< Part 2. > 딥러닝 시작하는 파트 토픽

[Ch01.] 딥러닝 네트워크의 역사

[CH01.] 01, 02

Lecture. 1

Artificial Neurons

- Part2 데이터 사이언스 필수 수학 (기초)
- Ch01 딥러닝 네트워크의 연산

CH00-01 Introduction

← Training Neural Networks

(model training)

이번강의 : model prediction → Loss Calculation

↓ Gradient Calculation

다음강의 : Parameter Update ← (Gradient Calculation), ↑ (model prediction)

특징 : 르넷이라는 모델 만들기, CNN의 가장 Classic 한 모델. (LeNet)

+ 최종적으로 모델 만드는지, 어떤 연산이 이루어지는지.

forward : AN → Dence → Sigmoid / Softmax → Loss → Con → Pooling → CNN

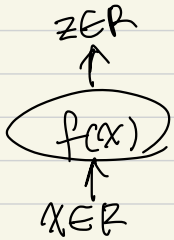
CH00_02 Colaboratory

Back End를 구글이 제공.
간단한 프로젝트까지 가능.

CH01_01. [이론정리] Parameteric Functions and Datasets.

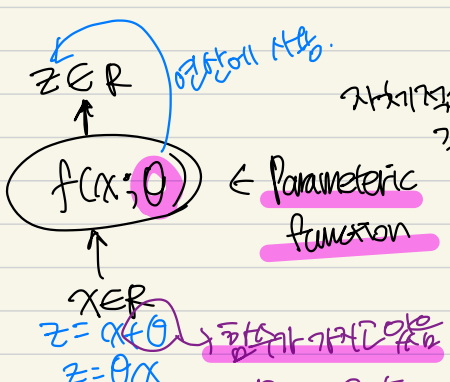
Lecture.1 Artificial Neurons.

- Parameteric Functions.

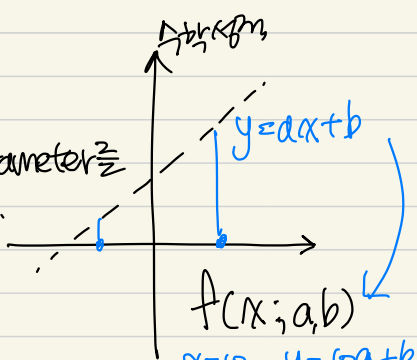


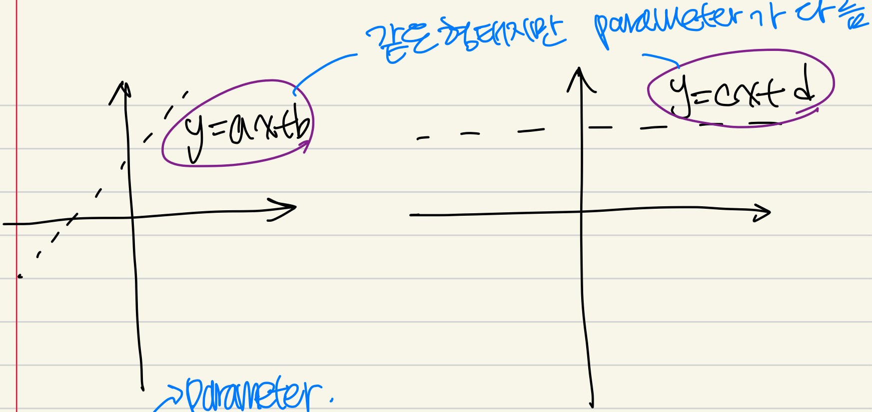
입력받은 규칙에따라 아웃풋 반환.

(우리가아는 함수) $z=e^x$ $z=\log(x)$



자기자신으로 Parameter를 갖고있음.





$f(x;\theta)$ → 에서 θ가 달라지면 아예 다른 함수가 된다. (θ → parameter.)

☆ → θ가 달라지면 연산자체가 달라짐.

① Parameter를 내재하고 있는 함수를 다룬다.
② θ가 단순 하나가 아니다. 바뀔.. 천만개 될수 있다.
③ x (입력값)이 고차원 정보이다. 벡터, 매트릭스..

- Tensor 개념.

