

# 正则化方法

一叶知秋

# 预备知识

## 课前

- Coursera吴恩达 《机器学习》 WEEK 6

<https://www.coursera.org/learn/machine-learning>

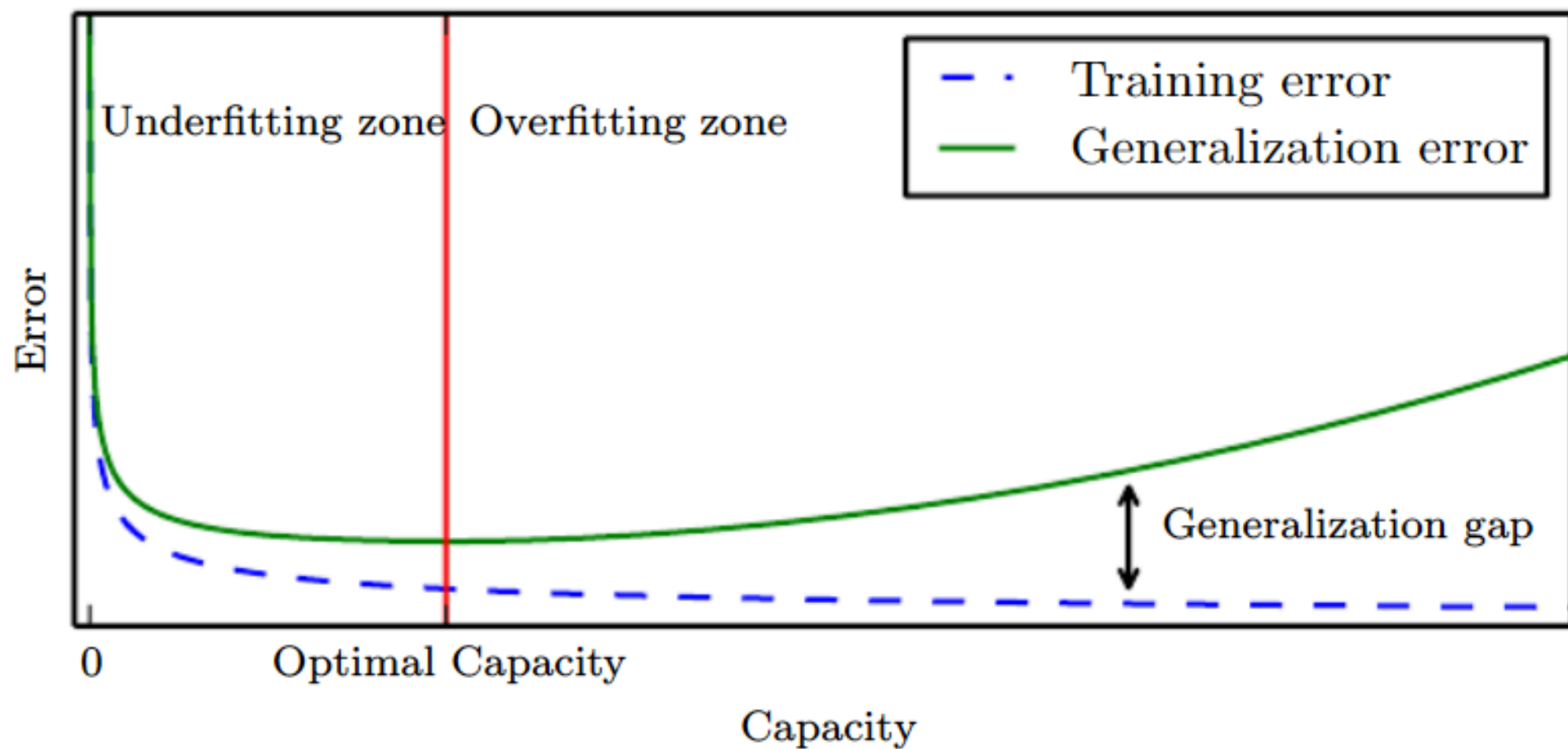
- 《数据挖掘导论》 4.4节

## 课后

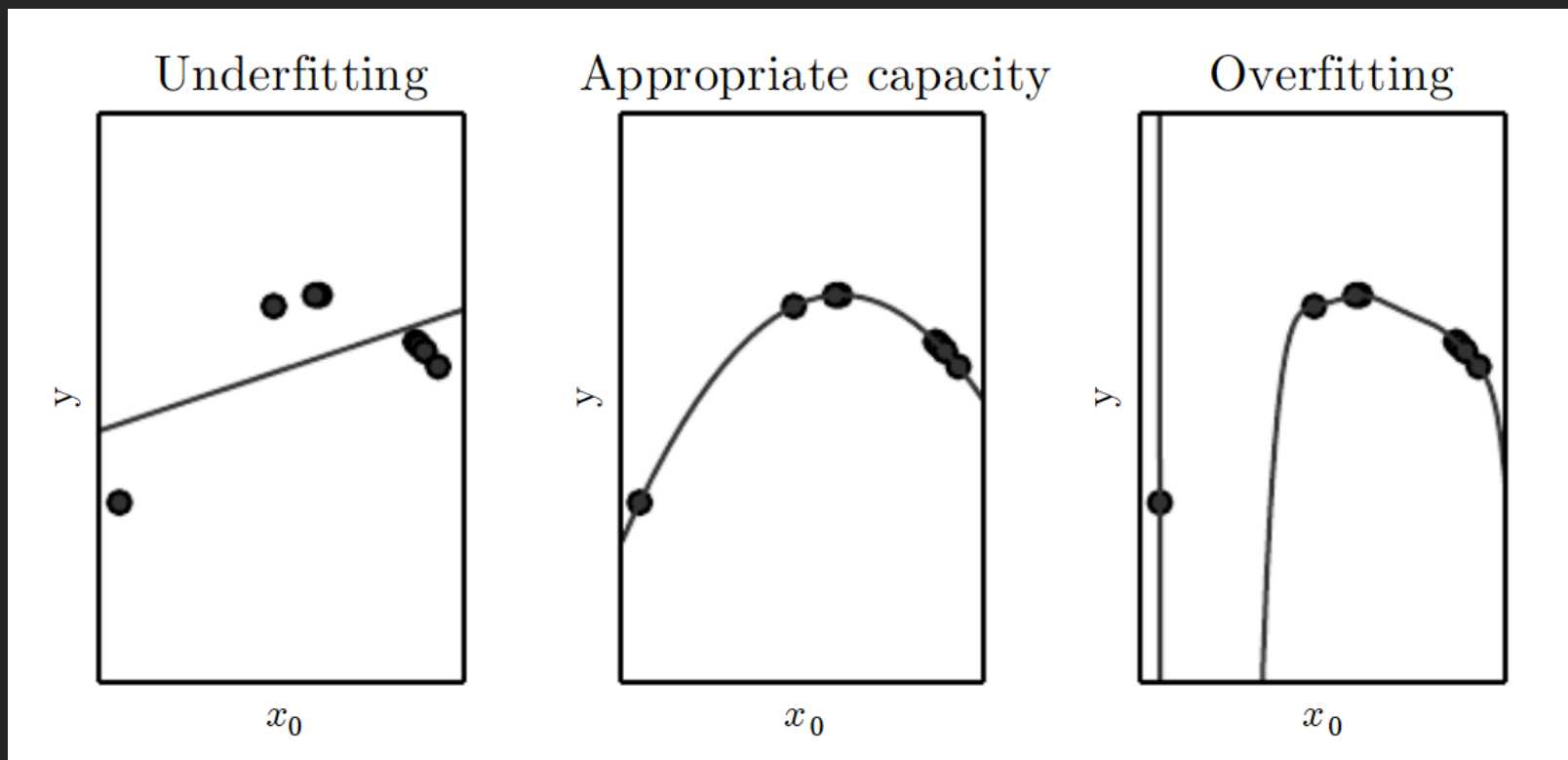
- 《Deep Learning》 第7章

<http://www.deeplearningbook.org/>

# 正则化的目的



# 模型的容量



奥卡姆剃刀定律 (Occam's Razor) : 如无必要, 勿增实体。

# 正则化方法

数据

模型

优化

# 更丰富的数据

数据

## 增加数据的数量和种类

- 获取更多的训练数据
- 人工合成更多的数据

数据的空间变换；人工数据合成；...

