

기계학습원론 중간고사
2025 봄

강의자: 조재민
강의조교: 이재웅

April 22, 2025

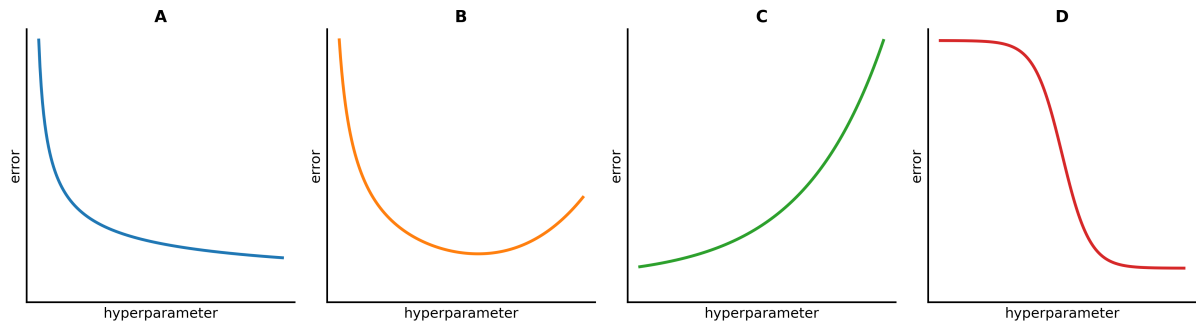
문제 1. 명제의 참(O) 또는 거짓(X)을 판단하시오. (정답: +2점, 오답: -2점, 미기재: 0점)

1. 임의의 $x, x' \in X, t \in [0, 1]$ 에 대해 $f(tx + (1 - t)x') \leq tf(x) + (1 - t)f(x')$ 인 함수 $f : X \rightarrow \mathbb{R}$ 에서, 어떤 지점의 그래디언트가 0이면 항상 전역 최솟값이다.
2. 어떤 함수가 볼록(convex)하면 최솟값을 지닌다.
3. 어떤 함수는 볼록(convex)하면서 오목(concave)하다.
4. k -fold cross-validation에서 k 가 커질수록, 정확도의 분산(variance)은 감소한다.
5. 선형 회귀(linear regression)의 비용 함수는 파라미터에 대해 볼록(convex)하며, 닫힌 형태의 해(closed-form solution)가 존재한다.
6. 로지스틱 회귀(logistic regression)의 비용 함수는 파라미터에 대해 볼록(convex)하지만, 닫힌 형태의 해(closed-form solution)는 존재하지 않는다.
7. 두 확률분포 P, Q 에 대해, $KL(P\|Q) > 0$ 이면 $KL(Q\|P) > 0$ 이다.
8. 평균이 μ , 표준편차가 σ 인 정규분포의 확률 밀도 함수는 $f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ 이다.
9. 표준 정규분포를 따르는 확률변수 X 에 대해 X 가 0일 확률은 $f(0, 0, 1)$ 이다.
10. 분산(Variance)이 증가하면 편향(Bias)은 감소한다.
11. 편향(Bias)이 증가하면 분산(Variance)은 감소한다.
12. L2 정규화는 파라미터의 절댓값을 작게 만들며 특히 정확히 0으로 만드는 선택 효과가 있다.
13. 정규화의 강도를 조절하는 하이퍼파라미터 λ 는 학습 도중에 튜닝된다.
14. 클래스가 2개 이상일 때, Softmax 함수의 출력값들은 항상 1보다 작고, 모두 양수이며 합은 1이다.
15. K-최근접이웃(K-Nearest Neighbor) 기반 모델들은 학습 시간은 느리지만 추론 시간이 빠르다는 장점이 있다.

문제 2. [8점] 다음은 로지스틱 회귀(Logistic regression)와 소프트맥스 회귀(Softmax regression)에 대한 설명이다. 옳은 문장을 모두 고르시오.

- a. 로지스틱 회귀는 선형 결정 경계(Linear decision boundary)를 생성한다.
- b. 로지스틱 회귀는 시그모이드 함수를 사용하여 선형 결합의 결과를 확률로 변환한다.
- c. 입력 벡터의 모든 성분에 같은 수를 더하더라도 소프트맥스 함수를 적용한 결과값은 변하지 않는다.
- d. 입력 벡터의 모든 성분에 같은 수를 곱하더라도 소프트맥스 함수를 적용한 결과값은 변하지 않는다.
- e. 로지스틱 회귀와 소프트맥스 회귀는 모두 실제 클래스와 예측 확률 분포 사이의 크로스 엔트로피(Cross entropy)를 최소화하도록 경사 하강법으로 학습된다.

문제 3. [8점] 여러 머신러닝 알고리즘에 대해, 하이퍼파라미터 값에 따라 훈련 오차(Training Error) 또는 테스트 오차(Testing Error)가 변할 수 있는 몇 가지 가능한 방식은 다음과 같다:



아래에서 제시하는 각각의 하이퍼파라미터와 관련하여, 위 그림 중 어떤 것이 훈련 오차와 테스트 오차에 대한 가장 일반적인 경향을 나타내는지 고르시오.

