

SUB_PJT_01 (AI)

서울 1반 6팀 이민아

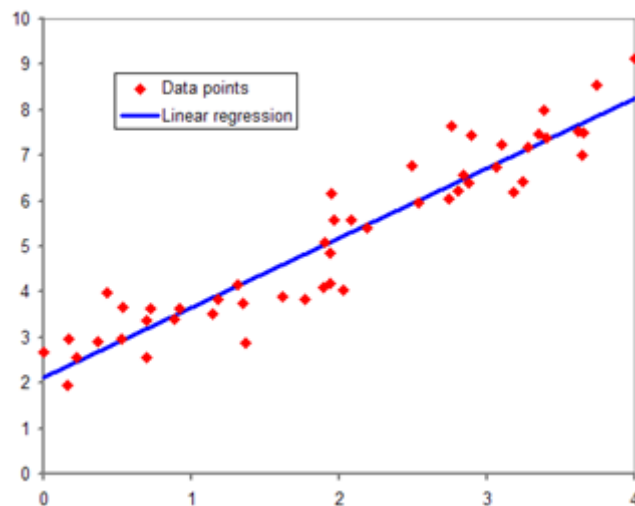
- 사전학습
 - [사전학습1 인공지능](#)
 - [사전학습2 회귀 및 경사하강법](#)
 - [사전학습3 신경망](#)
 - [사전학습4 파이썬 라이브러리](#)

사전학습2 회귀 및 경사하강법

1. 회귀

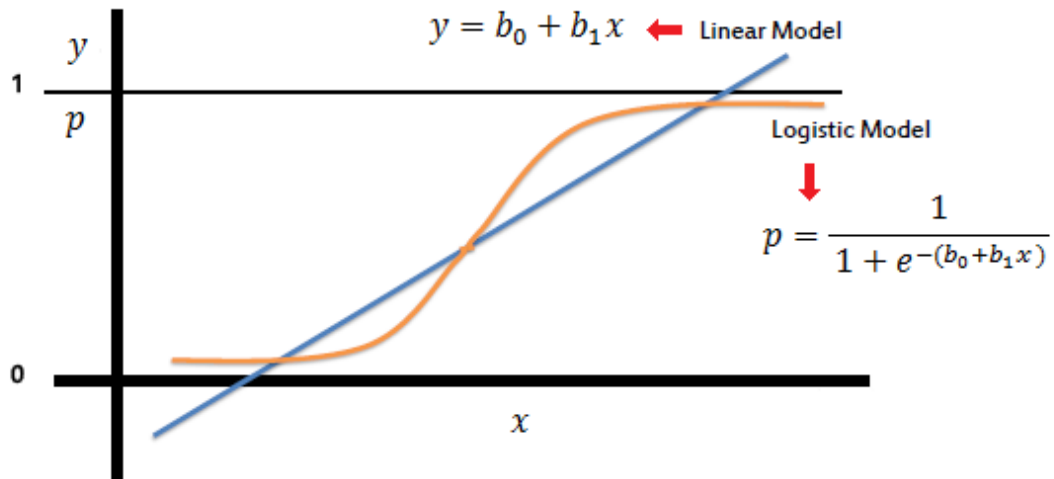
(1) 선형 회귀 (Linear Regression)

- 종속 변수 y 와 한개 이상의 독립 변수 x 와의 **선형 상관** 관계를 모델링하는 회귀분석 기법
- 독립 변수 ' x '를 사용하여 종속 변수 ' y '의 움직임 예측 및 설명
- **1차원** 방정식을 찾는 과정
- **활성화 함수**는 주로 **계단 함수** 사용



(2) 로지스틱 회귀 (Logistic Regression)

- 선형회귀에서 구하는 직선 대신 **S자 곡선**을 이용하여 분류의 정확도를 향상한 회귀 방법
- 선형 회귀를 사용하여 데이터를 0과 1 사이의 값으로 분류하는 알고리즘
- 독립변수에 따른 종속변수를 추정하는데 사용되나, 종속변수가 **이진**인 형태
- **분류**에 이용
- **활성화 함수**는 주로 Sigmoid, ReLU, Softmax 함수 사용



2. 경사하강법 (Gradient Descent Algorithm)

(1) 정의

- 손실 함수가 정의되었을 때, **손실 함수의 값이 최소화** 되는 지점
- **MSE(평균제곱오차/손실함수)**를 가중치(weight)에 대한 **미분 값이 감소**하는 방향으로 **가중치를 업데이트**하여 손실 함수의 최소값을 찾음

(2) 배치 (Batch)

- **Total Training Dataset**
- 단일 반복에서 기울기를 계산하는 데 사용하는 예의 총 개수
- **전체 데이터 셋에 대해 에러를 구한 뒤 기울기를 한번만 계산하여 모델의 파라미터를 업데이트**하는 방식
- 배치가 너무 커지면 단일 반복으로 계산하는데 오랜 시간 걸림
- 유형
 - BGD
 - SGD
 - MSGD

(3) Batch Gradient Descent (BGD)

- 전체 데이터에 대해서 업데이트가 한번에 이루어지므로 **업데이트 횟수 자체는 적다**
- 한번의 업데이트에 모든 Training Data Set을 사용하므로 **계산 자체는 오래 걸린다**
- Error Gradient 계산하므로 **Optimal로 수렴이 안정적**
- **Local Optimal에 수렴**될 경우 탈출하기 어려울 수 있다
- GPU를 활용한 병렬처리에 유리
- 전체 학습데이터에 대해 한번에 처리를 해야하므로 **많은 메모리가 필요**

(4) Stochastic Gradient Descent (SGD)

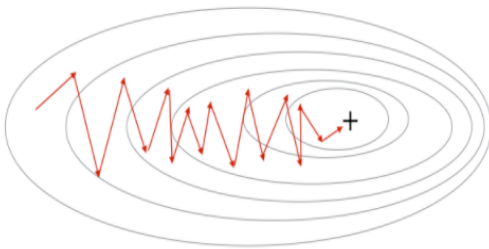
- 파라미터를 업데이트할 때 무작위로 샘플링된 학습 데이터를 하나씩만 사용하여 추출된 각각 **한개의 데이터**에 대해서 cost function(손실함수)의 Gradient 계산 후 업데이트 하는 알고리즘
- 확률적이기에 불안정하여 아래와 같이 **요동치**는 것을 볼 수 있으며 **자주 업데이트한다**
- 성능 개선 정도를 빠르게 확인

- 최소 cost에 수렴했는지의 판단이 상대적으로 어려움
- **Shooting**이 발생하므로 **Local Optimal** 빠질 리스크가 다소 적어진다
- 훈련 데이터 한개씩 처리하므로, GPU 성능에 대해서 **전부 활용이 불가하다**
- 때로는 **Global Optimal** 찾지 못할 수도 있다

(5) Mini-Batch Stochastic Gradient Descent (SGD, MSGD)

- **BGD와 SGD의 혼합**으로 SGD의 노이즈를 줄이면서 전체 배치보다 효율적이어서 널리 사용
- 전체 데이터셋에서 **m개의 데이터를 뽑아 mini-batch**를 구성
- 이것의 **평균 기울기**를 통해 모델을 업데이트
- BGD보다 (슈팅이 어느정도 발생하므로) **local optimal** 빠질 리스크 적다
- SGD보다는 GPU 성능을 활용한 **병렬처리가 가능하여 속도면에서 유리**
- 전체 데이터 아닌 일부 데이터만 메모리에 적재하여 사용하므로, 메모리 사용측면에서 BGD보다 원할

Stochastic Gradient Descent



Mini-Batch Gradient Descent

