CUAI 스터디_MTs

2022.04.04

발표자 : 정 승 욱

스터디 모임 인증

Ch 04 모델 훈련
회귀 Regression

1. 선형 회귀

2. 경사 하강법

3. 다항 회귀

4. 학습 곡선

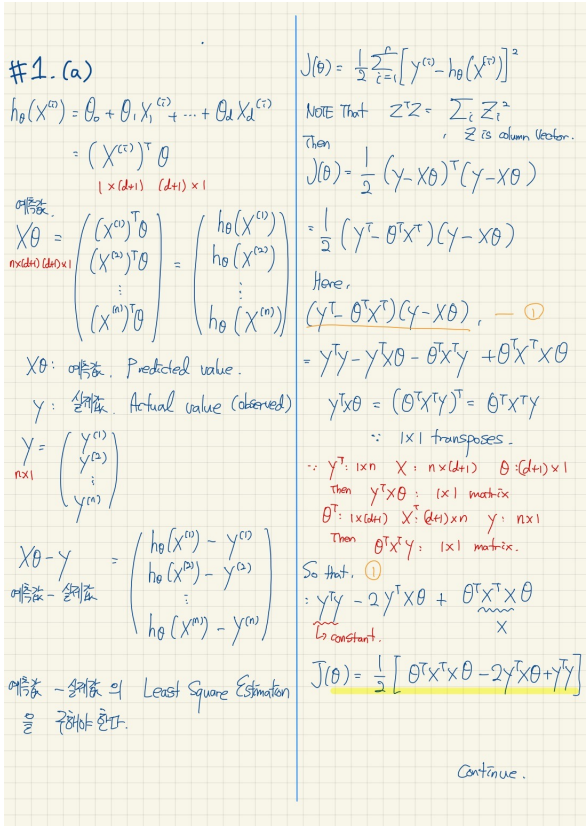
5. 규제 : 릿지, 라쏘, 엘라스틱 넷

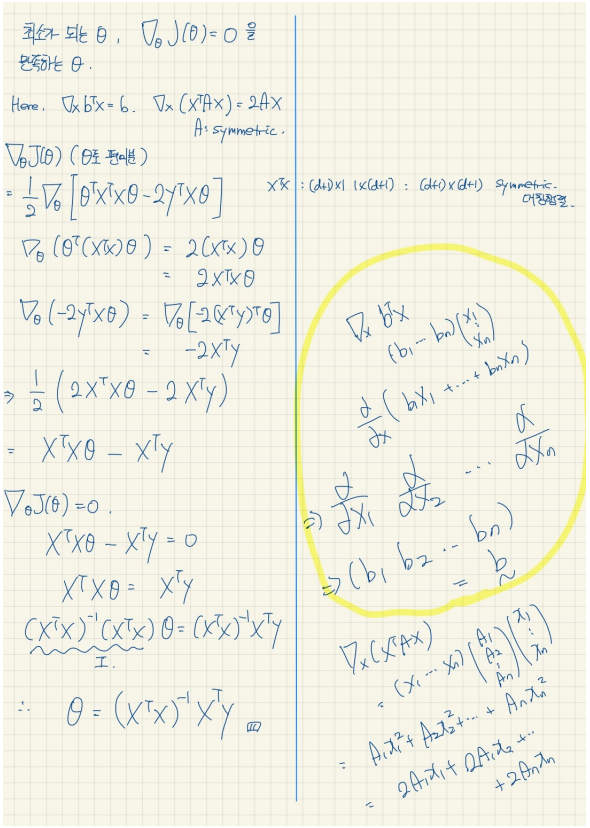
6. 로지스틱 회귀

정규 방정식, 계산복잡도

$$\hat{\theta} = (X^T X)^{-1} X^T y$$

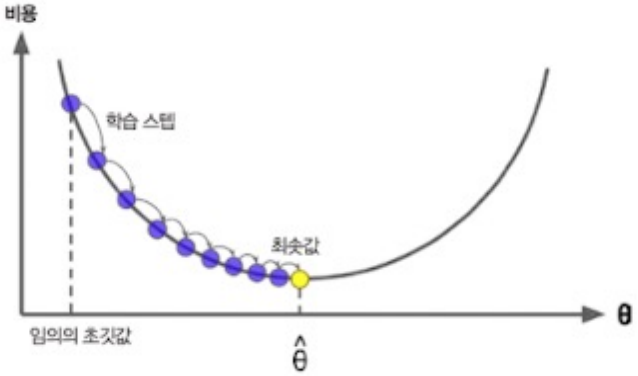
장점: 한번에 최적 계수가 나옴! But 비용 높음!





