

기계학습원론 중간고사  
2024 가을

강의자: 조재민  
강의조교: 이재웅, 정명원

October 18, 2024

**문제 1.** 길이가 1인 2차원 벡터  $x$ 와 양의 정부호(positive-definite)  $2 \times 2$  행렬  $A$ 에 대해  $f(x) = x^T A x$ 를 최소화 하는  $x$ 의 조건과 그 때의 최솟값을  $A$ 를 활용하여 설명하여라. 완결된 식이 아니더라도 좋다. 그리고 이 결과를 이용하여  $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 에 대해  $f(x)$ 의 최솟값을 구하여라.

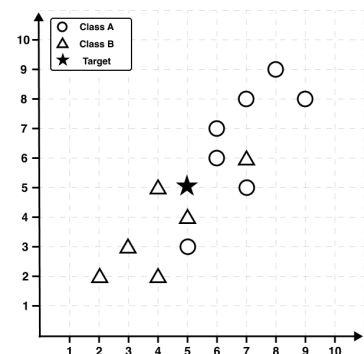
**문제 2.** 네 종류의 물건 (A, B, C, D)가 실제 팔릴 확률의 분포  $P$ 와 이를 기계 학습 모델이 예측한 결과  $Q$ 가 다음과 같다. 아래 값들을 구하여라.

$$P = [\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}], Q = [\frac{1}{2}, \frac{1}{8}, \frac{1}{4}, \frac{1}{8}]$$

1.  $P$ 의 엔트로피,  $H(P)$
2.  $P$ 와  $Q$ 의 교차 엔트로피,  $H(P, Q)$ .
3.  $P$ 와  $Q$ 의 Kullback-Leibler 발산,  $KL(P||Q)$

**문제 3.** 오른쪽 그림에 있는 타겟(별표)의 범주를  $k$ -최근접점 분류 알고리즘을 이용하여 A(원형) 또는 B(삼각형)의 클래스로 예측하고자 한다.

1.  $k = 3$ 일 때 타겟은 무슨 클래스로 분류되는가?
2.  $k = 5$ 일 때 타겟은 무슨 클래스로 분류되는가?
3.  $k = 5$ 일 때, 각 근접점이 L2 거리의 반비례하는 가중치를 가진다면, 타겟은 무슨 클래스로 분류되는가?



**문제 4.** 두 파라미터  $w_0$ 와  $w_1$ 에 대한 손실 함수가 다음과 같이 주어졌다.

$$E(w_0, w_1) = 3w_1^2 + 2w_0^2 + 5w_1w_0 - 4w_1 - 8w_0 + 7$$

1.  $\frac{\partial E}{\partial w_0}$ 와  $\frac{\partial E}{\partial w_1}$ 를  $w_0$ 와  $w_1$ 에 대한 식으로 나타내어라.
2. 초기 값  $w_0 = 2$ ,  $w_1 = 0$ 이고, 학습률(Learning rate)  $\alpha = 0.05$ 일 때, 경사 하강법(Gradient descent)을 한 번 수행한 후  $w_0$ 와  $w_1$ 의 값을 구하여라.
3. 손실 함수  $E$ 의 임계점을 찾아라.

**문제 5.** 다음과 같은 데이터가 주어져 있다:

$$(x_1, y_1) = (1, 2), \quad (x_2, y_2) = (2, 3), \quad (x_3, y_3) = (3, 5)$$

1. 오차를 실제 값과 예측 값의 차이의 제곱으로 두고, 오차의 합을 최소화하는 선형 회귀 직선  $y = w_0 + w_1x$ 을 편미분을 사용하여 구하여라.
2. 정규 방정식(Normal equation)을 사용하여 선형 회귀 직선  $y = w_0 + w_1x$ 을 구하여라.

**문제 6.** 당신은 10,000개의 데이터 포인트로 이진 분류 모델을 훈련하고 있으며, 현재 모델은 훈련 데이터에서 98% 정확도를, 검증 데이터에서 91% 정확도를 보이고 있다. 다음 중 이 모델의 일반화(Generalization) 성능을 향상시키는 방법으로 옳은 것을 모두 고르시오.

