인공신경망과딥러닝심화

Lecture 03. 딥러닝을 위한 기초 수학

동덕여자대학교
데이터사이언스 전공
권 범

목차

Copyright 2023. Dongduk Women's University all rights reserved.

시작하기 전에

❖ 딥러닝을 위한 기초 수학 (1/2)

- ‘딥러닝을 배운다’는 말에는 딥러닝의 실행법을 익히는 것뿐 아니라,
  딥러닝의 수학 원리를 공부한다는 의미도 담겨 있음
- 원리를 알아야 정확히 실행할 수 있기 때문에 딥러닝의 원리를 이해하는 것은
  좋은 코드를 만드는 것 이상으로 중요함
- 딥러닝의 수학 원리를 이해하기 위해서는 당연히 기본적인 수학 지식이 필요함
- 어떤 원리로 입력 값의 패턴을 분석하고 학습하는지 이해하려면
  그 배경이 되는 수학 연산을 살펴보아야 하고,
  여기에 사용되는 함수들을 알아야 하기 때문임

Copyright 2023. Dongduk Women's University all rights reserved.

시작하기 전에

❖ **딥러닝을 위한 기초 수학 (2/2)**

- 좋은 소식은 딥러닝 뒤에 있는 수학적 배경이
  다른 머신러닝과 비교했을 때 그다지 어렵지 않다는 것
- 딥러닝은 고등학교 수준의 수학만으로도 원리와 배경을 파악할 수 있음
- 조금 더 깊이 공부하더라도 대학교 교양 강좌 수준을 넘지 않는 범위에서
  딥러닝의 원리를 이해할 수 있음
- 이번 강의에서는 딥러닝을 이해하는 데 꼭 필요한 기초 수학을 공부할 예정

각 수학 공식이 딥러닝의 어느 부분에 활용되는지 참고하면서,
수학에 대한 두려움을 없애고 딥러닝 공부를 시작할 수 있길 바랍니다.

Copyright 2023. Dongduk Women's University all rights reserved.

01. 일차 함수, 기울기와 y-절편

Copyright 2023. Dongduk Women's University all rights reserved.

01. 일차 함수, 기울기와 y-절편

❖ **함수(Function)란?**

- 두 집합 사이의 관계를 설명하는 수학 개념
- 변수 x와 y가 있을 때, x가 변하면 이에 따라 y는 어떤 규칙으로 변하는지 나타냄
- 보통 함수를 나타낼 때는 function의 f와 변수 x를 사용해 $y = f(x)$라고 표시

Copyright 2023. Dongduk Women's University all rights reserved.

01. 일차 함수, 기울기와 y-절편

❖ 일차 함수란?

- y가 x에 관한 일차식으로 표현된 경우를 의미
- 예를 들어, 다음과 같은 함수식으로 나타낼 수 있음

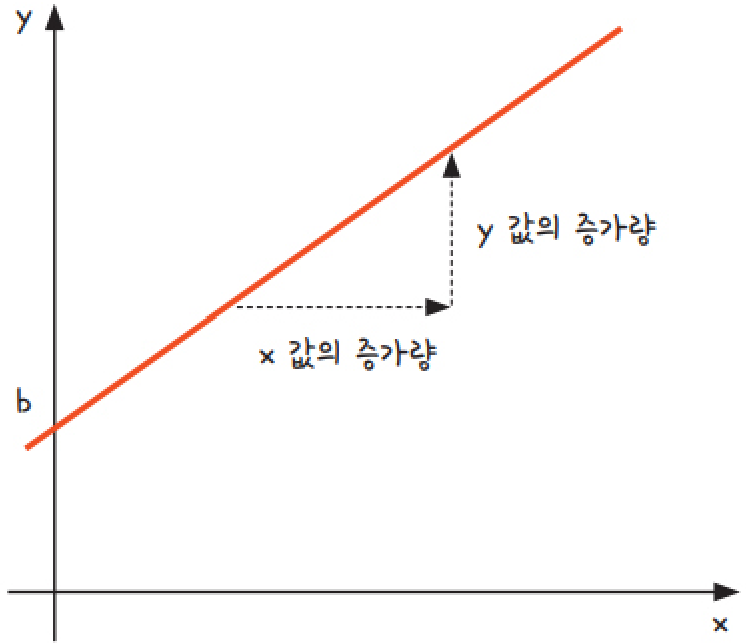
  $y = ax + b \ (a \neq 0)$
- x가 일차인 형태이며 x가 일차로 남으려면 a는 0이 아니어야 함

Copyright 2023. Dongduk Women's University all rights reserved.

01. 일차 함수, 기울기와 y-절편

❖ 기울기와 y-절편이란?

- 일차 함수식 $y = ax + b$에서 a는 **기울기**, b는 **y-절편**이라고 함
- 기울기는 기울어진 정도를 의미하는데, 아래 그림에서 x 값이 증가할 때 y 값이 어느 정도 증가하는지에 따라 그래프의 기울기 a가 정해짐
- y-절편은 그래프가 y축과 만나는 지점을 의미
- 아래 그림에서 y축과 만나는 y-절편이 바로 b



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 일차 함수, 기울기와 y-절편

- 딥러닝의 수학 원리를 배울 때 초반부터 이 식이 등장

  $y = ax + b \ (a \neq 0)$

- x가 주어지고 원하는 y 값이 있을 때 적절한 a와 b를 찾는 것

  이것이 바로 딥러닝을 설명하는 가장 간단한 표현

Copyright 2023. Dongduk Women's University all rights reserved.

