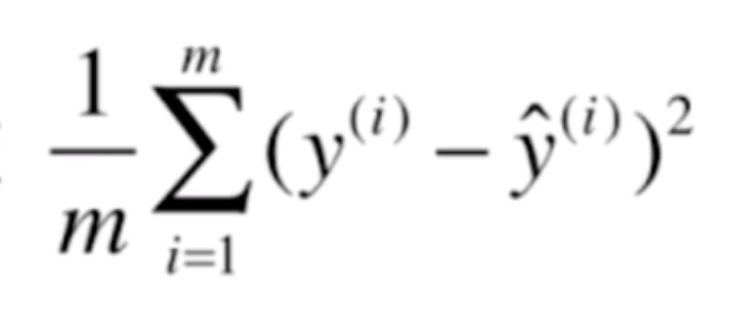


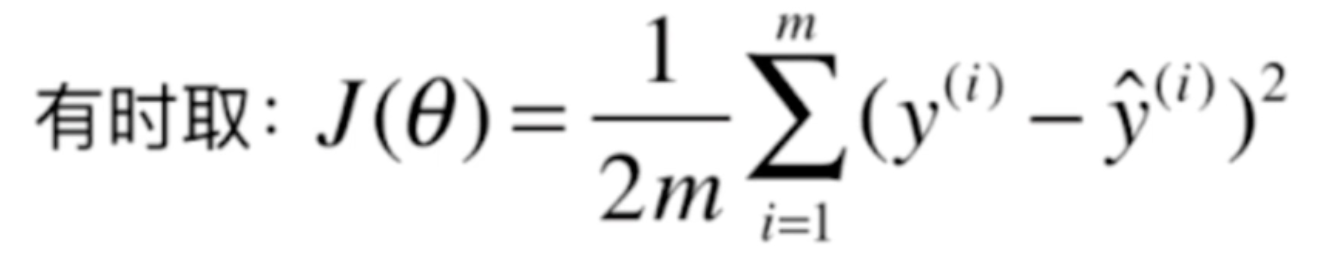
还是注意：对于损失函数来说，这里一堆的 theta 是未知数，而 yi 和 xi 都是已知数，是训练数据集。

其中的 Xbi 是一个 n + 1 的向量，第一个 x0 恒等于 1。theta也是一个n+1 的向量。

但是这个结果和样本的数量m相关，所以还要除以2



还记得公式吗？ 它就是线性回归问题的衡量指标之一：MSE。所以我们的损失函数就是：

但有些书会使用这个

这个 1/2 是为了中和掉矩阵求导时统一提取的系数2，实际运算过程中这个 1/2 倍系数对我们的梯度下降计算过程没什么影响。