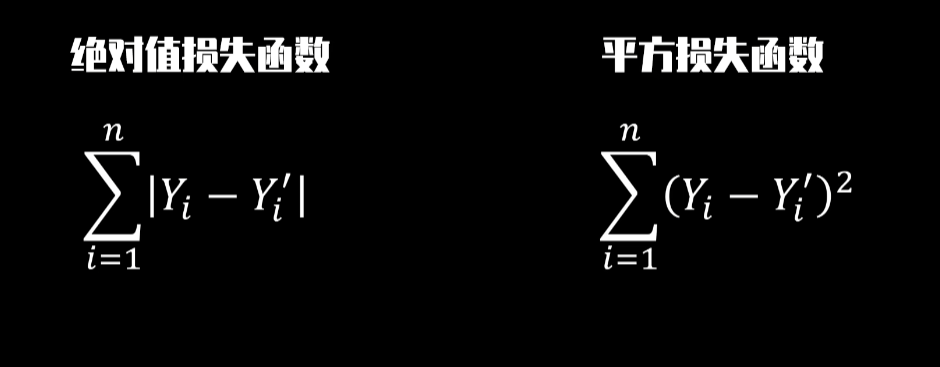
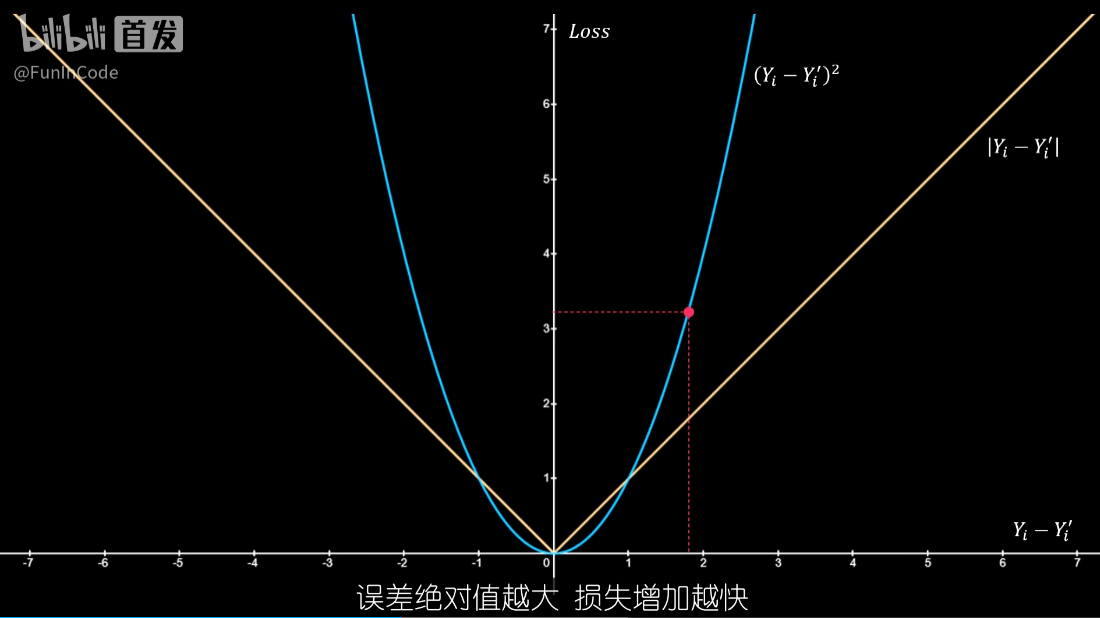
Y=wx+b