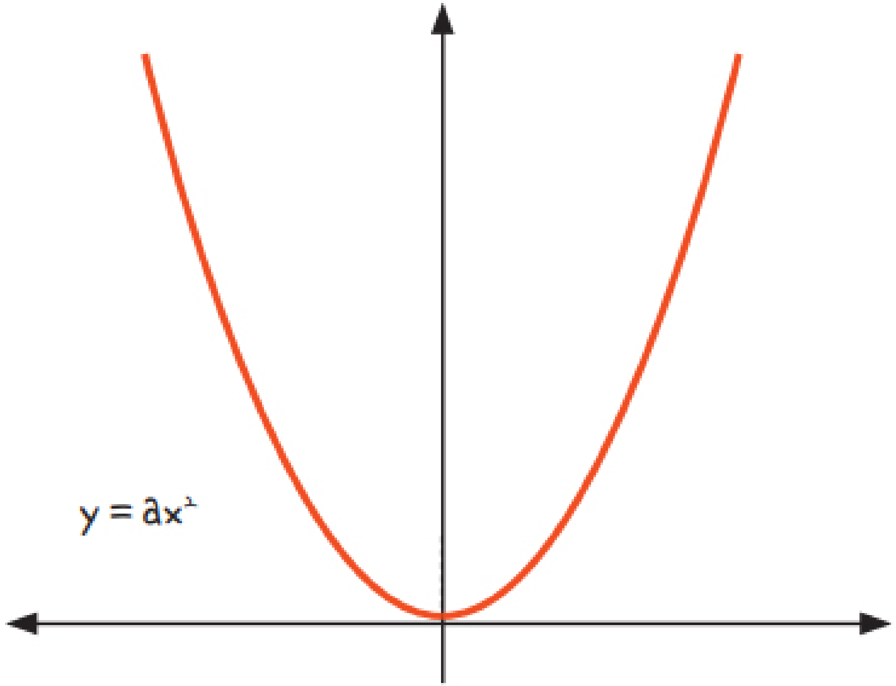
02. 이차 함수와 최솟값

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 이차 함수란?

- 이차 함수란 y가 x에 관한 이차식으로 표현되는 경우를 의미
- 다음과 같은 함수식으로 표현할 수 있음

  $y = ax^2\ (a \neq 0)$
- 이차 함수의 그래프는 아래 그림과 같이 포물선 모양
- $a > 0$이면 아래로 볼록한 그래프가 됨

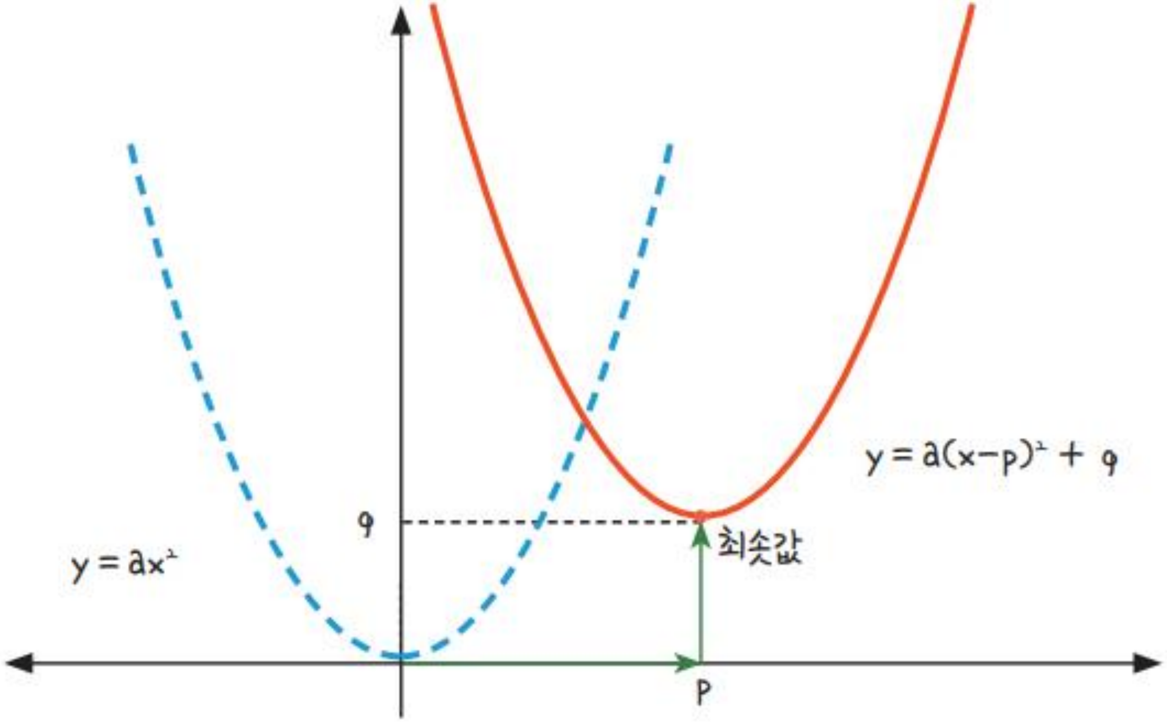


[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 이차 함수 그래프의 평행 이동과 최솟값 (1/2)

