## 通过提前终止训练来防止过拟合

通常在计算力小而数据量大的情况下采用，这种方法不尝试多种超参数，而是开始时采用一个小权值，然后在验证集上的效果变坏的时候直接终止训练。

具体操作的时候可能较难确定确切拐点，因为误差经常上下波动。通常会超出最佳时间节点，然后折返回去。

这种方法天然地限制了模型的复杂度，因为权值很小并且来不及变大就终止了。

## L2正则

通过向算是函数添加一个额外项来惩罚过大的权值，除非误差相对较大，否则保持较小权值。

将L2正则项视为损失函数的一部分，这么做的目的是防止网络使用不必要的权值。通过阻止模型利用多余的权值拟合采样误差，可以使模型泛化得更好。权值较小，则输出变化得比输入慢，模型更加“平稳”。

