# AI Programming 복습 2

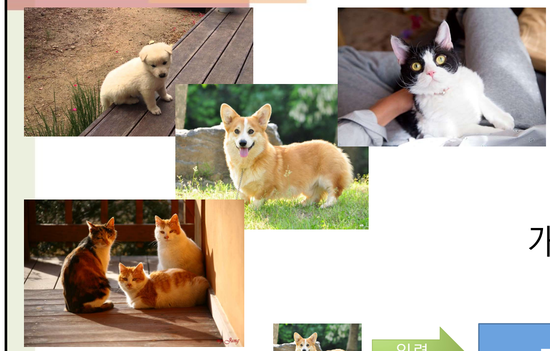
한국항공대학교
KOREA AEROSPACE UNIVERSITY

1

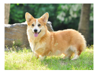
## 지난목차

- 분류 (classification)
- 인공 뉴런 (Artificial Neuron)
- 인공 뉴런 (Artificial Neuron)의 연결
- 인공신경망의 분류(classification)
- 인공신경망의 학습
- 객체 지향 프로그래밍 (Object Oriented Programming)
- 상속(Inheritance)
- Numpy

2

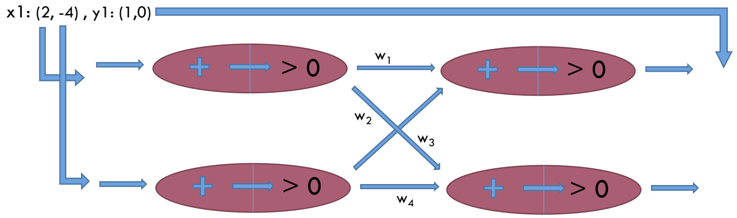


개 또는 고양이?

 입력 → F → 출력 → 개

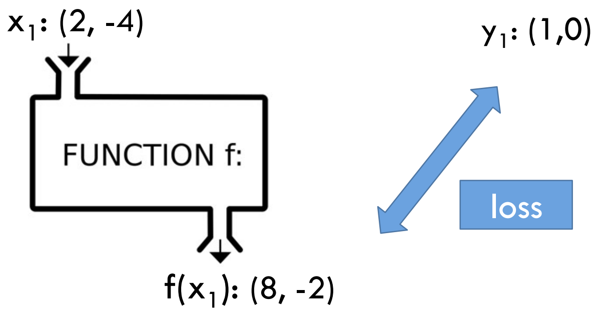
3

## 인공신경망을 이용한 분류




4

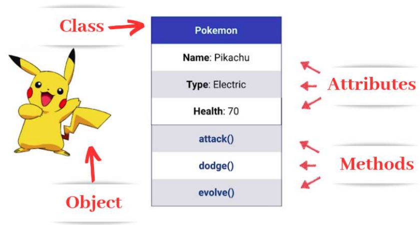
# 인공신경망을 이용한 분류




5

# 객체 (Object) 의 정의

- 객체 (Object):
  데이터(실체)와 그 데이터에 관련되는 동작(절차, 방법, 기능)을 모두 포함한 개념




6

# Numpy

Numpy

- 프로그램 라이브러리
- 수학의 함수 등 여러 계산을 편리하게 하기 위한 라이브러리

```
>>> import numpy as np
```

7

# 목차

- Linear Algrebra
- Derivative
- Gradient Descent
- Back-propagation
- Gradient Descent of Loss Function

8

# Linear Algebra

AI Programming

한국항공대학교
KOREA AEROSPACE UNIVERSITY

9

## Scalars

- 숫자

$$x \in \mathbb{R}$$

```
x = np.array(3.0)
y = np.array(2.0)

x + y, x * y, x / y, x ** y
```

10

## Vectors

- 숫자 여러 개

$$x \in \mathbb{R}^n$$

```
x = np.arange(4)
```

11

## Matrices

- 2차원 숫자

$$A \in \mathbb{R}^{n \times m}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

12

## Matrices

- 2차원 숫자

```
A = np.arange(20).reshape(5, 4)
```

```
array([[ 0.,  1.,  2.,  3.],
       [ 4.,  5.,  6.,  7.],
       [ 8.,  9., 10., 11.],
       [12., 13., 14., 15.],
       [16., 17., 18., 19.]])
```

13

## Matrix – Matrix Multiplication

$\mathbf{A} \in \mathbb{R}^{n \times k}$ $\mathbf{B} \in \mathbb{R}^{k \times m}$ $\mathbf{C} = \mathbf{AB}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{km} \end{bmatrix}$$