- 在训练数据中加入各种噪声

增加模型的反脆弱性

# 修改成本函数

优化

$$\tilde{J}(\theta; X, y) = J(\theta; X, y) + \alpha\Omega(\theta)$$

## L1-Norm (Lasso)

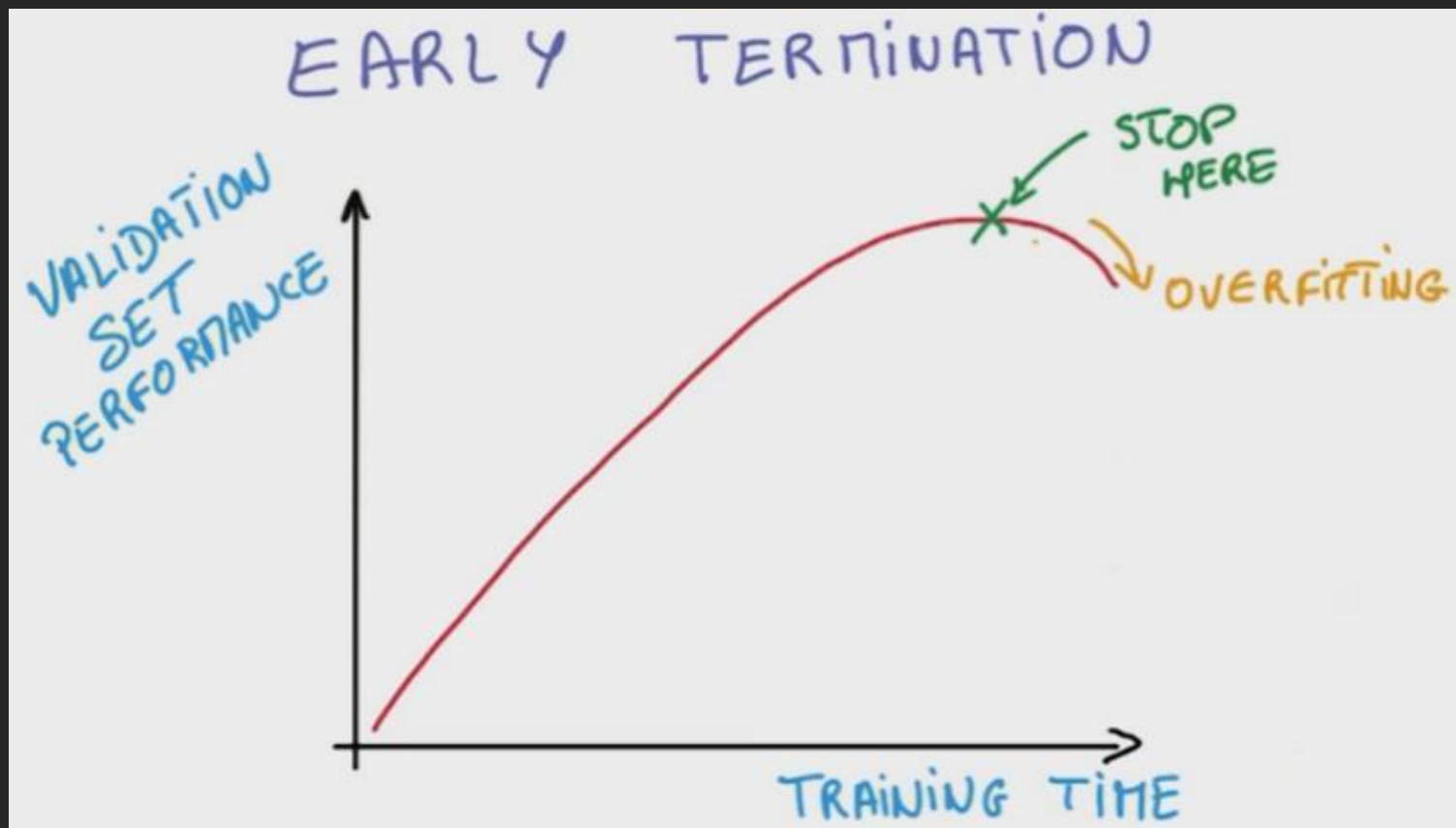
稀疏参数空间  
(少量非零参数)

