

과제명 : 지식 그래프 기반 음식점 광고 콘텐츠
(Title) 추천 알고리즘 개발

과제번호 :
(Project No.)

3) 선형회귀 (Linear Regression)

: 한개 이상의 독립변수 x 와 y 의 선형 관계를 모델링.

-1. 선형회귀

(1) 단순 선형회귀 (독립변수가 1개)

$$y = wx + b \quad (w: \text{가중치}, b: \text{편향})$$

(2) 다중 선형회귀 분석

$$y = w_1x_1 + w_2x_2 + \dots + w_nx_n + b$$

-2. 가설 세우기 (Hypothesis)

x 와 y 의 관계 유추를 위해 수학적으로 식을 세운 것.

-3. 비용함수 (Cost function): 평균 제곱 오차 (MSE)

함수의 값을 최소화하거나 최소화하면 목적함수.

최소화가 목적이면 비용함수, 손실 함수

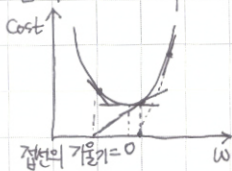
* 평균 제곱 오차

$$\left(\sum_{i=1}^n [y^{(i)} - H(x^{(i)})]^2 \right) \times \frac{1}{n} = \text{Cost}(w, b)$$

음수 오차와 양수 오차 모두 있을 수 있어 제곱한다.

비용함수는 MSE를 줄이는 것이 목적이다.

-4. 옵티마이저 (Optimizer): 경사하강법 (Gradient Descent)



Cost를 최소화하는 w 를 구하는 식

$$w := w - \alpha \frac{d}{dw} \text{Cost}(w) \quad (\alpha: \text{학습률}, 0 < \alpha < 1)$$

α 의 값이 너무 크면 발산할 수 있고

너무 작으면 수렴시간이 너무 오래걸린다.

4) 자동 미분

자동 미분 (TensorFlow)과 케라스를 통해 선형 회귀를 할 수 있다.

5) 로지스틱 회귀 (Logistic Regression)

-1. 이진 분류 (Binary Classification)

실제 값이 0과 1 두가지 값이므로 연속 값도

0과 1이 나오도록 함수를 규정한다.

-2. 시그모이드 함수 (Sigmoid function)

$$H(x) = \frac{1}{1 + e^{-(wx+b)}} = \text{Sigmoid}(wx+b) = \sigma(wx+b)$$

-3. 비용함수

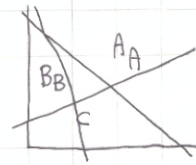
로지스틱 회귀는 MSE를 사용했을 때 로컬 미니멀에 빠질 수 있어서 새로운 함수를 만들어야 한다.

$$J(w) = -\frac{1}{n} \sum_{i=1}^n [y^{(i)} \log H(x^{(i)}) + (1-y^{(i)}) \log (1-H(x^{(i)}))]$$

로지스틱 회귀에서 찾아낸 비용함수

크로스 엔트로피 (Cross Entropy) 함수라 한다.

6) 소프트 맥스 회귀 (Softmax Regression)



$$X \rightarrow \square \xrightarrow{w} \square \rightarrow \bar{y}$$

$$X \rightarrow \square \xrightarrow{w} \square \rightarrow \bar{y}$$

$$X \rightarrow \square \xrightarrow{w} \square \rightarrow \bar{y}$$

하나하나 Sigmoid 함수를 써야 한다.

하지만 Softmax 함수를 이용하면

값이 0보다 크고 1보다 작고,

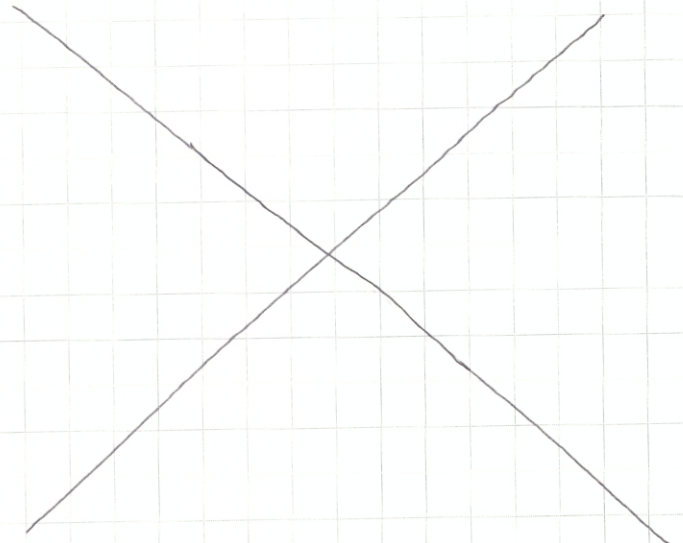
합이 1이 된다.

$$\text{Softmax}(z) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad \text{for } i = 1, 2, \dots, K$$

Cost 함수:

$$\text{Cost} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^K y_j^{(i)} \log(p_j^{(i)})$$

$$= -\frac{1}{n} \sum_{i=1}^n [y^{(i)} \log(p^{(i)}) + (1-y^{(i)}) \log(1-p^{(i)})]$$



기록자
(Recorded by)

이상준

서명
(Signature)

이상준

일자
(Date)

2022. 4. 9

점검자
(Reviewed by)

정선태

서명
(Signature)

정선태

일자
(Date)

2022. 4. 12

이 노트의 지적재산권은 숭실대학교 산학협력단에 있음
(Intellectual property right for this notebook belongs to Foundation of Soongsil)