$\mathbf{C} \in \mathbb{R}^{n \times m}$

14

## Tensor

- 3차원 이상 숫자

```
X = np.arange(24).reshape(2, 3, 4)
```

```
array([[[ 0.,  1.,  2.,  3.],
        [ 4.,  5.,  6.,  7.],
        [ 8.,  9., 10., 11.]],

       [[12., 13., 14., 15.],
        [16., 17., 18., 19.],
        [20., 21., 22., 23.]]])
```

15

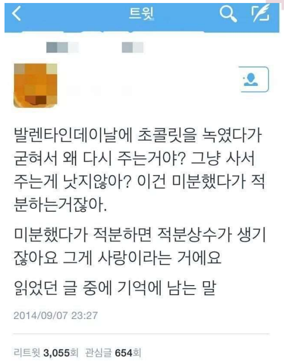
## Derivative

16

## Derivative

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

- 함수 f(x)
- 작은값 h
- 이에 대한 변화량!




17

## 기초 미분들

- $DC = 0$ ($C$ is a constant),
- $Dx^n = nx^{n-1}$ (the *power rule*, $n$ is any real number),
- $De^x = e^x$,
- $D\ln(x) = 1/x$.

18

## 기초 미분들

$$\frac{d}{dx}[Cf(x)] = C\frac{d}{dx}f(x),$$

$$\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x),$$

$$\frac{d}{dx}[f(x)g(x)] = f(x)\frac{d}{dx}[g(x)] + g(x)\frac{d}{dx}[f(x)],$$

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)\frac{d}{dx}[f(x)] - f(x)\frac{d}{dx}[g(x)]}{[g(x)]^2}.$$

19

## 변수가 여러 개 일때

$$y = f(x_1, x_2, \ldots, x_n)$$

$$\frac{\partial y}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}$$

20

## 편미분 표현

$$\frac{\partial y}{\partial x_i} = \frac{\partial f}{\partial x_i} = f_{x_i} = f_i = D_i f = D_{x_i} f.$$

21

## Gradient

$$f : \mathbb{R}^n \to \mathbb{R} \quad y = f(x_1, x_2, \ldots, x_n) \quad f(\mathbf{x})$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right]^\top$$

22

## Gradient 예시

$$f(x, y) = 2x + 5y$$

$$\nabla_{x,y} f(x, y) = \left[\frac{df(x,y)}{dx}, \frac{df(x,y)}{dy}\right]^T$$
$$= [2,5]^T$$

$$f(\mathbf{x}) = 2x_1 + 5x_2$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = [2,5]^T$$

23

## Chain Rule

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

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

24

## Multivariate Chain Rule

$y$ has variables $u_1, u_2, \ldots, u_m$,

$u_i$ has variables $x_1, x_2, \ldots, x_n$

$$\frac{dy}{dx_i} = \frac{dy}{du_1}\frac{du_1}{dx_i} + \frac{dy}{du_2}\frac{du_2}{dx_i} + \cdots + \frac{dy}{du_m}\frac{du_m}{dx_i}$$

25

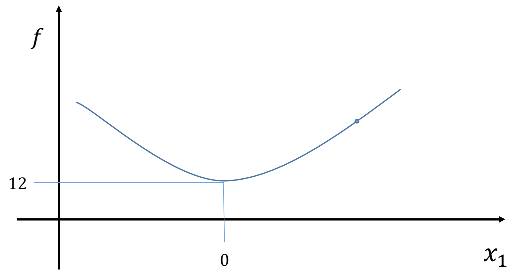
Minimization of Function

Gradient Descent

26

## Gradient Descent

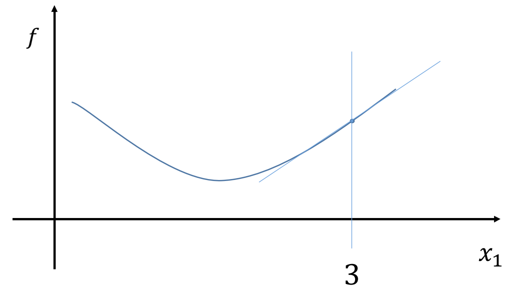
$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$




27

## Gradient Descent

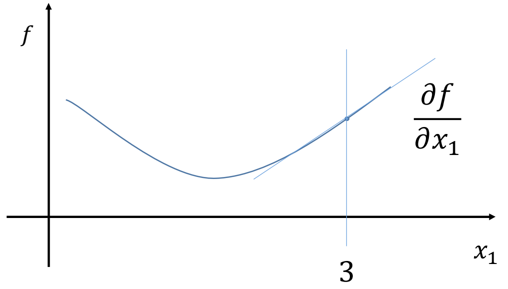
$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$




