

기계학습 (HYU-KITECH공동학과)

Subject : Machine Learning, Lecturer : Prof. Youngjin Choi,

Date : Mar. 24, 2022 (Contact e-mail : cyj@hanyang.ac.kr)

Problem 1 (20pt) 다음 함수에 대해 답하십시오

$$f(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + 4x_3^2 + 6x_1x_2 + 8x_2x_3 + 12x_1x_3 + 7x_1 + 8x_2 + 9x_3 + 10$$

(1.1, 10pt) 최소값 ?

(1.2, 10pt) 초깃값 $\mathbf{x}_0 = \begin{bmatrix} -1.0 \\ 0.0 \\ 0.0 \end{bmatrix}$ 에서 시작하여, 학습률 $\rho = 0.05$ 을 가지고 경사하강법을 연속 적용하여 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 구한 해가 최소값을 찾아가는지 확인하라 ?

Solution of Problem 1 (20pt) 다음 함수에 대해 답하십시오

$$f(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + 4x_3^2 + 6x_1x_2 + 8x_2x_3 + 12x_1x_3 + 7x_1 + 8x_2 + 9x_3 + 10$$

(1.1, 10pt) Jacobian and Minimum value are

$$\frac{\partial f}{\partial \mathbf{x}} = \mathbf{0} \quad \rightarrow \quad \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 6x_1 + 6x_2 + 12x_3 + 7 \\ 6x_1 + 4x_2 + 8x_3 + 8 \\ 12x_1 + 8x_2 + 8x_3 + 9 \end{bmatrix} = \mathbf{0} \quad \rightarrow \quad \therefore \quad x_1 = -1.667 \quad x_2 = 2.25 \quad x_3 = -0.875$$

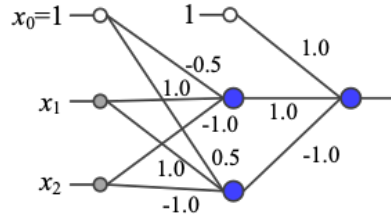
$\therefore \quad f(x_1, x_2, x_3) = 9.2292$

(1.2, 10pt) Gradient Decent Rule

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0 - \rho \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_0} = \begin{bmatrix} -1.0 \\ 0.0 \\ 0.0 \end{bmatrix} - 0.05 \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -1.05 \\ -0.1 \\ 0.15 \end{bmatrix} \quad \rightarrow \quad f(\mathbf{x}_1) = 5.2375 \\ \mathbf{x}_2 &= \mathbf{x}_1 - \rho \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_1} = \begin{bmatrix} -1.05 \\ -0.1 \\ 0.15 \end{bmatrix} - 0.05 \begin{bmatrix} 1.9 \\ 2.5 \\ -3.2 \end{bmatrix} = \begin{bmatrix} -1.145 \\ -0.225 \\ 0.31 \end{bmatrix} \quad \rightarrow \quad f(\mathbf{x}_2) = 4.1221 \\ \mathbf{x}_3 &= \mathbf{x}_2 - \rho \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_2} = \begin{bmatrix} -1.145 \\ -0.225 \\ 0.31 \end{bmatrix} - 0.05 \begin{bmatrix} 2.5 \\ 2.71 \\ -4.06 \end{bmatrix} = \begin{bmatrix} -1.27 \\ -0.3605 \\ 0.513 \end{bmatrix} \quad \rightarrow \quad f(\mathbf{x}_3) = 2.4437 \end{aligned}$$

함수값 $f(\mathbf{x})$ 은 지속적으로 낮아지고 있어서 언젠가는 최소값을 찾아갈 것으로 기대된다.

Problem 2 (40pt) 입력 $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 와 레이블 $y = 1$ 이 주어진다고 가정하자. 은닉층 및 출력층의 모든 활성화함수에 ReLU 적용해 보자.



(2.1, 10pt) 가중치 행렬 U^1 and U^2 ?

(2.2, 10pt) 출력 o ?

