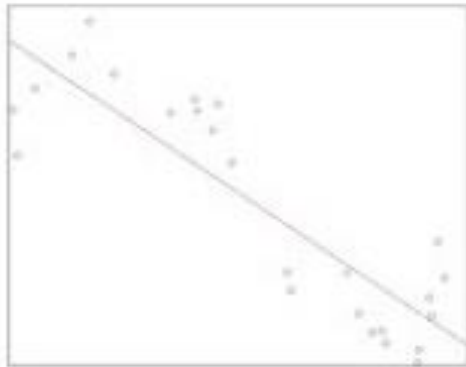
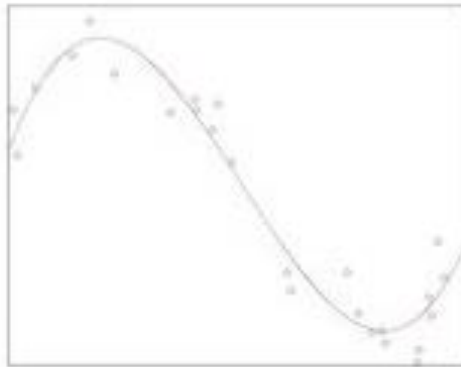


1. 모델 문제
2. Bias / Variance
3. Overfitting 해결방안
4. Regularization
5. Regularization 특성 평가

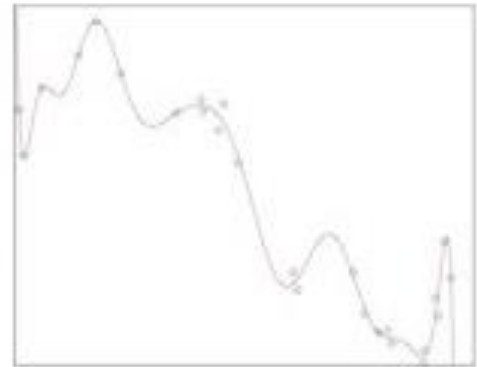
## 1. 모델 문제



underfit  
(degree = 1)



ideal fit  
(degree = 3)



overfit  
(degree = 20)

$$\rightarrow \theta_0 + \theta_1 x \quad \rightarrow \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 \quad \rightarrow \theta_0 + \theta_1 x + \dots + \theta_{20} x^{20}$$

$$\text{Error}(X) = \text{noise}(X) + \text{bias}(X) + \text{variance}(X)$$

(역자 주: noise는 데이터가 가지는 본질적인 한계치이기 때문에 irreducible error라고도 불리며, bias/variance는 모델에 따라 변하는 것이기에 reducible error라고 불린다)

## 2. Bias / Variance



**Bias**는 데이터 내에 있는 모든 정보를 고려하지 않음으로 인해, 지속적으로 잘못된 것들을 학습하는 경향을 말한다. underfitting 이라고도 한다.

반대로 **Variance**는 데이터 내에 있는 예러나 노이즈까지 잘 잡아내는 highly flexible models 에 데이터를 fitting시킴으로써, 실제 현상과 관계 없는 random한 것들까지 학습하는 알고리즘의 경향을 의미한다. 이는 overfitting과 관계되어 있다.

### 3. Overfitting 해결방안

- feature 를 줄일 수 있다
  - 수동으로 feature를 선택 (차원 축소)
  - Model Selection algorithm
- Regularization
  - 모든 feature를 유지하지만 얼마나 각 feature가 예측에 기여할지 변경
- More training data

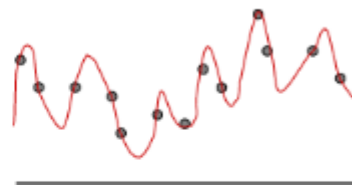
### Regularization

- The minimization

$$\min_f |Y_i - f(X_i)|^2$$

may be attained with zero errors.

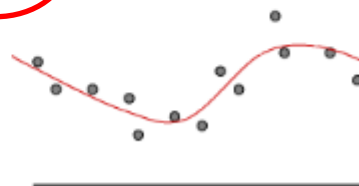
But the function may not be unique.



- Regularization

$$\min_{f \in H} \sum_{i=1}^n |Y_i - f(X_i)|^2 + \lambda \|f\|_H^2$$

- Regularization with smoothness penalty is preferred for uniqueness and smoothness.
- Link with some RKHS norm and smoothness is discussed in Sec. IV.



II-26

**“ Let’s not have too big numbers in the weight!!”**

1. 왜 Regularization 해야 할까? (정규화의 궁극적인 목표)
2. W가 작아지도록 학습한다는 것은 무슨 의미?

