강의 5 - Back-propagation 첫번째

이번 강의는 실질적인 Back-propagation을 하기 전에 이론적으로 Chaine Rule을 적용하는지 배웠습니다. 이를 바탕으로 수식으로 배운 것을 복습해봅시다.

1) 주피터 노트북에서 수식 사용하기

이번 과제는 주피터 노트북에 수식을 입력해야 합니다. 기초적인 수식 입력 방법에 대해서 알아봅시다.

ㄱ. 마법의 $ 표시

주피터 노트북에서 2개의 $ 사이에 수식을 입력합니다. $ f(x) = 2x $를 입력해봅시다.

정답) f(x) = 2x

ㄴ. 각종 Symbol들

다음과 같은 symbol를 이용할 수 있습니다.

$$\cdot, \alpha, \beta, \sum, \Delta$$

이와의 여러가지 기호들을 latex symbols라고 찾으면 됩니다. Gradient와 Jacobain 상징인 기호를 찾아봅시다

정답) $\nabla_{\vec{u}} l, \nabla_{\vec{u}} \vec{l}$

ㄷ. 위첨자와 아래 첨자

위 첨자는 ^ 기호로 아래 첨자는 _ 기호로 만들 수 있습니다.

1.x 의 2승 표시를 해봅시다.

2.더불어 첨자가 2개 이상인 경우에는 어떻게 할 수 있는지 찾아봅시다. 즉 e 의 $i \cdot \theta$승을 해봅시다

정답 1) x^2

정답 2) $e^{i \cdot \theta}$

ㄹ. 매트릭스를 만들어 봅시다.

매트릭스는 \pmatrix 로 만들 수 있습니다.

예시를 보고 새로운 3x3 매트릭스를 만들어 보세요.

예시 $\begin{pmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{pmatrix}$

정답) $\begin{pmatrix} x_{00} & x_{01} & x_{02} \\ x_{10} & x_{11} & x_{11} \\ x_{20} & x_{21} & x_{21} \end{pmatrix}$

ㅁ. 분수

분수는 $\setminus frac\{\}\{\}$으로 나타낼 수 있습니다. 1/2을 분수로 나타내봅시다

정답) $\frac{1}{2}$

ㅂ. 최종 점검

다음 이미지를 수식으로 변경하시오

정답)$L_j = ReLU(\sum_i^n w_{ij} \cdot L_i + B_j)$

2) 미분하기

위에서 배운 수식 입력 방법을 가지고 계산을 해봅시다!

ㄱ. 다음 함수의 $\frac{dy}{dx}$를 구하시오

$$y = f(u) \qquad u = g(x)$$

정답) $\frac{dy}{du}\frac{du}{dx}$

ㄴ. 다음 함수의 ∇y를 $\vec{x}$의 원소로 풀어서 쓰시오.

단, $\vec{x} = [x_0, x_1, x_2, x_3]$이다.

$$y = f(x)$$

정답) $\nabla y = [\frac{\partial y}{\partial x_0}, \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial y}{\partial x_3}]$

ㄷ. 다음 함수의 ∇y를 $\vec{u}$와 $\vec{x}$의 원소로 풀어서 쓰시오.

단, $\vec{u}$는 $[u_0, u_1, u_2]$ 이며 $\vec{x}$는 $[x_0, x_1]$ 이다.

$$y = f(\vec{u}) \qquad \vec{u} = g(\vec{x})$$

정답) $\nabla y = [\frac{\partial y}{\partial x_0}, \frac{\partial y}{\partial x_1}]$ 이며 각 원소는 아래와 같다

$$\frac{\partial y}{\partial x_0} = \frac{\partial y}{\partial u_0}\frac{\partial u_0}{\partial x_0} + \frac{\partial y}{\partial u_1}\frac{\partial u_1}{\partial x_0} + \frac{\partial y}{\partial u_2}\frac{\partial u_2}{\partial x_0}$$
$$\frac{\partial y}{\partial x_1} = \frac{\partial y}{\partial u_0}\frac{\partial u_0}{\partial x_1} + \frac{\partial y}{\partial u_1}\frac{\partial u_1}{\partial x_1} + \frac{\partial y}{\partial u_2}\frac{\partial u_2}{\partial x_1}$$

ㄹ. 다음 함수의 Jacobian 인 $\nabla \vec{y}$를 $\vec{y}$와 $\vec{x}$의 원소로 풀어서 쓰시오.

