HW W4

요약

머신러닝 모델 훈련의 목표는 예측 성능을 최대화 하는 것이고, 이 성능을 비용함수를 이용하여 계산한다. 회귀 모델에서 일반적으로 MSE를 비용함수로 사용한다.

경사하강법 – 임의의 값에서 시작, MSE의 최소점을 찾아 스텝 반복

학습률이 너무 적은 경우 수렴속도 느림, 너무 큰 경우 발산

경사하강법의 종류

배치 경사하강법 -> mb = m, 에포크마다 1스텝, 느리고 메모리 차지 많아 사용 잘 안함

확률적 경사하강법 -> mb = 1, 데이터마다 편차가 크므로 파라미터의 동요가 심함

두 방법의 타협점으로 미니배치 경사하강법을 이용한다.

다항 회귀는 선형 회귀 모델로 치환할 수 있다

Train – valid – test

편향 – 데이터셋과 맞지 않는 모델링 -> 과소적합

분산 - 모델이 훈련 데이터에 민감하게 반응하는 정도 -> 과대적합

분산을 잡기 위한 모델 규제

릿지 회귀, 라쏘 회귀 – mse에 파라미터 값이 큰 정도를 더하여 비용함수를 정함

릿지, 라쏘 차이??

두 모델 절충 – 엘라스틱 넷

조기 종료 – 훈련 중 검증셋에 대한 비용함수의 값이 다시 커지는 순간(= 과대 적합, 진동이 아님이 확인될 때) 훈련 종료, 기억해둔 최적의 모델 사용

규제를 적용하면 분산이 줄고 편향이 는다.

수백만 개의 특성을 가진 훈련 세트에서는 어떤 선형 회귀 알고리즘을 사용할 수 있는가? – 경사하강법