## 4. Regularization

$$\lambda \|f\|_H^2$$

Regularization strength

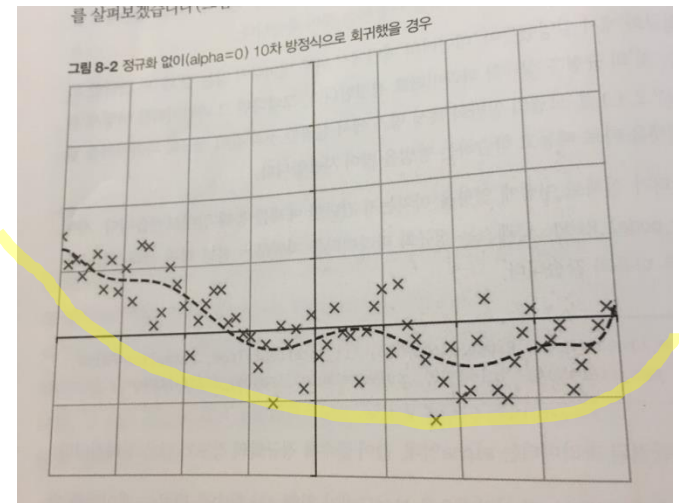
- Model 의 bias 와 variance 사이의 균형을 조절
- Regularization strength : 제약을 얼마나 강하게 걸지 결정해주는 값

Ex)