28

## Gradient Descent

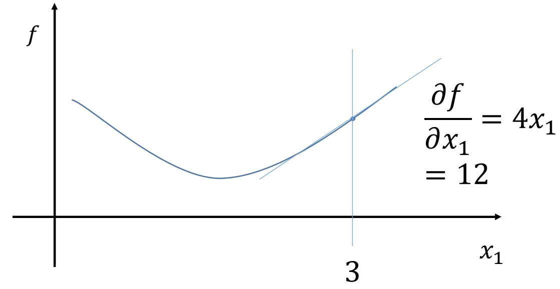
$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$




29

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

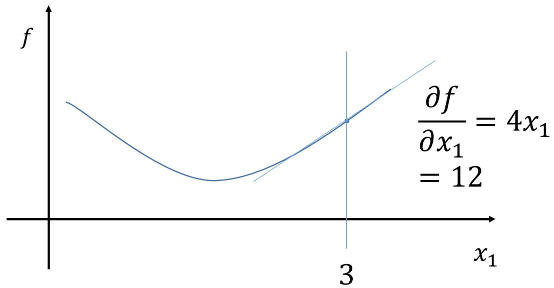



30

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

$f(X)$의 값을 줄이기 위한 $x_1$의 방향은?

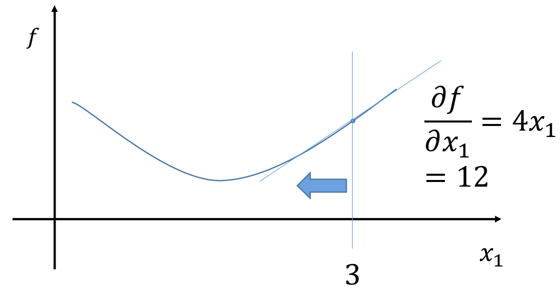



31

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

$f(X)$의 값을 줄이기 위한 $x_1$의 방향은?



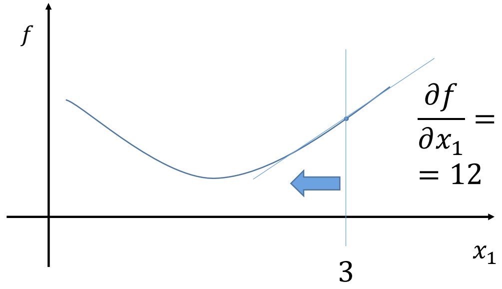

32

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

$f(X)$의 값을 줄이기 위한 $x_1$의 방향은?

3 → 2.99




$$\frac{\partial f}{\partial x_1} = 4x_1 = 12$$

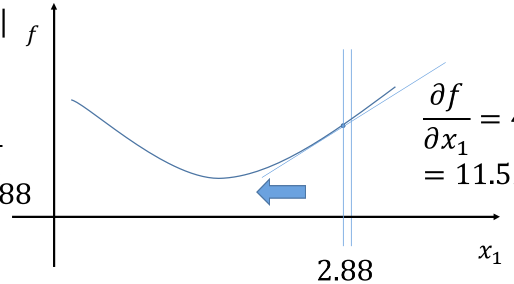
33

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

$f(X)$의 값을 줄이기 위한 $x_1$의 방향은?

$$3 \rightarrow 3 - 0.01 \times \frac{\partial f}{\partial x_1} = 3 - 0.01 \times 12 = 2.88$$




$$\frac{\partial f}{\partial x_1} = 4x_1 = 11.52$$

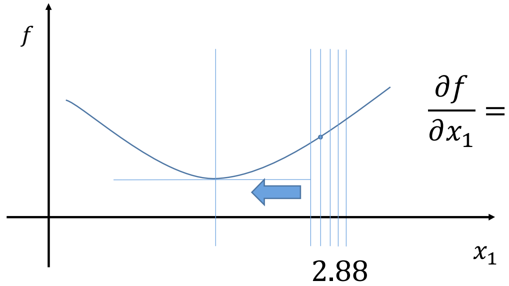
34

## Gradient Descent

$$f(X; x_2 = 2, x_3 = 3, x_4 = 4) = 2x_1^2 + 12$$

$f(X)$의 값을 줄이기 위한 $x_1$의 방향은?

$$x_1 \leftarrow x_1 - \eta \frac{\partial f}{\partial x_1}$$




$$\frac{\partial f}{\partial x_1} = 0$$

35

## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

36

## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

$$X \leftarrow X \; - \eta \, \nabla_X f$$



## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

$$X \leftarrow X \; - \eta \, \nabla_X f$$

$$X = [x_1, x_2, x_3, x_4]$$



## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

$$X \leftarrow X \; - \eta \, \nabla_X f$$

$$X = [x_1, x_2, x_3, x_4]$$

$$\nabla X = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial x_4}\right]$$



## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

$$X \leftarrow X \; - \eta \, \nabla_X f$$

$$X = [x_1, x_2, x_3, x_4]$$

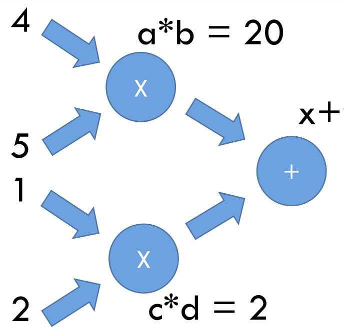
$$\nabla X = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial x_4}\right]$$

$$x_1 \leftarrow x_1 - \eta \frac{\partial f}{\partial x_1}$$

## Back-propagation

41

## Chain Rule

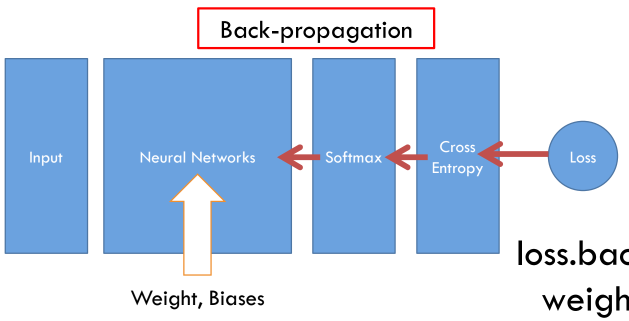



$$\frac{\partial f}{\partial a} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial a}$$

a.grad
b.grad
c.grad
d.grad

42

## 큰 그림



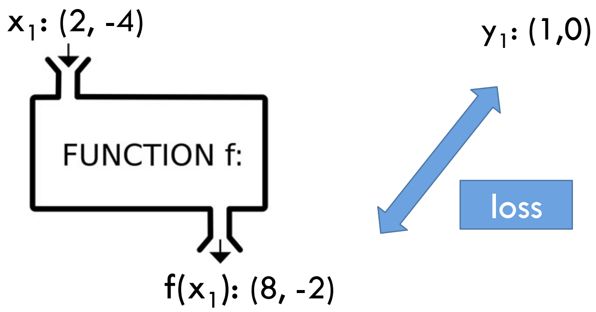

loss.backward()
weight.grad
bia.grad

43

## Gradient Descent of Loss Function

44

## 인공신경망을 이용한 분류




45

## Gradient Descent

$$f(X)$$

$f(X)$의 값을 줄이기 위한 $X$의 방향은?

$$X \leftarrow X - \eta \nabla_X f$$

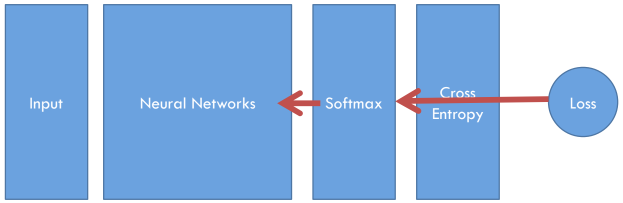
$$X = [x_1, x_2, x_3, x_4]$$

$$\nabla X = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial x_4}\right]$$

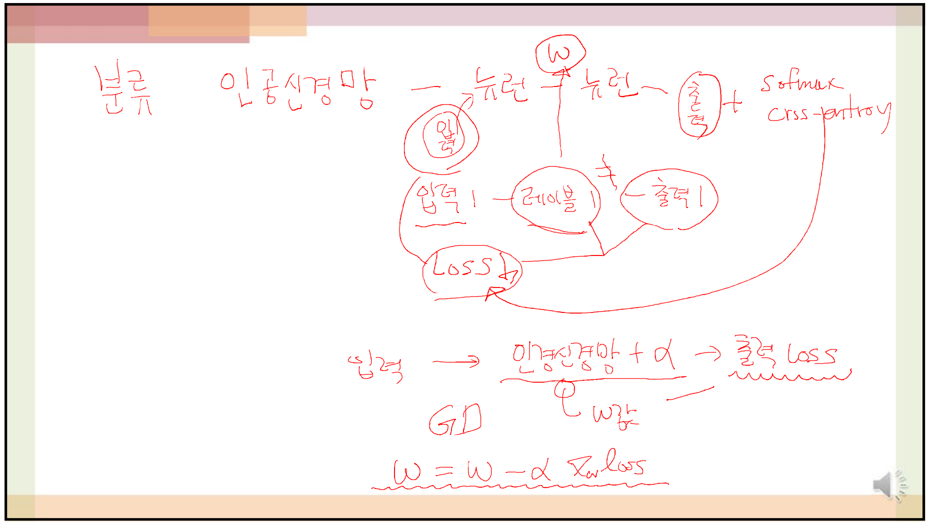
46

## Back-propagation

$$w = w - \eta \nabla_w \text{loss}$$




47




48

# 목차

- Linear Algrebra
- Derivative
- Gradient Descent
- Back-propagation
- Gradient Descent of Loss Function

49