**提前停止**

//参考深度学习 7.8

在每次验证集误差有所改善后，我们存储模型参数的副本。当验证集上的误差在事先指定的循环次数内没有进一步改善时，算法就会终止。当训练算法终止时，我们返回这些参数而不是最新的参数。

**提前终止为何具有正则化的效果**

1. 验证集误差的学习曲线是一个 U 型曲线（考虑偏差方差，泛化误差，过拟合角度）

2. 提前终止可以将优化过程的参数空间限制在初始参数值 θ 0 的小邻域内

3. 提前终止相当于 L 2 正则化，证明如下