원래 데이터 패턴 :  $y = x^2 - 2x - 1$

모델 :  $h = w_0 + w_1x + w_2x^2 + \dots + w_{10}x^{10}$

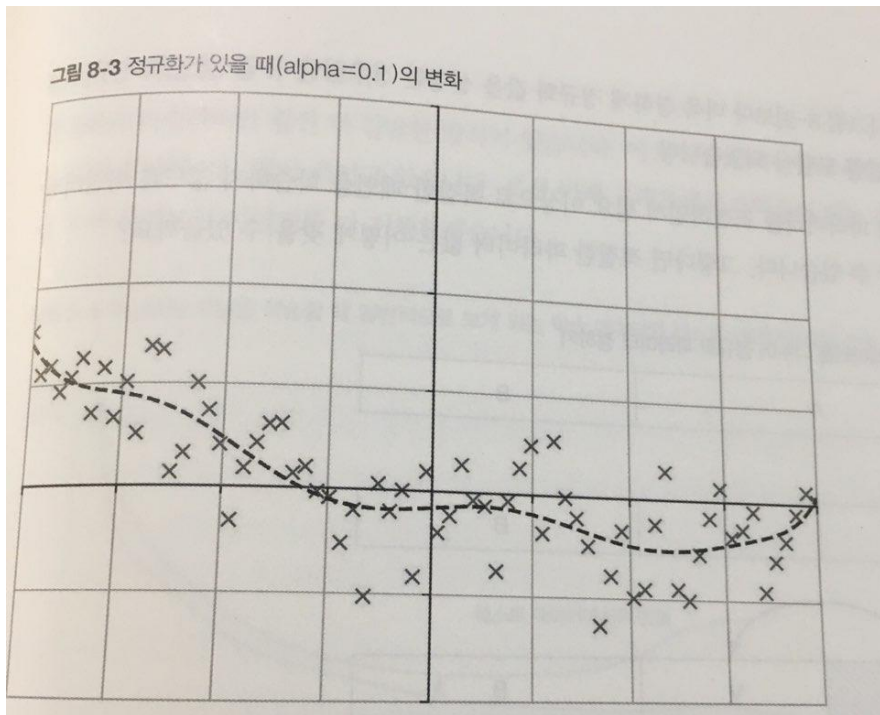
➡ **Overfitting!!**



정규화가 없을 때:  $\lambda = 0$

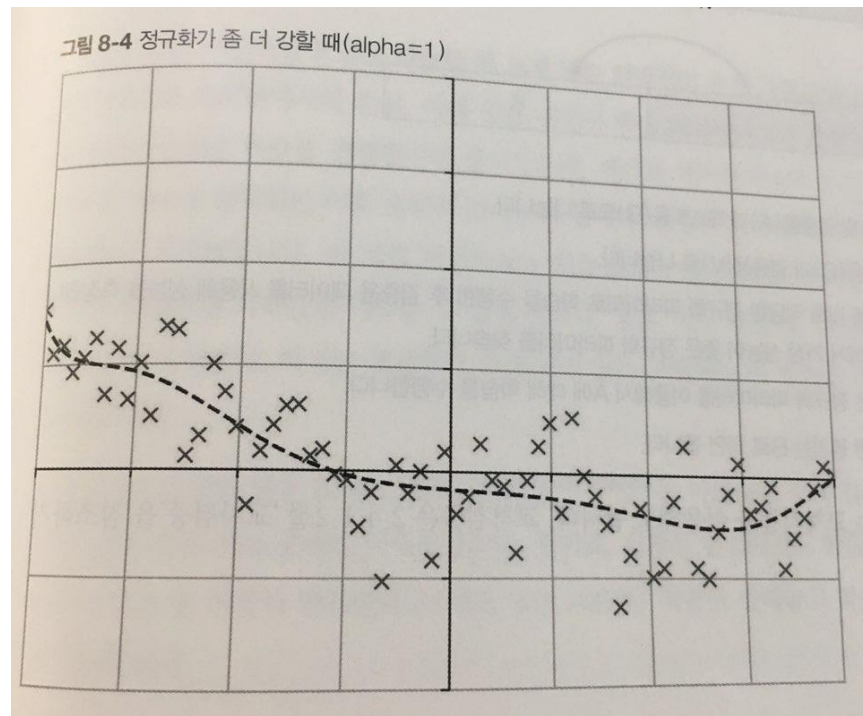
## 4. Regularization

그림 8-3 정규화가 있을 때( $\alpha=0.1$ )의 변화



정규화가 있을 때:  $\lambda = 0.1$

그림 8-4 정규화가 좀 더 강할 때( $\alpha=1$ )

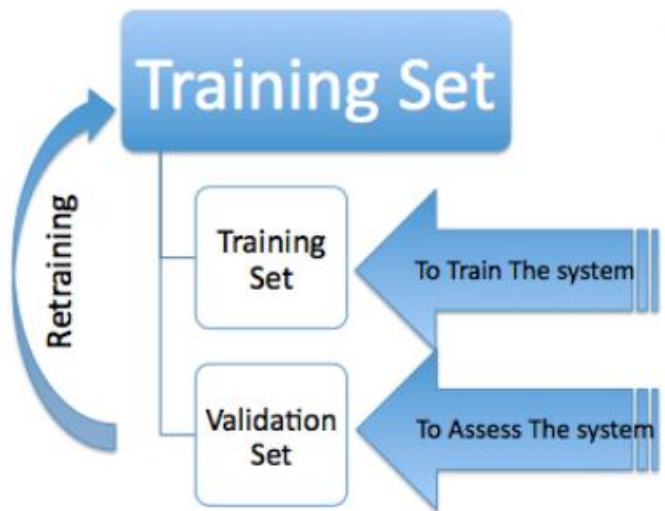


정규화가 강할 때:  $\lambda = 1$

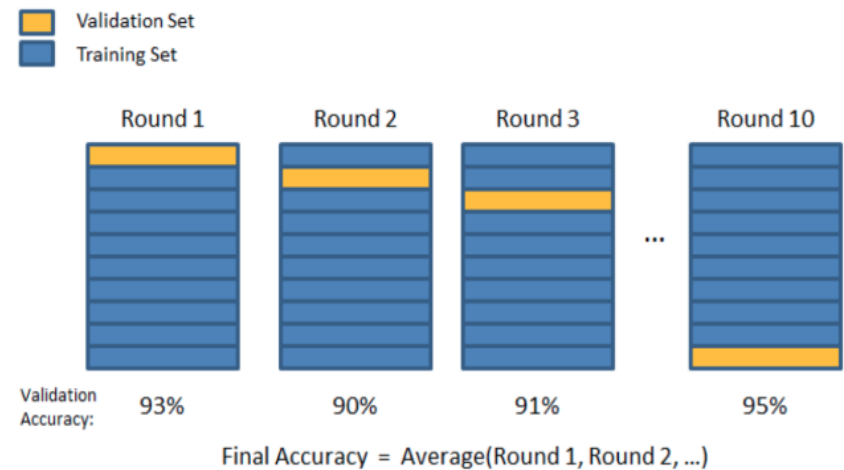
참고도서) 처음 배우는 머신러닝 - 한빛미디어

## 5. Regularization 특성 평가

- Train set & Validation set



- Cross- Validation (CV)





1. 쉽게 이해해보는 bias-variance tradeoff

<http://bywords.tistory.com/entry/%EB%B2%88%EC%97%AD-%EC%9C%A0%EC%B9%98%EC%9B%90%EC%83%9D%EB%8F%84-%EC%9D%B4%ED%95%B4%ED%95%A0-%EC%88%98-%EC%9E%88%EB%8A%94-biasvariance-tradeoff>

2. bias-variance decomposition

<https://ratsgo.github.io/machine%20learning/2017/05/19/biasvar/>

3. Overfitting

<http://sanghyukchun.github.io/59/>