

Chap4. 훈련모델

4.1)

유사역행렬을 구할 수 있는 함수는 무엇인가?

`np.linalg.pinv()`

4.2)

경사하강법(GD)에서 학습률이 너무 클 때와 너무 작을 때의 문제점을 각각 쓰세요.

학습률이 클 때 : 골짜기를 가로질러 반대편으로 건너뛰게 이전보다 더 높은 곳으로 올라가게 되어 알고리즘이 더 큰 값으로 발산하게 만들어 적절한 해법을 찾지 못할 수 있다.

학습률이 작을 때 : 알고리즘이 수렴하기 위해 반복을 많이 진행해야 하므로 시간이 오래 걸린다.

4.3)

비선형 데이터를 선형 모델을 사용하여 학습시키는 방법은?

각 특성의 거듭제곱을 새로운 특성으로 추가하고, 이 확장된 특성을 포함한 데이터셋에 선형 모델을 훈련시킨다.

4.4)

모델이 데이터에 과대적합 또는 과소적합 되었는 지 파악 할 수 있는 방법 2가지를 쓰세요.

1. 교차검증을 사용. 훈련 데이터에서 성능이 좋지만 교차 검증 점수가 나쁘다면 모델이 과대적합 된 것이다.
2. 훈련 세트와 검정 세트의 모델 성능을 훈련 세트 크기의 함수로 나타낸 학습 곡선을 살펴본다.

4.5)

릿지회귀, 라쏘회귀, 엘라스틱넷의 차이점을 서술하세요.

규제항의 차이가 있다.

릿지회귀는 규제가 포함된 선형 회귀 버전이다. 규제항 $\alpha \sum_{i=1}^n \theta_i^2$ 이 비용함수로 추가된다,

라쏘회귀는 선형 회귀의 또 다른 규제된 버전으로 가중치 벡터 \mathbf{l}_1 노름을 사용한다.

엘라스틱넷은 릿지회귀와 라쏘회귀를 절충한 모델이다. 규제항은 릿지와 회귀의 규제항을 단순히 더해서 사용하며, 혼합 정도는 혼합 비율 r 을 사용해 조절한다. $r=0$ 이면 엘라스틱넷은 릿지회귀와 같고, $r=1$ 이면 라쏘회귀와 같다.

4.6)

로지스틱 회귀의 비용함수(로그 손실)에서 최솟값을 계산하는 방법은?

비용함수가 볼록 함수이므로 경사 하강법을 사용한다.