- $y = ax^2$의 그래프를 x축 방향으로 p만큼, y축 방향으로 q만큼 평행 이동시키면 아래 그림과 같이 점 (p, q)를 꼭짓점으로 하는 포물선이 됨
- 이때 포물선의 맨 아래에 위치한 지점이 **최솟값**이 되는데, 딥러닝을 실행할 때는 이 최솟값을 찾아내는 과정이 매우 중요함



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 이차 함수 그래프의 평행 이동과 최솟값 (2/2)

- 이 최솟값은 이후 강의에 소개할 '최소 제곱법' 공식으로 쉽게 알아낼 수 있음
- 딥러닝을 실제로 실행할 때 만나는 문제에서는 대부분 최소 제곱법을 활용할 수가 없음

여러 개의 입력을 처리하기에는 무리가 있기 때문임

딥러닝에서는 **미분**과 **기울기**를 이용함

Copyright 2023. Dongduk Women's University all rights reserved.

03. 미분, 순간 변화율과 기울기

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 미분에 대해 알아보기 (1/2)

- 딥러닝을 이해하는 데 가장 중요한 수학 원리는 **미분**이라고 할 수 있음
- 조금 전 딥러닝은 결국 일차 함수의 a와 b 값을 구하는 것인데,
  a와 b 값은 이차 함수 포물선의 최솟값을 구하는 것

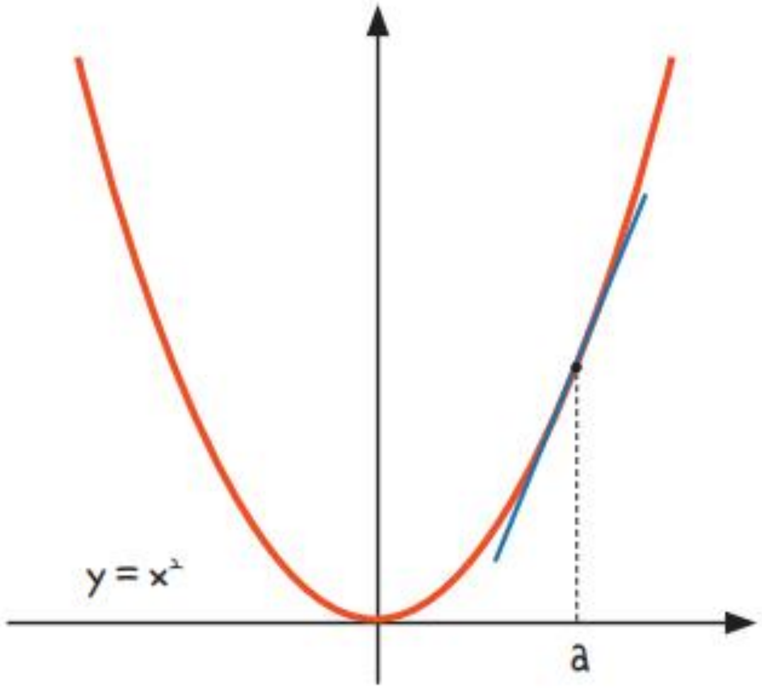
이 최솟값을 미분으로 구하기 때문에, 미분이 딥러닝에서 중요한 것

미분과 기울기의 개념을 먼저 알아보겠습니다.

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 미분에 대해 알아보기 (2/2)

- 아래 그림과 같이 $y = x^2$이라는 그래프가 있다고 가정
- x축에 있는 한 점 a에 대응하는 y의 값은 a^2
- 이때 a가 오른쪽이나 왼쪽으로 조금씩 이동한다고 상상
- 이에 따라 y도 조금씩 변화할 것



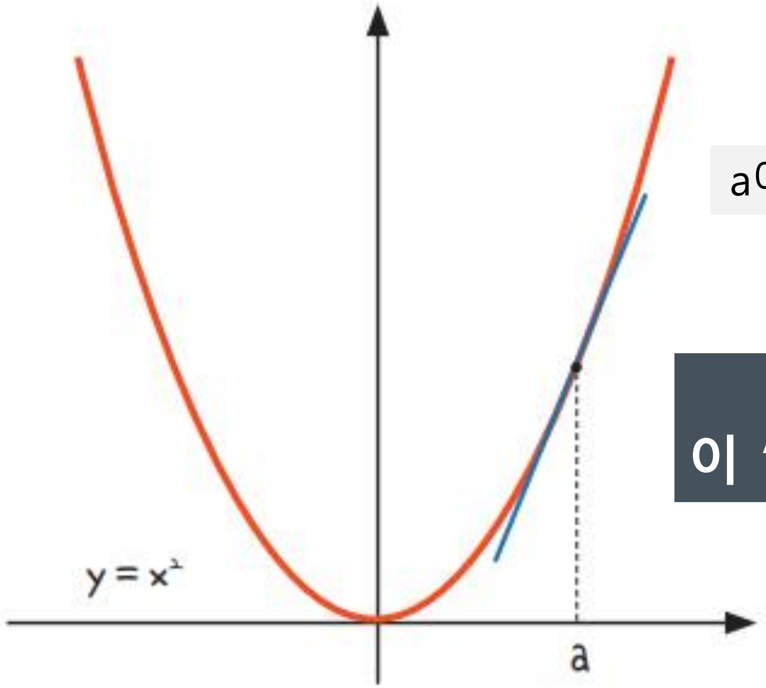
[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