(2.3, 10pt) 평균제곱오차를 목적함수로 사용한다면, 오류 역전파 계산 과정 중의 ΔU^2 ?

(2.4, 10pt) 오류 역전파 계산 과정 ΔU^1 ?

Solution of Problem 2 (40pt) 다음 함수에 대해 답하시오

(2.1, 10pt)

$$U^1 = \begin{bmatrix} -0.5 & 1.0 & -1.0 \\ 0.5 & 1.0 & -1.0 \end{bmatrix} \qquad U^2 = \begin{bmatrix} 1.0 & 1.0 & -1.0 \end{bmatrix}$$

(2.2, 10pt)

$$\mathbf{z} = \tau_1(U^1 \mathbf{x}) = \text{ReLU}(U^1 \mathbf{x}) = \text{ReLU} \left(\begin{bmatrix} -0.5 & 1.0 & -1.0 \\ 0.5 & 1.0 & -1.0 \end{bmatrix} \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \right) = \begin{bmatrix} 0.0 \\ 0.5 \end{bmatrix}$$

$$o = \tau_2(U^2 \mathbf{z}) = \text{ReLU}(U^2 \mathbf{z}) = \text{ReLU} \left(\begin{bmatrix} 1.0 & 1.0 & -1.0 \end{bmatrix} \cdot \begin{bmatrix} 1.0 \\ 0.0 \\ 0.5 \end{bmatrix} \right) = 0.5$$

(2.3 20pt) Using $\delta_1 = (y - o)\tau'(o) = 0.5 \cdot 1 = 0.5$

$$\Delta U^2 = - \begin{bmatrix} \delta_1 z_0 & \delta_1 z_1 & \delta_1 z_2 \end{bmatrix} = - \begin{bmatrix} 0.5 & 0.0 & 0.25 \end{bmatrix}$$

(2.4, 10pt) Using $\eta_1 = \tau'(z_{sum1})\delta_1 u_{11}^2 = 0.0$ and $\eta_2 = \tau'(z_{sum2})\delta_1 u_{12}^2 = -0.5$

$$\Delta U^1 = - \begin{bmatrix} \eta_1 x_0 & \eta_1 x_1 & \eta_1 x_2 \\ \eta_2 x_0 & \eta_2 x_1 & \eta_2 x_2 \end{bmatrix} = - \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ -0.5 & -0.5 & -0.5 \end{bmatrix}$$

Problem 3 (40pt) 아래와 같이 주어진 5×7 이미지에 대하여 다음과 같은 2가지 콘볼루션 연산을 수행하라?

0	0	0	0	0	0	0
0	1	1	1	1	1	0
0	0	0	0	0	1	0
0	1	1	1	1	1	0
0	0	0	0	0	0	0

(3.1, 10pt) 커널

-1	-1	-1
0	0	0
1	1	1

 를 보폭 $s=2$ 로 적용한 결과를 쓰시오. 이때 0의 덧대기를 하고 바이어스로 0를 사용한다.

(3.2, 10pt) 커널

-1	0	1
-1	0	1
-1	0	1

 를 보폭 $s=2$ 로 적용한 결과를 쓰시오. 이때 0의 덧대기를 하고 바이어스로 0를 사용한다.

(3.3, 10pt) 3.1의 결과에 0의 덧대기 하고 3×3 마스크를 가지고 보폭 $s=1$ 로 최대풀링 (max pooling) 하라

(3.4, 10pt) 3.2의 결과에 0의 덧대기 하고 3×3 마스크를 가지고 보폭 $s=1$ 로 최대풀링 (max pooling) 하라

Solution of Problem 3 (40pt)

(3.1, 10pt)

1	3	3	1
0	0	0	0
-1	-3	-3	-1

(3.2, 10pt)

1	0	0	-1
2	0	1	-3
1	0	0	-1

(3.3, 10pt)

3	3	3	3
3	3	3	3
0	0	0	0

(3.4, 10pt)

2	2	1	1
2	2	1	1
2	2	1	1