## L2-Norm (Ridge)

限制参数范围  
(参数接近于零)

# Early-Stopping

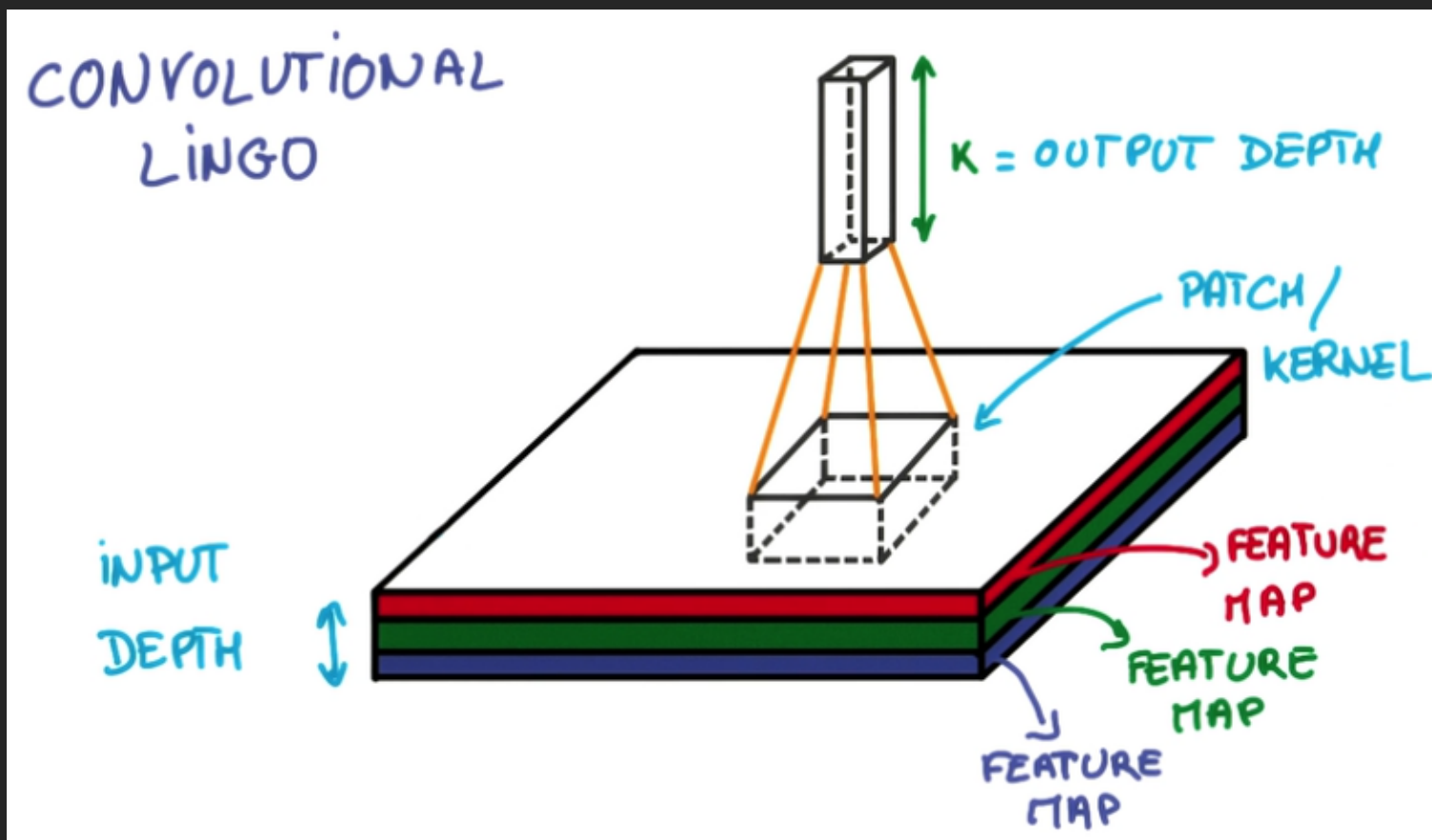
优化





# 参数共享

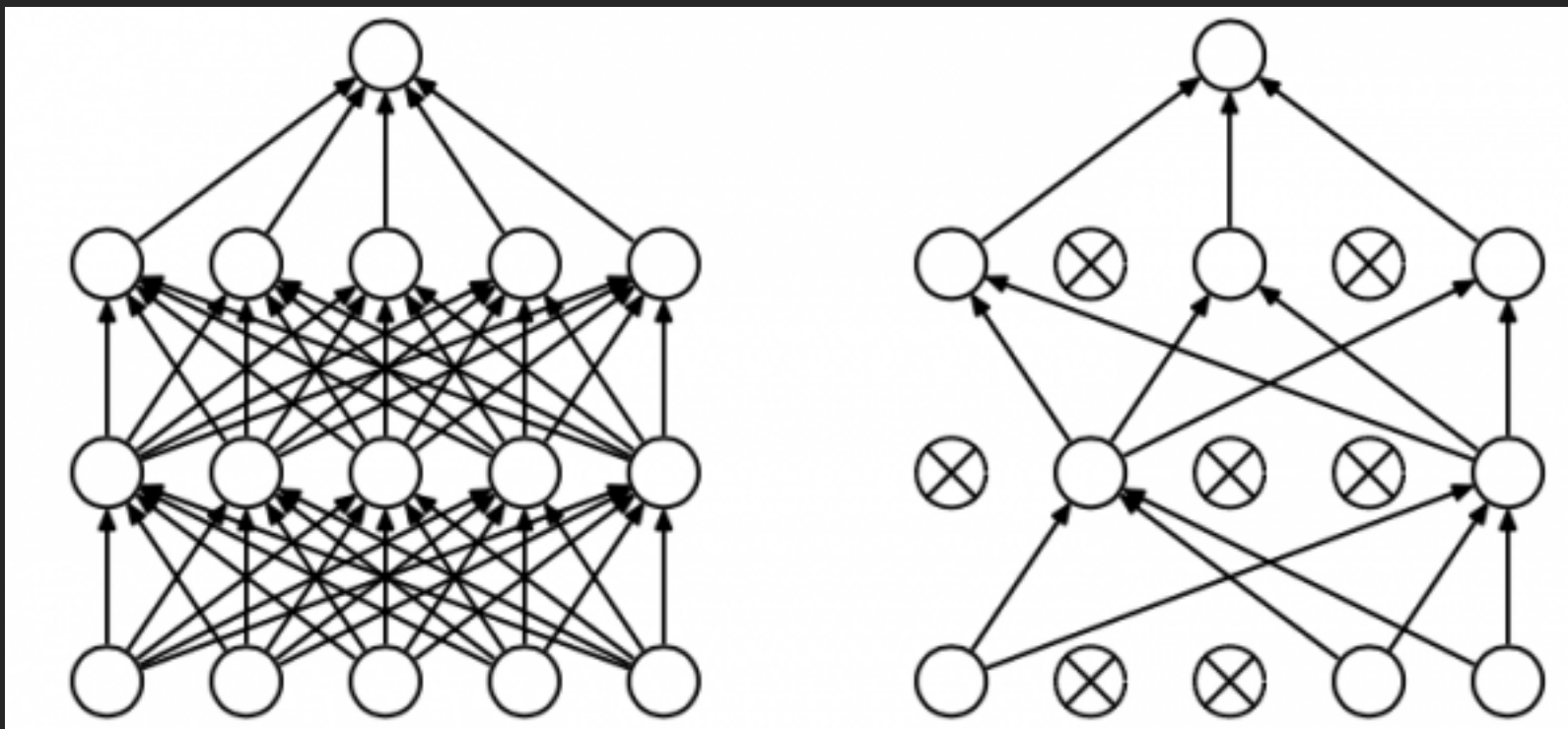
模型



卷积神经网络

# 系综方法

模型



Dropout

# 其他方法

- 多任务学习
- Batch Normalization
- 非监督/半监督学习
- 生成式对抗网络

...



实战经验

大模型 + 正则化

>

小模型

# Next Class – 卷积神经网络

## 参考教材

Deep Learning

<http://www.deeplearningbook.org/>

<https://github.com/HFTrader/DeepLearningBook>

