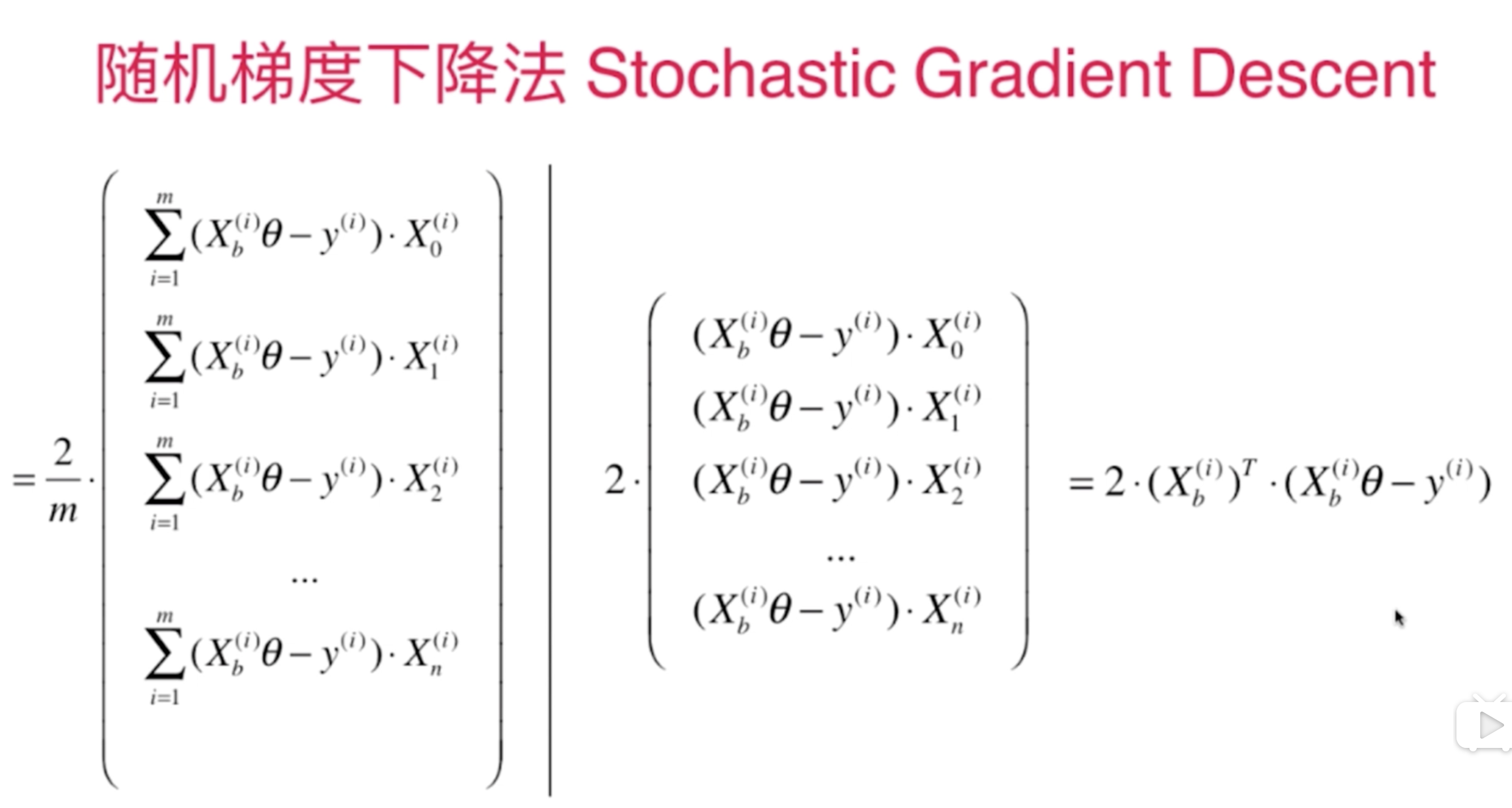
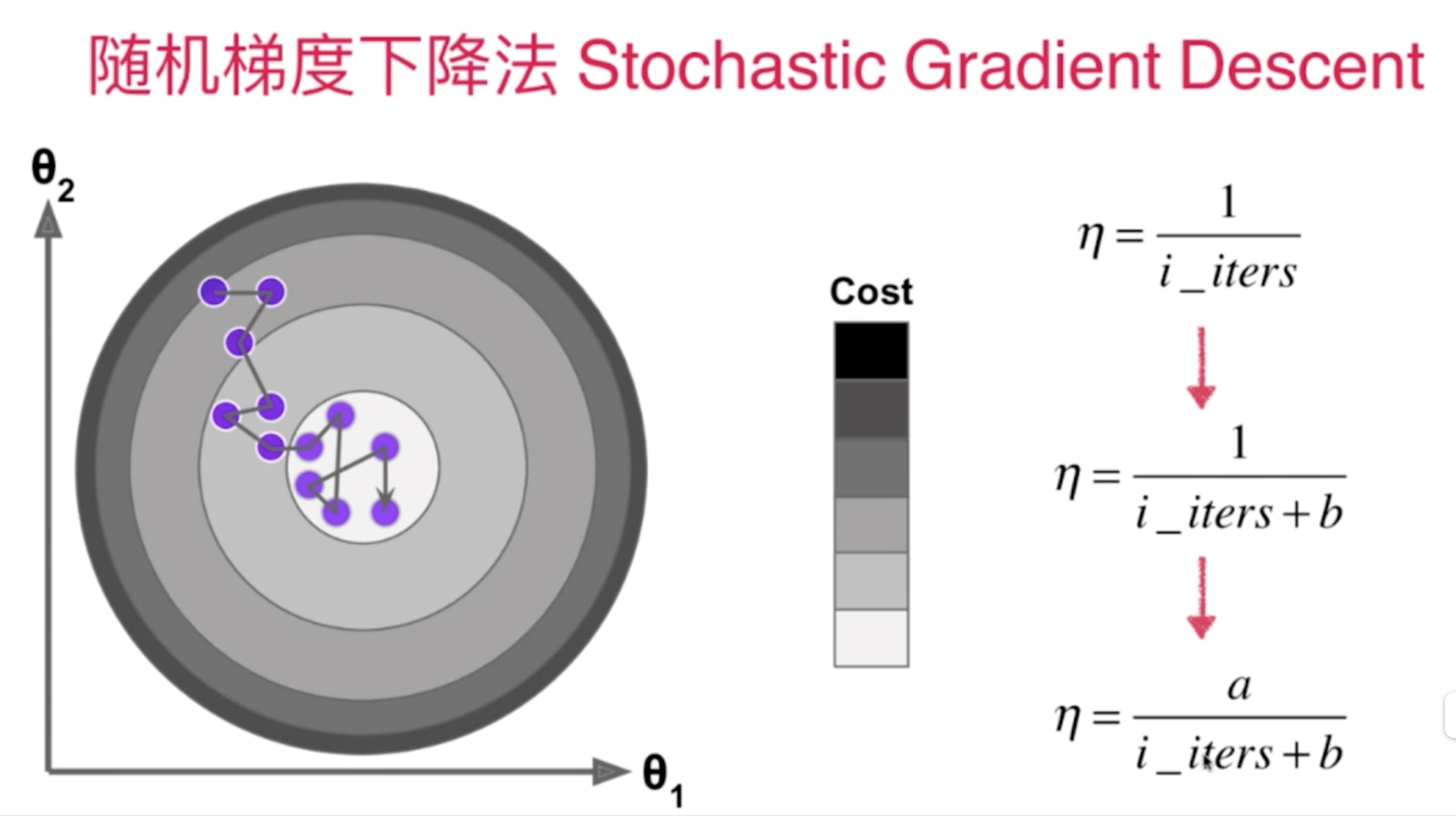


如果样本数量(m)太大，计算梯度也是很耗时的。



每次只随机取一个样本



eta 也要随着循环次数变小，a 和 b都是 超参数，用来调教eta 递减的速度。根据经验值可以默认用 a=5, b=50