w是权重，也是斜率

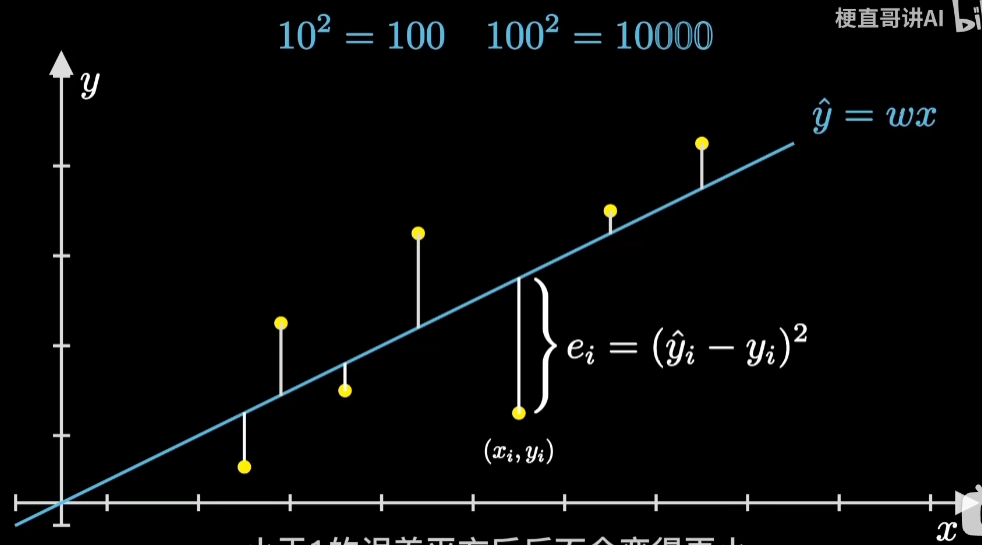
b的偏置值

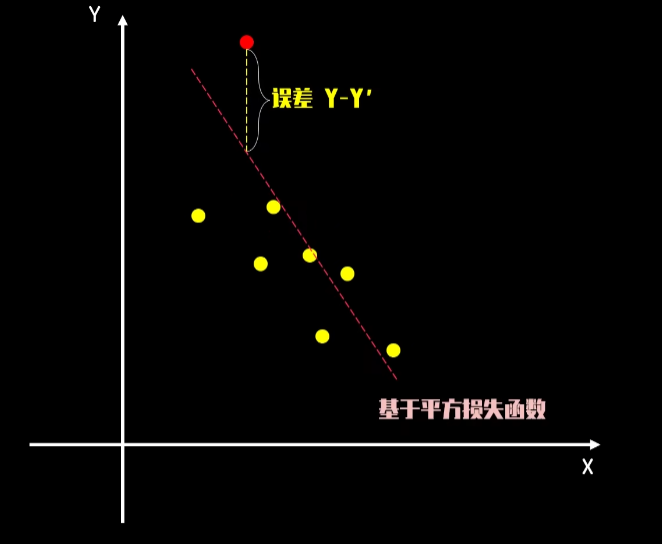


不同的损失函数具有不同的特性

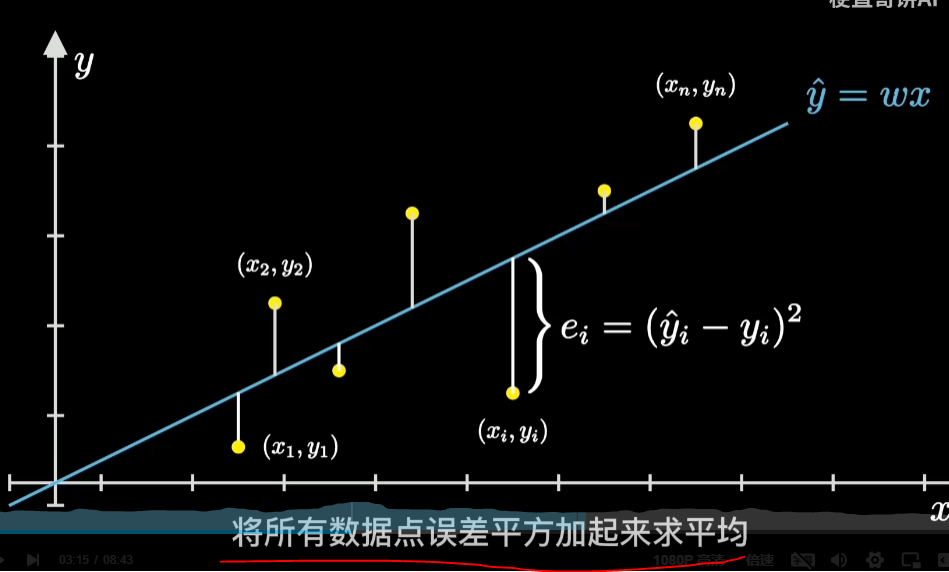


在平方损失函数中，大误差值会对模型产生大的影响，因为x移动小距离，会导致y发生巨变。

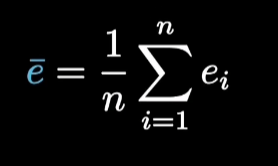




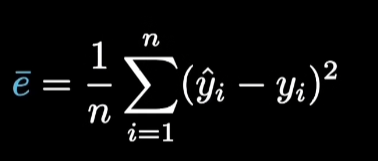