03. 미분, 순간 변화율과 기울기

❖ 순간 변화율과 기울기

- 상상력을 조금 더 발휘해 이번에는 a가 미세하게 '0에 가까울 만큼' 움직였다고 가정
- y 값 역시 매우 미세하게 변화를 할 텐데, 이번에는 너무 미세해서 실제로 움직이는 것이 아니라 방향만 드러내는 정도의 순간적인 변화만 있을 것 이 순간의 변화를 **순간 변화율**이라고 함
- 순간 변화율은 어느 쪽을 향하는 방향성을 지니고 있으므로, 이 방향을 따라 직선을 길게 그려 주면 그래프와 맞닿는 접선이 그려짐 이 선이 바로 이 점에서의 **기울기**가 됨



a에서의 순간 변화율이 기울기입니다.

미분을 한다는 것은 쉽게 말해 이 '순간 변화율'을 구한다는 것입니다.

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

03. 미분, 순간 변화율과 기울기

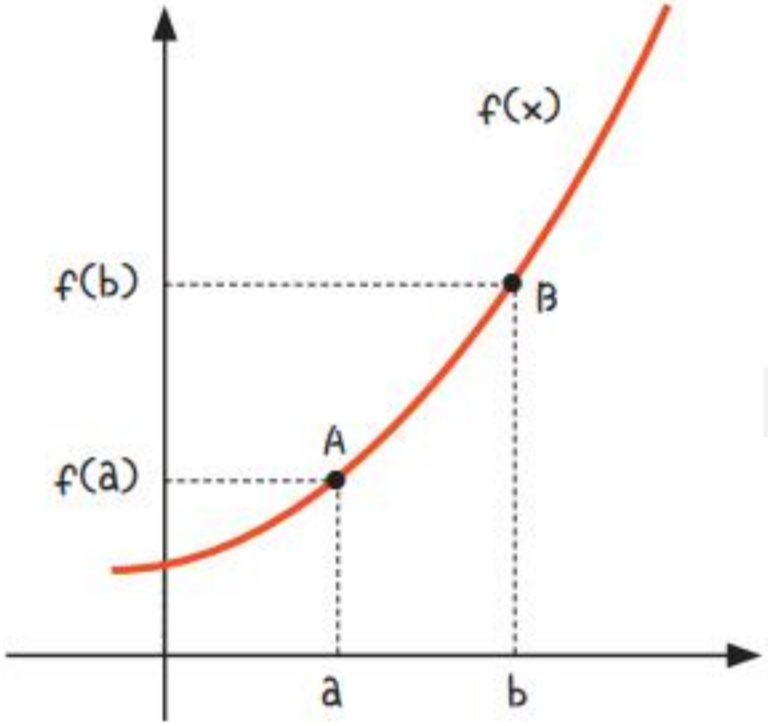
❖ 미분 계수

- 어느 순간에 어떤 변화가 일어나고 있는지 숫자로 나타낸 것을 **미분 계수**라고 하며, 이 미분 계수는 곧 그래프에서의 기울기를 의미
- 이 기울기가 중요한 것은 **기울기가 0일 때**, 즉 x축과 평행한 직선으로 그어질 때가 바로 그래프에서 **최솟값인 지점이 되기 때문**임

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 순간 변화율을 구하는 방법 (1/6)

- 어떤 함수 $f(x)$가 아래 그림과 같이 주어졌다고 하자
- 이 함수에 x축 위의 두 실수 a와 b를 대입하면 두 점 A, B는 그림과 같이 각각 A$(a, f(a))$, B$(b, f(b))$에 해당하는 곳에 표시