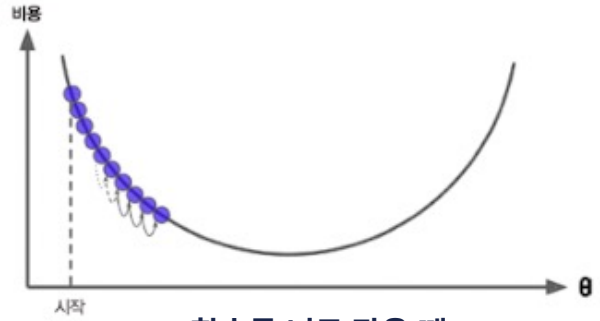
https://www.youtube.com/watch?v=sDv4f4s2SB8

경사 하강법



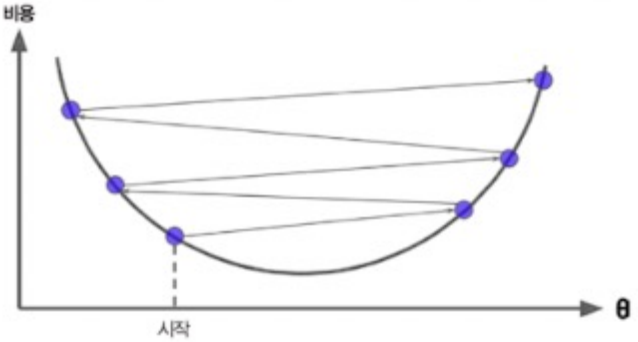
$$\theta^{(nextstep)} = \theta - \eta \nabla_\theta MSE(\theta)$$

$$\eta = Learning\ Rate$$



학습률 너무 작을 때

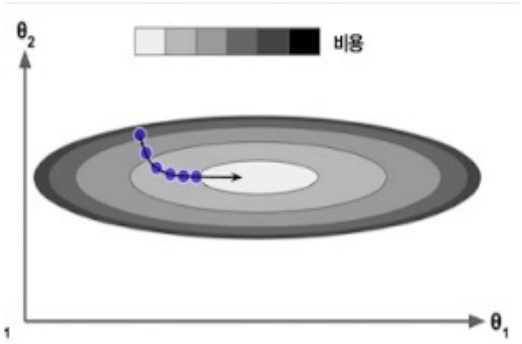
:

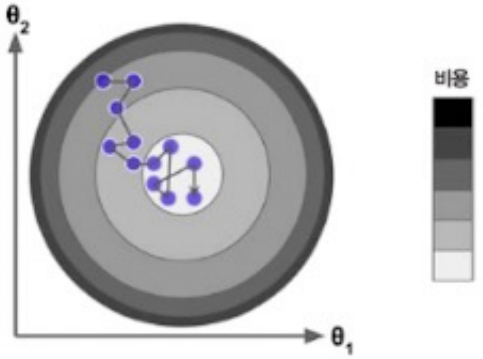


학습률 너무 클 때

배치 경사 하강법 &
확률적 경사 하강법
& 미니 배치 경사 하강법

배치 : 매 스텝 모든 샘플
확률 : 매 스텝 1개 샘플
미니 배치 : 매 스텝 작은 랜덤 샘플





https://www.youtube.com/watch?v=sDv4f4s2SB8
https://www.youtube.com/watch?v=vMh0zPT0tLI

Epoch & Batch & Iteration

1 Epoch : 모든 데이터 셋을 한 번 학습
1 iteration : 1회 학습

minibatch : 데이터 셋을 batch size 크기로 쪼개서 학습

ex) 총 데이터가 100개, batch size가 10이면,
1 iteration = 10개 데이터에 대해서 학습
1 Epoch = 100/batch size = 10 iteration

