进行back propagation，也就是计算computation graph中每个节点相对于损失函数（这里表示为loss）的导数。如何计算这些梯度，使用的就是BP算法，其实也就是求导的链式法则。

本质就是：梯度下降+链式法则