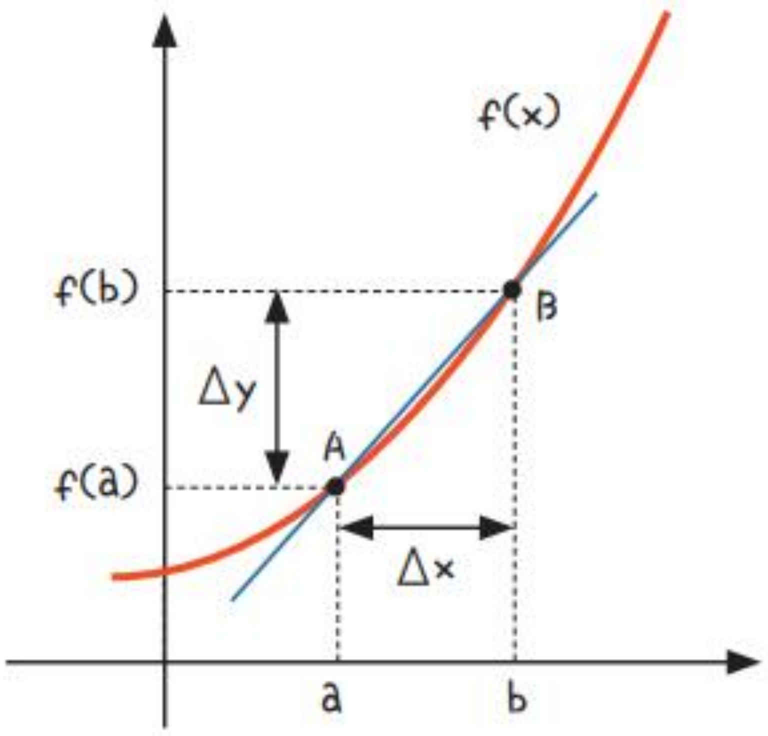
함수 f(x)의 x축 위에 두 실수 a와 b를 대입

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 순간 변화율을 구하는 방법 (2/6)

- 이때 두 점 A와 B를 이어 직선을 만들면 아래 그림과 같이 두 점 A와 B를 지나는 직선의 기울기가 그려짐
- 여기서 Δ(델타)는 변화량을 나타내는 기호



A와 B를 지나는 직선은 이 두 점 간의 기울기,
즉 평균 변화율을 의미

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 순간 변화율을 구하는 방법 (3/6)

- 이 그래프에서 x 값의 증가량은 $b-a$이고, y 값의 증가량은 $f(b)-f(a)$
- 이를 Δ를 써서 표현하면 x 값의 증가량은 Δx로,
  y 값의 증가량은 $f(a+\Delta x)-f(a)$로 나타낼 수 있음

$$\text{직선의 기울기} = \frac{y\ \text{값의 증가량}}{x\ \text{값의 증가량}}$$

- A와 B를 지나는 직선의 기울기는 다음과 같이 표현할 수 있음

$$\text{직선 AB의 기울기} = \frac{y\ \text{값의 증가량}}{x\ \text{값의 증가량}} = \frac{f(b)-f(a)}{b-a} = \frac{f(a+\Delta x)-f(a)}{\Delta x}$$

Copyright 2023. Dongduk Women's University all rights reserved.

03. 미분, 순간 변화율과 기울기

❖ 순간 변화율을 구하는 방법 (4/6)

- 이때 직선 AB의 기울기를 A와 B 사이의 **평균 변화율**이라고도 함
- 미분을 배우고 있는 우리에게 필요한 것은 **순간 변화율**
- 순간 변화율은 x의 증가량(=Δx)이 0에 가까울 만큼
  아주 작을 때의 순간적인 기울기를 의미하므로,
  극한(Limit) 기호를 사용해 다음과 같이 나타냄

$$\lim_{\Delta x \to 0} \frac{f(a+\Delta x) - f(a)}{\Delta x}$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 순간 변화율을 구하는 방법 (5/6)