平方损失函数中，大误差变得更大，小误差变得更小。



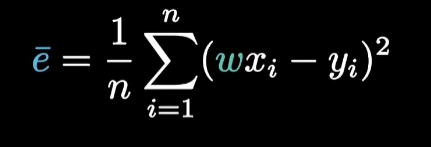
上图的字指的就是均方差

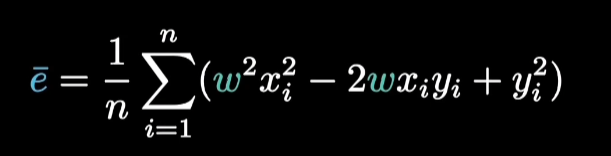


推导如下：

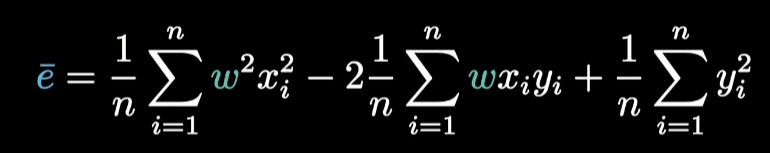


推导



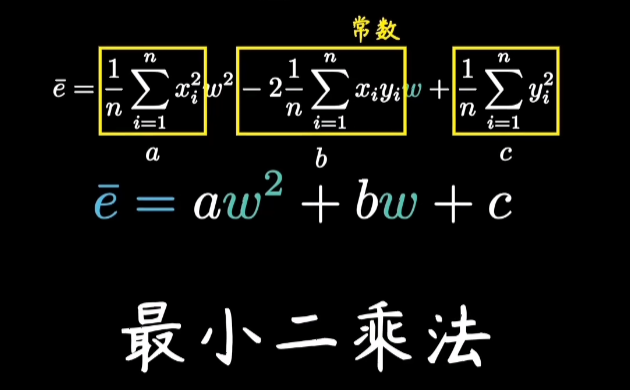


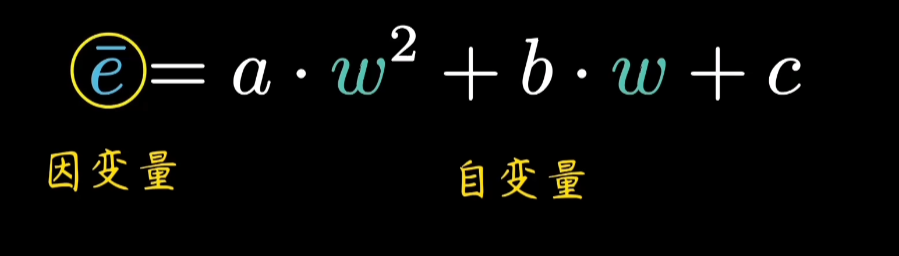