1. 로지스틱 회귀에서 정규화 항의 가중치(Regularization Weight, λ):
 - i. 훈련 오차 (Training Error): ☐ A ☐ B ☐ C ☐ D
 - ii. 테스트 오차 (Testing Error): ☐ A ☐ B ☐ C ☐ D
2. 경사 하강법(Gradient Descent)의 학습률(Learning Rate, η) (반복 횟수는 고정):
 - i. 훈련 오차 (Training Error): ☐ A ☐ B ☐ C ☐ D
 - ii. 테스트 오차 (Testing Error): ☐ A ☐ B ☐ C ☐ D
3. 최근접 이웃(k-NN) 분류에서의 이웃 수(Number of Neighbors, k):
 - i. 훈련 오차 (Training Error): ☐ A ☐ B ☐ C ☐ D
 - ii. 테스트 오차 (Testing Error): ☐ A ☐ B ☐ C ☐ D
4. 경사 하강법의 에포크 수(Number of Epochs):
 - i. 훈련 오차 (Training Error): ☐ A ☐ B ☐ C ☐ D
 - ii. 테스트 오차 (Testing Error): ☐ A ☐ B ☐ C ☐ D

문제 4. [8점] 두 파라미터 w_0 와 w_1 에 대한 손실 함수가 다음과 같다:

$$E(w_0, w_1) = 3w_0^2 + 2w_1^2 + 4w_0w_1 - 8w_0 - 10w_1 + 6$$

1. 손실 함수 $E(w_0, w_1)$ 를 최소화하는 파라미터 w_0, w_1 의 값을 구하여라.
2. 초기 값 $(w_0, w_1) = (1, -1)$, 학습률 $\alpha = 0.1$ 일 때, 경사 하강법을 한 번 수행한 후의 w_0, w_1 값을 구하여라.

문제 5. [8점] 다음 표는 실제 클래스와 모델이 예측한 클래스의 혼동 행렬(confusion matrix)을 나타낸 것이다:

예측 \ 실제	고양이	강아지	물고기
고양이	5	2	1
강아지	1	6	1
물고기	1	2	5

1. 강아지 클래스에 대한 Precision, Recall, F1-score를 구하여라.
2. F1-score는 Precision과 Recall의 산술평균과 비교할 때 항상 (크다 / 크거나 같다 / 작거나 같다 / 작다).

문제 6. [8점] 빨강, 초록, 노랑의 세 가지 색상의 상자가 있다. 빨강 상자에는 사탕 3개, 초콜릿 2개, 껌 1개가 있고, 초록 상자에는 사탕 2개, 초콜릿 2개, 껌 2개가 있으며, 노랑 상자에는 사탕 1개, 초콜릿 2개, 껌 3개가 있다. 상자를 $p(\text{빨강}) = 0.5$, $p(\text{초록}) = 0.3$, $p(\text{노랑}) = 0.2$ 의 확률로 무작위로 선택하고, 선택된 상자에서 하나의 물건을 꺼내려 한다. (단, 각 물건은 동일한 확률로 선택된다.)

1. 초콜릿을 선택할 확률은 얼마인가?
2. 선택한 물건이 껌이라는 것을 관찰했을 때, 이 물건이 노랑 상자에서 나왔을 확률은 얼마인가?

문제 7. [10점] 다음 목적함수를 최대화하는 문제를 라그랑주 승수법(Lagrange multiplier)을 사용하여, 최적해 (x, y, z) 와 최적값 $f(x, y, z)$ 를 구하여라.

$$f(x, y, z) = xy + yz + 2xz, \quad \text{s.t.} \quad x + y + z = 6$$

문제 8. [10점] 다음 데이터를 고려하자:

$$(x_1, y_1) = (1, 3), \quad (x_2, y_2) = (2, 5), \quad (x_3, y_3) = (3, 6)$$

행렬 X , 벡터 \mathbf{y} , 파라미터 벡터 \mathbf{w} , 손실 함수 $E(\mathbf{w})$ 를 다음과 같이 정의한다:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, \quad E(\mathbf{w}) = (\mathbf{y} - X\mathbf{w})^\top (\mathbf{y} - X\mathbf{w})$$

1. 위 손실 함수를 최소화하는 \mathbf{w} 에 대한 정규방정식(normal equation)을 유도하여라.
2. 정규방정식을 사용하여 선형 회귀 직선 $y = w_0 + w_1x$ 을 구하여라.

문제 9. [10점] 동전을 10번 던졌더니 앞면이 7번, 뒷면이 3번 나왔다고 하자. 동전이 앞면이 나올 확률을 θ 라 할 때, 다음 물음에 답하시오.

1. 동전 던지기의 결과에 대한 Likelihood를 θ 에 대한 식으로 나타내시오.
2. θ 에 대한 Maximum Likelihood Estimation (MLE)을 구하시오.
3. θ 가 사전 분포 $P(\theta) \sim \text{Beta}(4, 4)$ 를 따른다고 할 때, Maximum A Posteriori (MAP) 추정값을 구하시오.