

对于牛顿法而言，显然鞍点会是一个问题。梯度下降旨在朝“下坡路”移动，而非明确寻求临界点。而牛顿法的目标是寻求梯度为零的点。没有适当的修改，牛顿法会跳进一个鞍点。高维空间中鞍点的扩散或许解释了在神经网络训练中为什么二