- 여기서 $\lim_{\Delta x \to 0}$ 는 'x의 증가량이 0에 가까울 만큼 작을 때'라는 뜻

기울기

$$기울기 = \frac{y \text{ 값의 증가량}}{x \text{ 값의 증가량}}$$

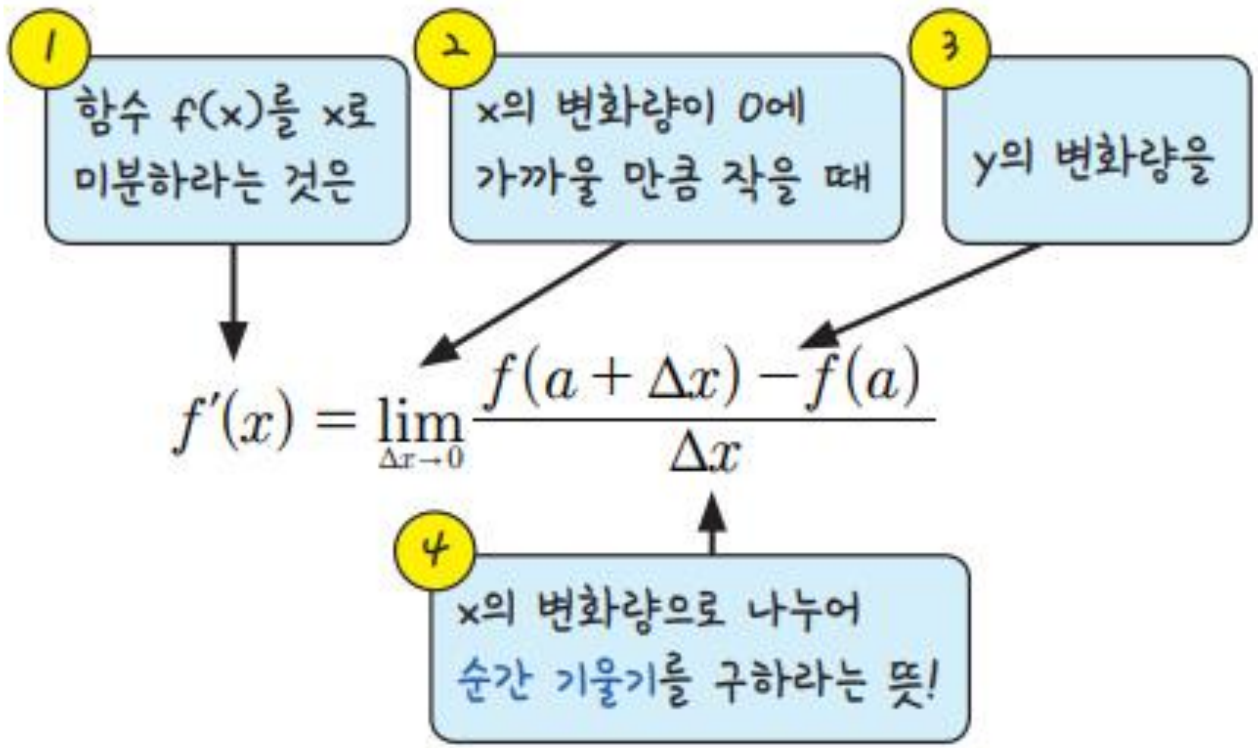
순간 기울기

$$순간\ 기울기 = \lim_{\Delta x \to 0} \frac{y \text{ 값의 증가량}}{x \text{ 값의 증가량}} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx}$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 순간 변화율을 구하는 방법 (6/6)

- "함수 $f(x)$를 미분하라"는 것을 $f'(x)$ 또는 $\frac{d}{dx}f(x)$로 표기하는데, 함수 $f(x)$를 미분하는 공식을 알기 쉽게 정리하면 다음과 같음



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 미분의 기본 공식

- 다음은 딥러닝을 공부하는 과정 중에 자주 만나게 되는 중요한 다섯 가지 미분의 기본 공식

미분의 기본 공식

1 | $f(x) = x$일 때 $f'(x) = 1$

2 | $f(x) = a$에서 a가 상수일 때 $f'(x) = 0$

3 | $f(x) = ax$에서 a가 상수일 때 $f'(x) = a$

4 | $f(x) = x^a$에서 a가 자연수일 때 $f'(x) = ax^{a-1}$

5 | $f(g(x))$에서 $f(x)$와 $g(x)$가 미분 가능할 때 $\{f(g(x))\}' = f'(g(x)) \times g'(x)$

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

04. 편미분

Copyright 2023. Dongduk Women's University all rights reserved.

04. 편미분

❖ 편미분이란? (1/3)

- 미분과 더불어 딥러닝을 공부할 때 가장 자주 접하게 되는 또 다른 수학 개념은 바로 편미분
- 미분과 편미분 모두 '미분하라'는 의미에서는 다를 바가 없음

"편미분"

**여러 가지 변수가 식 안에 있을 때, 모든 변수를 미분하는 것이 아니라
우리가 원하는 한 가지 변수로만 미분하고 그 외에는 모두 상수로 취급하는 것**

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 편미분이란? (2/3)

- 예를 들어, $f(x) = x$와 같은 식을 미분할 때는 변수가 x 하나뿐이어서 미분하라는 의미에 혼란이 없음
- 하지만 다음 식을 보면 여기에는 변수가 x와 y, 이렇게 두 개 있음

$$f(x, y) = x^2 + xy + a \ (a\text{는 상수})$$

- 이 중 어떤 변수로 미분해야 하는지 정해야 하므로 편미분을 사용하는 것
- 만일 이 식처럼 여러 변수 중에서 x에 관해서만 미분하고 싶다면,
  함수 f를 'x에 관해 편미분하라'고 하며 다음과 같이 식을 씀

$$\frac{\partial f}{\partial x}$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 편미분이란? (3/3)

- 앞에 나온 함수 $f(x, y) = x^2 + yx + a$를 x에 관해 편미분하는 과정은 어떻게 될까?
- 먼저 바로 앞에서 배운 미분의 성질 4에 따라 x^2항은 $2x$가 됨
- 미분법의 기본 공식 3에 따라 yx는 y가 됨
- 마지막 항 a는 미분의 성질 1에 따라 0이 됨
- 이를 정리하면 다음과 같이 나타낼 수 있음

$f(x, y) = x^2 + xy + a$ 일때

$$\frac{\partial f}{\partial x} = 2x + y$$

Copyright 2023. Dongduk Women's University all rights reserved.

05. 지수와 지수 함수

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 지수란?