단, $\vec{y} = [y_0, y_1, y_2] \quad \vec{x} = [x_0, x_1, x_2, x_3]$이다.

$$\vec{y} = f(\vec{x})$$

정답) $\nabla \vec{y} = \begin{pmatrix} \frac{\partial y_0}{\partial x_0} & \frac{\partial y_0}{\partial x_1} & \frac{\partial y_0}{\partial x_2} & \frac{\partial y_0}{\partial x_3} \\ \frac{\partial y_1}{\partial x_0} & \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_0} & \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \end{pmatrix}$

ㅁ. 위 문제ㄷ의 결과를 Jacobian인 $\nabla_{\vec{x}}\vec{u}$과 Gradient인 $\nabla_{\vec{u}}y$ 로 표기하고 각각$\nabla_{\vec{x}}\vec{u}$ 와 $\nabla_{\vec{u}}y$를 y와 $[u_0, u_1, u_2]$과 $[x_0, x_1]$로 표기하세요

정답) $\nabla y = \nabla_{\vec{u}}y \cdot \nabla_{\vec{x}}\vec{u}$ 과

$\nabla_{\vec{u}}y = \frac{\partial y}{\partial u_0} + \frac{\partial y}{\partial u_1} + \frac{\partial y}{\partial u_2}$

$$\nabla_{\vec{x}}\vec{u} = \begin{pmatrix} \frac{\partial u_0}{\partial x_0} & \frac{\partial u_0}{\partial x_1} \\ \frac{\partial u_1}{\partial x_0} & \frac{\partial u_1}{\partial x_1} \\ \frac{\partial u_2}{\partial x_0} & \frac{\partial u_2}{\partial x_1} \end{pmatrix}$$

ㅂ. 다음 함수의 $\nabla_{\vec{w}}\vec{y}$를 $\vec{y}$와 $\vec{w}$의 원소로 풀어서 쓰시오

단, $\vec{y} = [y_0, y_1, y_2]$ $\quad \vec{w} = \begin{pmatrix} w_{00} & w_{01} & w_{02} \\ w_{10} & w_{11} & w_{12} \\ w_{20} & w_{21} & w_{22} \end{pmatrix}$이다.

$$\vec{y} = f(\vec{w})$$

정답) $\nabla_{\vec{w}}\vec{y} = \begin{pmatrix} \frac{\partial y_0}{\partial w_{00}} & \frac{\partial y_0}{\partial w_{01}} & \frac{\partial y_0}{\partial w_{02}} & \frac{\partial y_0}{\partial w_{10}} & \frac{\partial y_0}{\partial w_{11}} & \frac{\partial y_0}{\partial w_{12}} & \frac{\partial y_0}{\partial w_{20}} & \frac{\partial y_0}{\partial w_{21}} & \frac{\partial y_0}{\partial w_{22}} \\ \frac{\partial y_1}{\partial w_{00}} & \frac{\partial y_1}{\partial w_{01}} & \frac{\partial y_1}{\partial w_{02}} & \frac{\partial y_1}{\partial w_{10}} & \frac{\partial y_1}{\partial w_{11}} & \frac{\partial y_1}{\partial w_{12}} & \frac{\partial y_1}{\partial w_{20}} & \frac{\partial y_1}{\partial w_{21}} & \frac{\partial y_1}{\partial w_{22}} \\ \frac{\partial y_2}{\partial w_{00}} & \frac{\partial y_2}{\partial w_{01}} & \frac{\partial y_2}{\partial w_{02}} & \frac{\partial y_2}{\partial w_{10}} & \frac{\partial y_2}{\partial w_{11}} & \frac{\partial y_2}{\partial w_{12}} & \frac{\partial y_2}{\partial w_{20}} & \frac{\partial y_2}{\partial w_{21}} & \frac{\partial y_2}{\partial w_{22}} \end{pmatrix}$

ㅅ. 다음 인공뉴런 1개에 대한 내용이다.

$\vec{I} = [I_0, I_1]$ 이며 입력값이고,

$\vec{w} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}$ 이며 입력값에 곱해지는 weight 값이며,

$h = \vec{I} \cdot \vec{w}$,

$f = ReLU(h)$, 단 $ReLU(x) = \begin{cases} 0 & x \le 0 \\ x & x > 0 \end{cases}$

일 때 $\nabla_{\vec{w}}f$를 I_0, I_1, h로 표기하시오

정답) $\nabla_{\vec{w}}f = \begin{cases} [0, 0] & h \le 0 \\ [I_0, I_1] & h > 0 \end{cases}$

ㅇ. 위 문제에서 $\nabla_{\vec{I}}f$ 를 구하시오

정답) $\nabla_{\vec{I}}f = \begin{cases} [0, 0] & h \le 0 \\ [W_0, W_1] & h > 0 \end{cases}$

ㅈ. 다음 인공 뉴런 1개에 대한 내용이다

$\vec{I} = [2, 3]$ 이며 입력값이고,

$\vec{w} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ 이며 입력값에 곱해지는 weight 값이며,

$h = \vec{I} \cdot \vec{w}$,

$f = ReLU(h)$, 단 $ReLU(x) = \begin{cases} 0 & x \le 0 \\ x & x > 0 \end{cases}$

일 때 $\nabla_{\vec{w}} f$를 벡터로 표기하시오

정답) [2, 3]

ㅊ. 위 ㅈ 문제에서 $\nabla_{\vec{i}} f$ 를 구하시오

정답) [4, 5]

In [ ]:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js