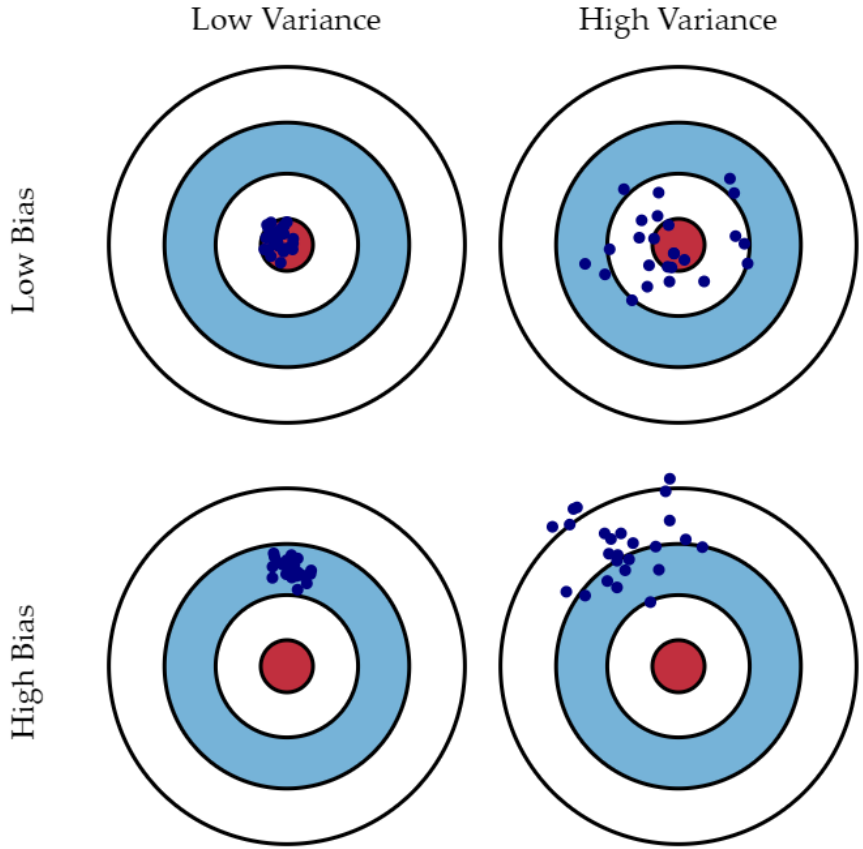
https://www.slideshare.net/w0ong/ss-82372826

분산과 편향

편향: 잘못된 가정
분산: 작은 변동에 모델이 민감

데이터 안의 진동: 엡실론으로 표기!

편향 종류!! bias vs variance & `ϵ`인지!



규제 : 릿지, 라쏘, 엘라스틱넷

-> 가중치에 제한, 과대적합 방지

$$J(\boldsymbol{\theta}) = \mathrm{MSE}(\boldsymbol{\theta}) + \alpha\frac{1}{2}\sum_{i=1}^{n}\theta_i^2$$

릿지: 덜 중요한 특성의 가중치 낮춤 l2 규제!

$$J(\boldsymbol{\theta}) = \mathrm{MSE}(\boldsymbol{\theta}) + \alpha\sum_{i=1}^{n}|\theta_i|$$

라쏘: 덜 중요한 특성의 가중치 삭제 l1 규제!

$$J(\boldsymbol{\theta}) = \mathrm{MSE}(\boldsymbol{\theta}) + r\alpha\sum_{i=1}^{n}|\theta_i| + \frac{1-r}{2}\alpha\sum_{i=1}^{n}\theta_i^2$$

릿지와 라쏘의 절충: 엘라스틱 넷

혼합 비율 r, 1-r
- 특성 수가 훈련 샘플 수 보다 많거나
- 특성 몇 개가 강하게 연관 되어있을 때 라쏘가 문제를 일으킴

로지스틱 회귀:
범주형 자료 회귀, 분류 모델

이진 분류 하는 문제! 1 또는 0

시그모이드 함수를 이용

$$log(Odds(p)) = wx + b$$

오즈 Odds

실패 비율 대비 성공 비율

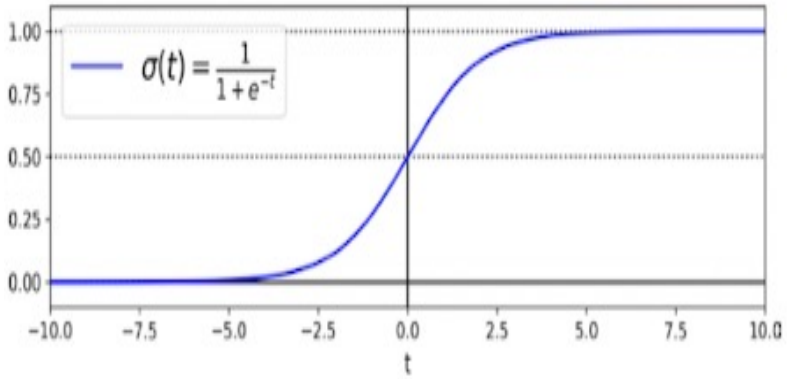
$$\frac{p}{1-p}$$

로짓 Logit Log+Odds

오즈에 자연로그!

$$L = ln\frac{p}{1-p}$$

$$p = \frac{e^{-L}}{e^{-L}+1}$$



https://www.youtube.com/watch?v=yIYKR4sgzI8

감사합니다!