让球求和符号（sigma）进入，合并同类项



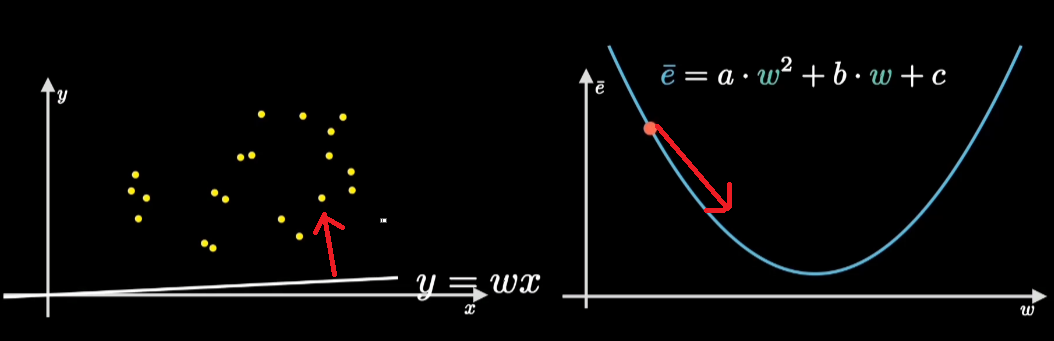
因为所有的

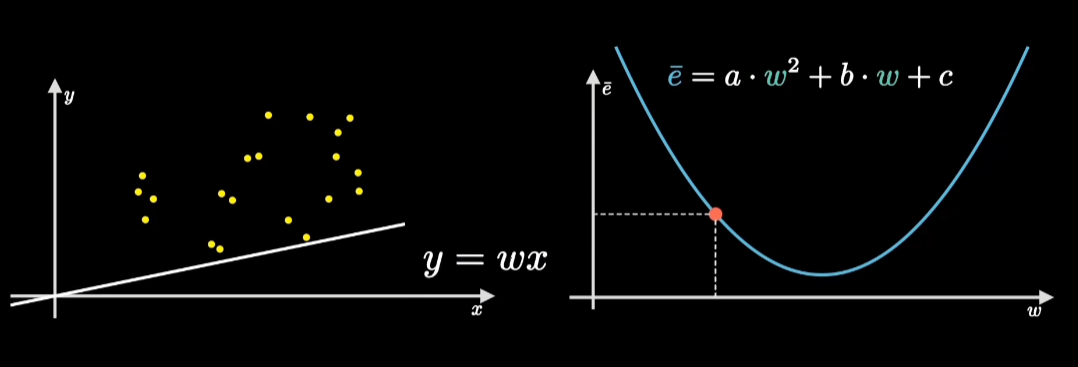
Xi,yi，n都是已知的，这个推导过程就是最小二乘法

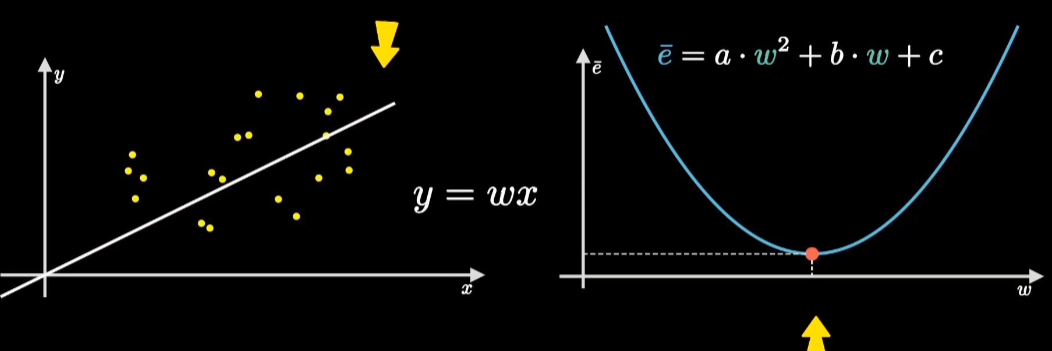




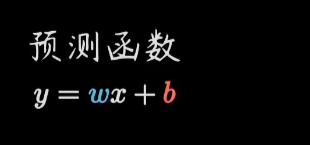
他们的关系图如下。损失函数和代价函数的关系



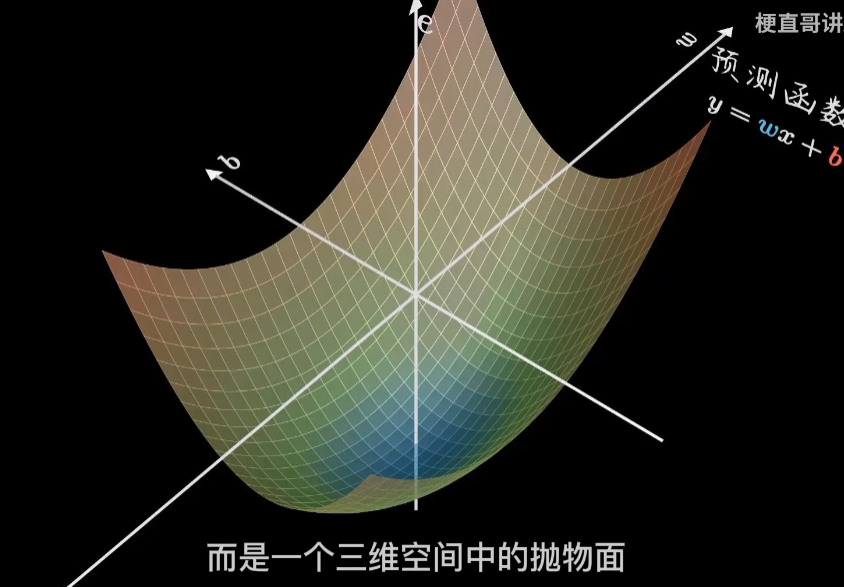