(scalar, vector, metrix) → (general 표현) Tensor

- Zeroth order tensor (0차원) : scalar

$a, b \in \mathbb{R}$

$a+b : \mathbb{R}\times\mathbb{R} \rightarrow \mathbb{R}$: $(5,10) \in \mathbb{R}\times\mathbb{R}$ → $15 \in \mathbb{R}$

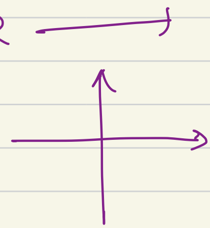
$5+10 \rightarrow 15$

$\mathbb{R}\times\mathbb{R}$ → Cartesian Product

$a \cdot b : \mathbb{R}\times\mathbb{R} \rightarrow \mathbb{R}$

$\mathbb{R}$

$\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$



$B = \{0, 1\}$

$B\times B = \{(0,0), (0,1), (1,0), (1,1)\}$

- First-Order tensor (vector) (zeroth-order operation들과 같음)

$a \in \mathbb{R}$

$\vec{u}, \vec{v} \in \mathbb{R}^n$ $\quad \vec{u} = \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$ $\quad \vec{v} = \begin{pmatrix} 50 \\ 100 \\ 1000 \end{pmatrix}$

$a\vec{u} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$

$2 \times \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 10 \end{pmatrix}$

$\mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$

$\vec{u} + \vec{v} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$

$\vec{u} \circ \vec{v} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \Rightarrow \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \circ \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix} = \begin{pmatrix} 20 \\ 60 \\ 120 \end{pmatrix}$

$\vec{u}^T \cdot \vec{v} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

$\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = 2 + 6 + 12 = 20$

- Second-order tensor (Matrix)

$M \in \mathbb{R}^{2 \times 3}$

$M = \begin{pmatrix} 2 & 3 & 4 \\ 10 & 20 & 30 \end{pmatrix}$

$a \in \mathbb{R}$ $\quad \vec{u} \in \mathbb{R}^n$

$M, N \in \mathbb{R}^{m \times n}$, $O \in \mathbb{R}^{n \times o}$

$aM : \mathbb{R} \times \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$

$2 \times \begin{pmatrix} 10 & 20 \\ 30 & 40 \end{pmatrix} = \begin{pmatrix} 20 & 40 \\ 60 & 80 \end{pmatrix}$

$\mathbb{R}$ $\quad \mathbb{R} \times \mathbb{R}^{2 \times 2}$

$M + N : \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$

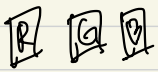
$M \circ N : \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$

$M \cdot \vec{u} : \mathbb{R}^{m \times n} \times \mathbb{R}^n = \mathbb{R}^m$

$MO : \mathbb{R}^{m \times n} \times \mathbb{R}^{n \times o} \to \mathbb{R}^{m \times o}$

- Third-Order Tensor Operations.

$\mathbb{R}^{2\times n\times 3}$ $\mathbb{R}^{28\times 28\times 3}$ (MNIST)



$a \in R$ $M, N \in R^{m\times n\times o}$

$aM : R \times R^{m\times n\times o} \to R^{m\times n\times o}$

$M+N : R^{m\times n\times o} \times R^{m\times n\times o} \to R^{m\times n\times o}$

$M \circ N : R^{m\times n\times o} \times R^{m\times n\times o} \to R^{m\times n\times o}$

- Dataset(X, Data)

$\vec{x}^T = (x_1\ x_2\ x_3\ \cdots\ x_{l_i})$: 가로로 쓸 때 row vector로 표현.

독립변수 ⇐ (공복 혈당량, 혈압, 수면시간 등등)

$\vec{x}$

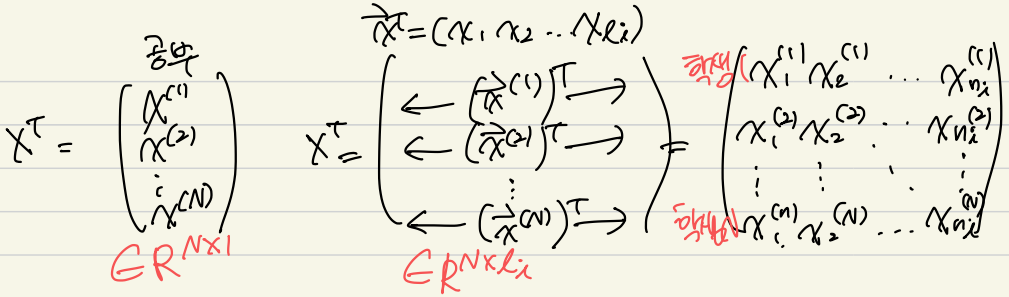
입력 1 $(\vec{x}^{(1)})^T = (x_1^{(1)}\ x_2^{(1)}\ \cdots\ x_{l_i}^{(1)})$