- a. 데이터 증강 (Data augmentation)
- b. 훈련 데이터에 대한 손실 최소화
- c. L2 정규화 하이퍼파라미터  $\lambda$ 값을 0.01에서 0.1로 증가
- d. 조기 종료 (Early stopping)
- e. 검증 데이터의 일부를 훈련 데이터에 추가
- f. 모델의 복잡도 증가
- g. 새로운 데이터를 수집하여 기존 훈련 데이터셋에 추가
- h. 5-Fold Cross-Validation 사용

**문제 7.** 다음은 선형 회귀(Linear regression)와 로지스틱 회귀(Logistic regression)에 대한 설명이다. 옳은 문장을 모두 고르시오.

- a. 두 모델 모두 입력 특성에 대한 선형 결합(Linear combination)을 사용한다.
- b. 로지스틱 회귀는 선형 결정 경계(Linear decision boundary)를 생성한다.
- c. 로지스틱 회귀는 시그모이드 함수를 사용하여 선형 결합의 출력을 확률로 변환한다.
- d. 로지스틱 회귀는 회귀 문제를 위한 지도 학습 알고리즘이다.
- e. 두 모델 모두 경사 하강법(Gradient descent)으로 학습될 수 있다.

문제 8. 다음과 같은 다중 클래스 분류 결과가 주어져 있다.

실제 레이블 (Ground-truth): [A, B, C, A, B]  
 모델 예측 (Model prediction): [B, B, A, A, A]

클래스 A에 대한 Precision, Recall, 그리고 F1-score를 계산하여라.

문제 9. 행성 A에 사는 외계인들의 키는 평균이  $\mu$ 이고 분산이  $\sigma^2$ 인 가우시안 분포  $N(\mu, \sigma^2)$ 를 따른다고 한다.  $\sigma^2$ 는 알려져 있다.  $N$ 개의 외계인 표본을 수집하여, 키의 평균  $\mu$ 를 추정하고자 한다. 이 때, 우도 함수는 아래와 같이 나타내어진다. 단,  $X_i$ 는  $i$ 번째 표본의 키이다.

$$P(X | \mu) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right)$$

또한, 외계인 전문가가  $\mu$ 에 대한 사전 확률이  $N(\mu_0, \sigma^2)$ 를 따른다고 귀띔해 주었다. 즉, 사전 확률 밀도 함수는 다음과 같다.

$$P(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma^2}\right)$$

베이즈 정리에 따라 아래와 같은 사실을 알고 있다.

$$P(\mu | X) \propto P(X | \mu)P(\mu)$$

1. 최대우도추정법(MLE)을 사용하여  $\mu$ 의 추정치  $\hat{\mu}_{MLE}$ 를 구하시오.
2. 최대사후확률(MAP) 추정법을 사용하여  $\mu$ 의 추정치  $\hat{\mu}_{MAP}$ 를 구하시오.
3. 표본의 수  $N$ 이 증가함에 따라 MAP 추정치는 어디에 수렴하는가?

문제 10. 성균관대학교 자연과학캠퍼스에는 제1공학관, 제2공학관, 반도체관, 이렇게 세 개의 건물이 있다. 각 건물의 학생 분포는 다음과 같다:

건물	소프트웨어학과	수학과	컴퓨터교육과
제1공학관	4명	3명	3명
제2공학관	1명	1명	0명
반도체관	3명	3명	4명

총장님이 자연과학캠퍼스 건물 중 한 곳을 방문하여 무작위로 학생을 만난다. 방문 확률은  $P(\text{제1공학관}) = 0.2$ ,  $P(\text{제2공학관}) = 0.2$ ,  $P(\text{반도체관}) = 0.6$ 이다. 총장님이 만난 학생이 소프트웨어학과 소속일 때, 총장님이 반도체관을 방문했을 확률은 얼마인가?