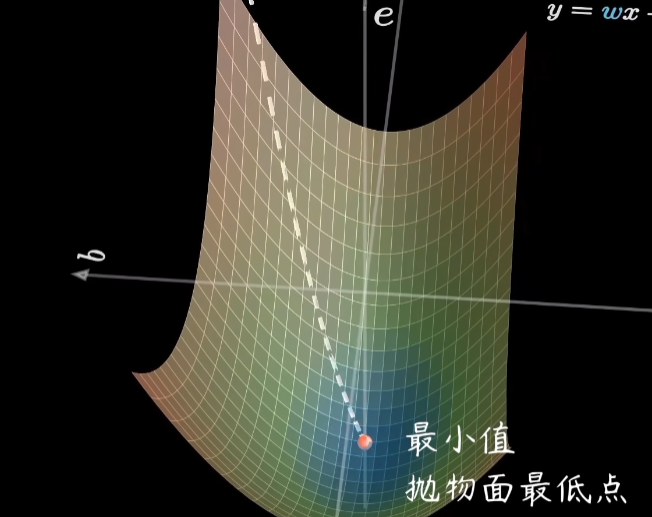


当我们的预测函数不再是y=wx，而是y=wx+b，我们需要调整的参数就变成了w，b

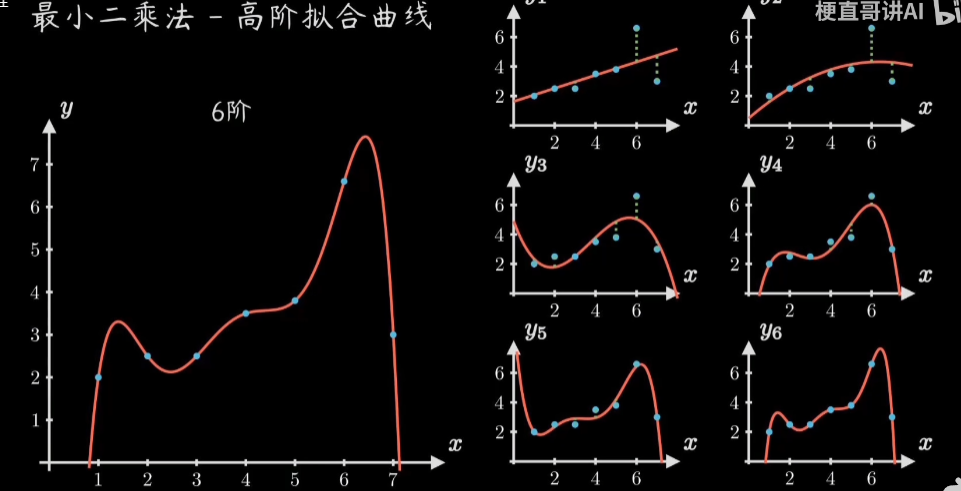


图形就不再是简单的抛物线，而是一个三维平面





选择不同的二次函数，通过最小二乘法可以得到不同的拟合曲线



机器学习的本质就是你设定了怎么样的代价函数，模型就会想你设定的方向去学习，只需要思考为什么需要这种代价就行了。

最小二乘法（最小平方法），后来又推出了最大似然估计法