입력 2 $(\vec{x}^{(2)})^T = (x_1^{(2)}\ x_2^{(2)}\ \cdots\ x_{l_i}^{(2)})$

⋮

입력 N $(\vec{x}^{(N)})^T = (x_1^{(N)}\ x_2^{(N)}\ \cdots\ x_{l_i}^{(N)})$

length of input → l_i



new vector 추가 → Matrix

$(\vec{x})^T = (x_1\ x_2\ x_3\ x_4 \dots)$

$$X^T = \begin{pmatrix} \leftarrow (\vec{x}^{(1)})^T \rightarrow \\ \vdots \\ \leftarrow (\vec{x}^{(N)})^T \rightarrow \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_{n_i}^{(1)} \\ \vdots & & & \vdots \\ x_1^{(N)} & x_2^{(N)} & \cdots & x_{n_i}^{(N)} \end{pmatrix} \in R^{N\times l_i}$$

$X^T = (x_1\ x_2\ x_3\ \cdots\ x_{l_i})$

$$X^T = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_{n_i}^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_{n_i}^{(2)} \\ \vdots & & & \\ x_1^{(m)} & x_2^{(m)} & \cdots & x_{n_i}^{(n)} \end{pmatrix} \in R^{N\times l_i}$$

CH01_02 Artificial Neuron

- Affine Function with one Feature.

  - Weighted Sum
    $z = xw$
  - Affine Transformation
    $z = xw + b$

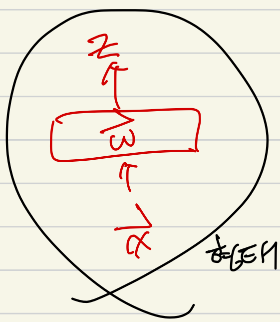
$z \in R$
↑
$f(x; w, b)$ $w, b \in R$
↑
$x \in R$

$z = x \cdot w + b$.

- Affine Functions with n Features

  - Weighted Sum
    $z = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n = (\vec{w})^T \vec{x} = (\vec{x})^T \vec{w}$
    (x_1: 키, x_2: 몸무게, x_n: 나이)
    더 많이 사용.
    w: 중요도

$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$



형태

  - Affine Transformation
    $z = (\vec{x})^T \vec{w} + b$
    l_i개 l_i개 1개

$z \in R$
↑
$f((\vec{x})^T; \vec{w}, b)$
↑
$(\vec{x})^T = (x_1\ x_2 \cdots x_{l_i}) \in R^{1 \times l_i}$

★l_i개 사용

$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{l_i} \end{pmatrix} \in R^{l_i \times 1}$, $b \in R$

★1개 사용.

- Activation Function

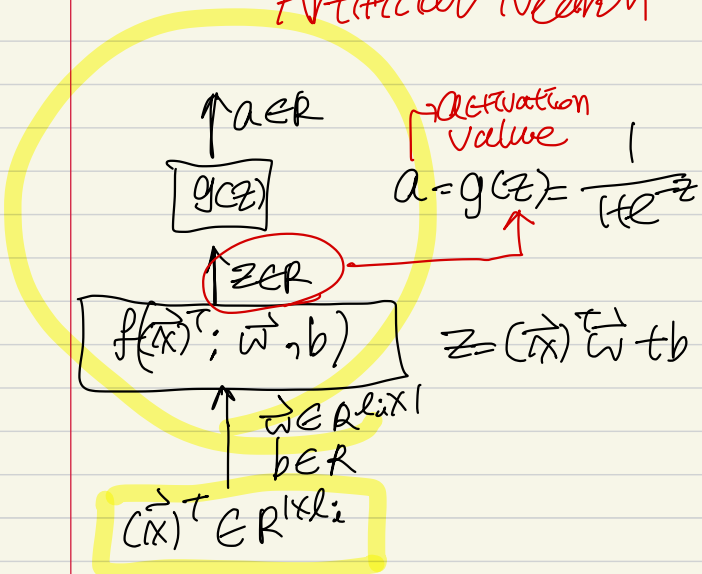
Sigmoid $g(x) = \sigma(x) = \frac{1}{1+e^{-x}}$

