단순 선형 회귀를 위한 일반화된 예측 함수(wave 데이터셋)

https://t1.daumcdn.net/cfile/tistory/999093335A04FD5F09

X[0]은 **특성(feature)**, https://t1.daumcdn.net/cfile/tistory/999F76335A05002A14은 **예측 값**, b는 **편향(bias)**, w[0]은 **가중치(weight) 또는 계수(coefficient)** 라고 부름.

여러 개의 특성을 활용한 회귀모델(다중 선형 회귀)

목표 : 데이터데이터셋들의 특성들(x[0], x[1] … x[p]) 과 라벨값(y) 사이의 관계를 잘 설명해낼 수 있는 적합한 특성 가중치(w[0], w[1] … w[p]) 와 b(편향) 을 찾는 것

어떻게 적절한 가중치와 편향을 찾을까..?

선형 회귀는 라벨 값(y)과 예측값(https://t1.daumcdn.net/cfile/tistory/999F76335A05002A14) 사이의 평균제곱오차(mean squared error, MSE)를 최소화하는 파라미터 w와 b를 찾는다.

(실제 라벨값과 예측값이 작으면 작을수록 예측성능이 좋은 것 이기 때문)

https://t1.daumcdn.net/cfile/tistory/997372335A0520F312

https://t1.daumcdn.net/cfile/tistory/99E3AF335A05331227

MSE를 최소화하는 파라미터 w와 b를 찾는 것이 선형회귀의 목적

수식을 대입하면 w에 관한 식이 하나 생김.

우리는 w의 2차원 수식이 생길텐데, 경사하강법(gradient descent)를 사용하여 w의 최소값을 찾음

선형 회귀 모델은 과대적합(overfitting)될 때가 종종 있음

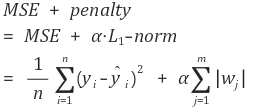
* 샘플의 특성값들과 라벨값의 관계를 너무 자세하게 분석하기 때문에

그래서 우리는 가중치에 대한 규제(regularization)가 필요함

릿지(Ridge) => L2 규제

라쏘(Lasso) => L1 규제

L1-norm 패널티를 가진 선형 회귀 방법



는 패널티의 효과를 조절해주는 파라미터



https://t1.daumcdn.net/cfile/tistory/99602D335A053A9908

**MSE + penalty 를 최소로 만드는 w, b 찾기**

MSE 항이 작아질수록, 라벨값들과 예측값들의 차이가 작아짐

L1-norm이 작아질수록많은 가중치들이 0이 되거나 0에 가까워짐

알파가 너무 작으면 과대적합(복잡도가 큼)

알파가 너무 크면 과소적합(복잡도가 너무 작음)