

一般的优化和我们用于训练算法的优化的一个重要不同：训练算法通常不会停止在局部极小点。反之，机器学习通常优化替代损失函数，但是在基于提前终止（第7.8节）的收敛条件满足时停止。通常，提前终止使用真实潜在损失函数，如验证集上的 0-1 损失，并设计为在过拟合发生之前终止。与纯优化不同的是，提前终止时替代损失函数仍然有较大的导数，而纯优化终止时导数较小。