- 다음과 같은 형태를 의미

$$a^{\square}$$

- ✓ 여기서 a를 '밑'이라 하고 $\square$를 '지수'라고 함
- ✓ a를 $\square$만큼 반복해서 곱한다는 뜻

$$\text{밑}^{\text{지수}}$$

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

05. 지수와 지수 함수

❖ 지수 함수란? (1/2)

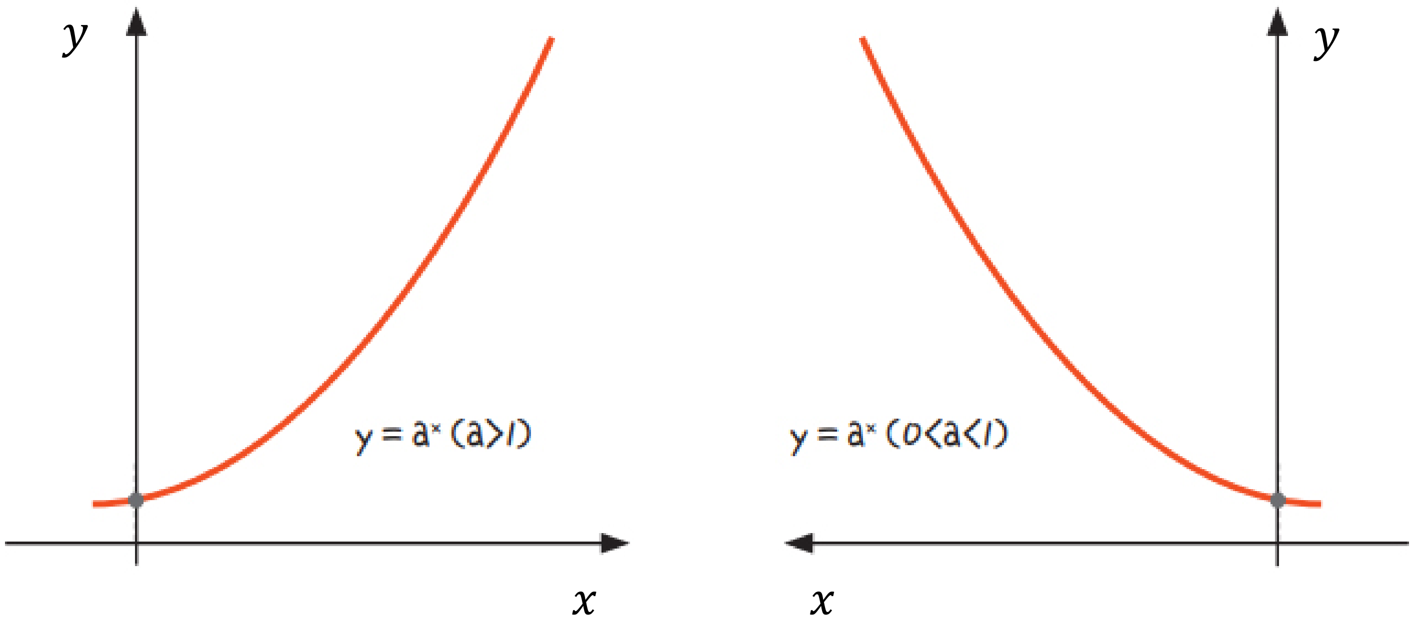
- 변수 x가 지수 자리에 있는 경우를 의미
- 식으로 나타내면 다음과 같은 형태

$$y = a^x \ (a \neq 1, a > 0)$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 지수 함수란? (2/2)

- 지수 함수에서는 밑(a) 값이 무엇인지가 중요함
- 이 값이 1이면 함수가 아님
- 또한, 0보다 작으면 허수를 포함하게 되므로 안 됨
- 밑의 값은 $a > 1$이거나 $0 < a < 1$, 둘 중 하나가 되어야 함
- 이 두 가지 경우의 그래프는 각각 아래 그림과 같음



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

06. 시그모이드 함수

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 시그모이드 함수에 대해 알아보기 (1/3)

- 딥러닝의 내부를 보면 입력받은 신호를 얼마나 출력할지를 계산하는 과정이 무수히 반복
- 이때 출력 값으로 얼마나 내보낼지를 계산하는 함수를 **활성화 함수**라고 함
- 활성화 함수는 딥러닝이 발전함에 따라 여러 가지 형태로 개발되어 왔는데, 그중 가장 먼저 배우는 중요한 함수가 바로 **시그모이드 함수**
- 시그모이드 함수는 지수 함수에서 밑 값이 자연 상수 e인 함수를 의미
- 자연 상수 e는 '자연 로그의 밑', '오일러의 수' 등 여러 이름으로 불리는데, 파이(π)처럼 수학에서 중요하게 사용되는 무리수이며 그 값은 대략 2.718281828...