Tanh $g(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$

ReLU $g(x) = \mathrm{ReLU}(x) = \max(0, x)$

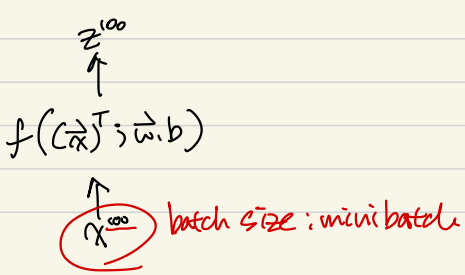
affine | activation fcn.

Artificial Neuron

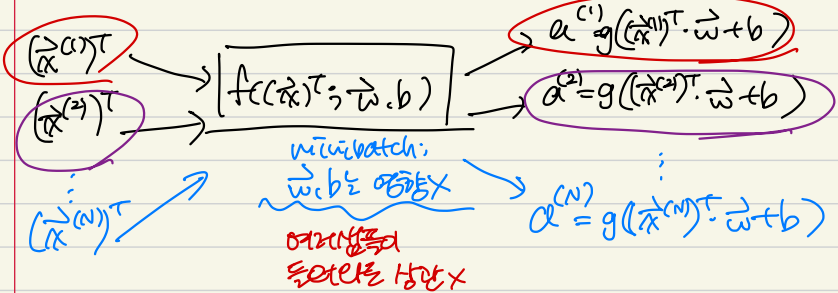


$$y(\vec{x}, \vec{w}, b) = g(\vec{x}^{T}\vec{w} + b)$$

- Minibatch in Artificial Neurons

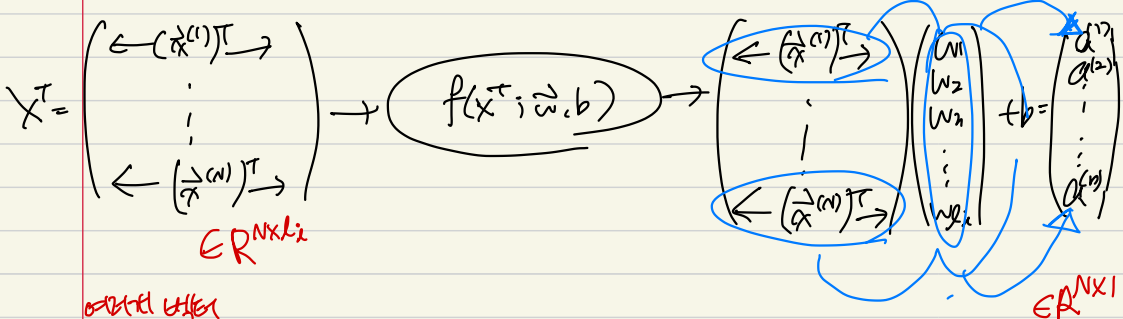


- Minibatch in ANN



★ $\vec{w}$는 가변의 값들.
(minibatch가 들어오더라도 $\vec{w}$는 변하지않기때문)

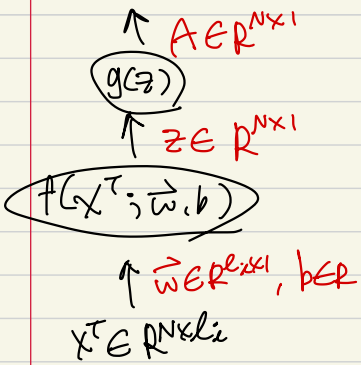
$f(x;\theta) = x+\theta$
$f(2;\theta) = 2+\theta$
$f(10;\theta) = 10+\theta$
$\rightarrow \theta$ 는 가변의 값들 (학습함)



여러개의 데이터
한의 Matrix로
묶어 → minibatch

∴ minibatch input → Weight/Bias → Affine function → Activation Function

$X^T \in R^{N \times l_i}$ $\quad \vec{z} = X^T \vec{w} + b$ $\quad z \in R^{N \times 1}$ $\quad A \in R^{N \times 1}$



$X^T \cdot \vec{w} + b$

CHOL-0강 [강화강의] Artificial-Neurons