

Overfitting & Underfitting

- 일반화: 모델이 학습 후에 새로운 데이터에 대해 정확하게 예측할 수 있는 능력

1) 과소적합: '모델의 용량이 너무 작아' or '훈련 집합이 너무 작아' 오류가 클수록 많은 경우

- 방지: 비선형 모델과 같이 용량이 큰 모델 사용

충분한 훈련 집합 사용

2) 과적합: 훈련 집합에 대해 거의 완벽하게 근사화 함.

→ 용량이 너무 크기 때문에 학습과정에서 잡음까지 수용

→ 데이터가 모여있지 않고 특성상성 있을 때 (~~잡음~~ 잡음이 많다)
위험하다.

- 방지: 충분히 많은 학습데이터 사용

- 모델의 용량이 커질수록 가중치 값도 커진다.

3) 과소·과대 적합 해결법

1. 검증집합을 이용한 모델 선택

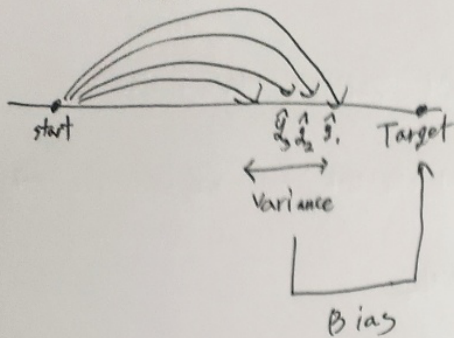
- 검증 집합에 대해 최고 성능을 갖는 모델 선택

2. 적당한 용량의 모델 선택

- 최소한의 잡음을 수용하는 용량

3. 규제 λ : ~~과~~ 과적합을 막기 위해 λ parameter 제한

Bias - Variance



Variance : 데이터들이 퍼져있는 정도

Bias : 점들과 예측한 결과가 정답과 차이가 나는 정도

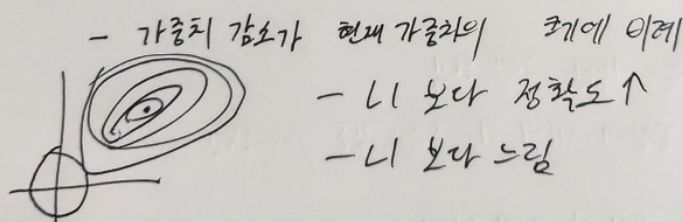
Regularization (규제) : 일반화 오류를 줄이려는 의도를 가지고 학습 알고리즘을 수정하는 방법

- 명시적 규제 : 목적함수나 신경망 수정 (직접적)
- 암시적 규제 : 과적합을, 데이터 증가, 등 간접적으로 영향을 미침.

1) 목적함수 규제

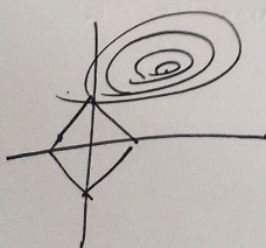
- 규제항 ~~항~~ : 가중치 감소 : 용량을 크게 하고 수치적 용량을 제한
- 큰가중치를 작은가중치로 유지 → 최용량을 원점으로 당김
- ~~L1 norm~~ L1 norm 방식, L2 norm 방식

① L2 norm 방식



- 가중치 감소가 현재 가중치의 제곱에 비례
- L1 보다 정확도 ↑
- L1 보다 느림

② L1 norm 방식



- 가중치 감소가 고정값만큼 줄인 후 업데이트
- 가중치가 0이 되는 효과가 발생
- L2 norm 보다 빠르다.

