

}
多层神经网络通常有像悬崖一样的斜率较大区域，如图8.3所示。这是由于几个较大的权重相乘导致的。遇到斜率极大的悬崖结构时，梯度更新会很大程度地改变参数值，通常会跳过这类悬崖结构。