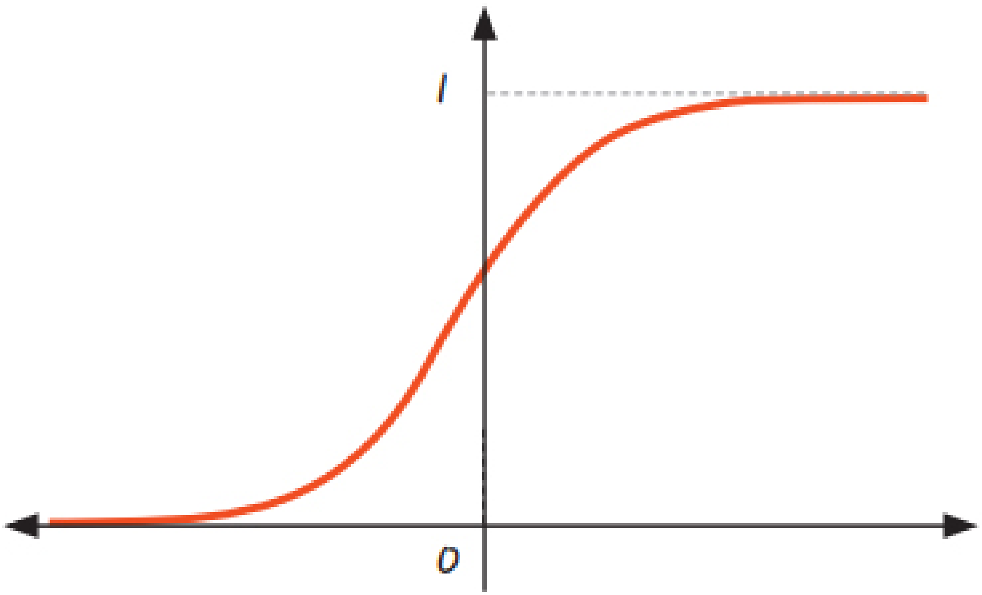
Copyright 2023. Dongduk Women's University all rights reserved.

❖ 시그모이드 함수에 대해 알아보기 (2/3)

- 자연 상수 e가 지수 함수에 포함되어 분모에 들어가면 시그모이드 함수가 되는데, 이를 식으로 나타내면 다음과 같음

$$f(x) = \frac{1}{1+e^{-x}}$$

- 시그모이드 함수를 그래프로 그려 보면 아래 그림과 같이 S자 형태로 나타남

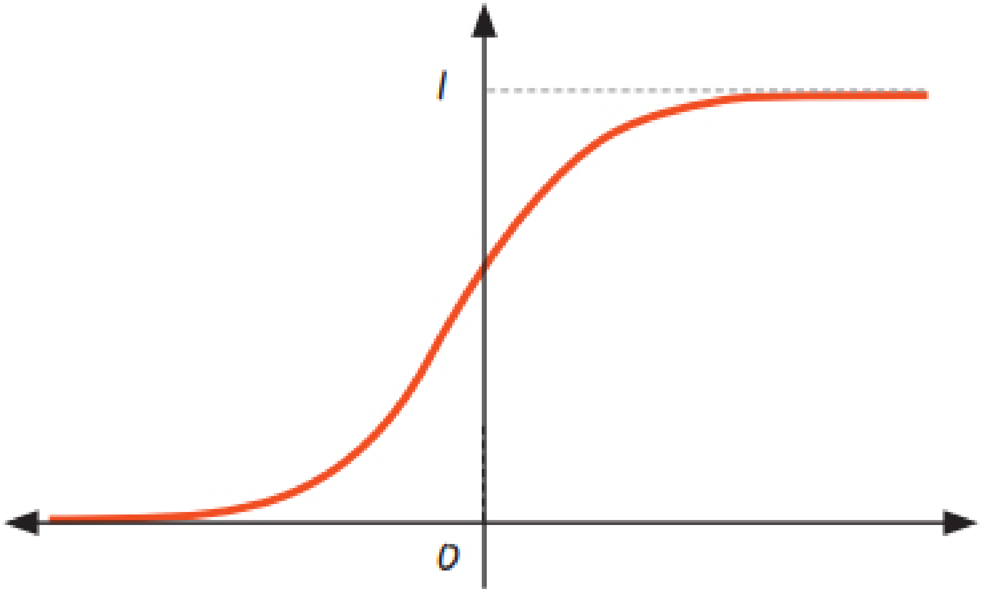


[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 시그모이드 함수에 대해 알아보기 (3/3)

- x가 큰 값을 가지면 $f(x)$는 1에 가까워지고,
  x가 작은 값을 가지면 $f(x)$는 0에 가까워짐
- S자 형태로 그려지는 이 함수의 속성은 0 또는 1,
  두 개의 값 중 하나를 고를 때 유용하게 쓰임



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

07. 로그와 로그 함수

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 로그에 대해 알아보기 (1/2)

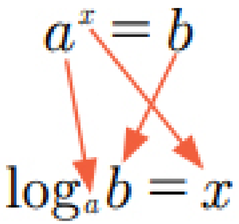
- 우선, **로그**를 이해하려면 먼저 지수부터 이해해야 함
- a를 x만큼 거듭제곱한 값이 b라고 할 때,
  이를 식으로 나타내면 다음과 같음

  $$a^x = b$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 로그에 대해 알아보기 (2/2)

- 이때 a와 b를 알고 있는데 x를 모른다고 가정
- x는 과연 어떻게 구할 수 있을까?
- 이 x를 구하기 위해 사용하는 방법이 로그
- 영어로 Logarithm이라고 하는데 앞 세 글자 log를 사용해서 표시하며,
  지수식에서 a와 b의 위치를 다음과 같이 바꾸어 쓰면 됨



- 로그가 지수와 이렇게 밀접한 관계가 있듯이 로그 함수 역시 지수 함수와 밀접한 관계에 있음
- 바로 **역함수**의 관계   역함수는 x와 y를 서로 바꾸어 가지는 함수

[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 로그 함수와 지수 함수의 관계 (1/2)

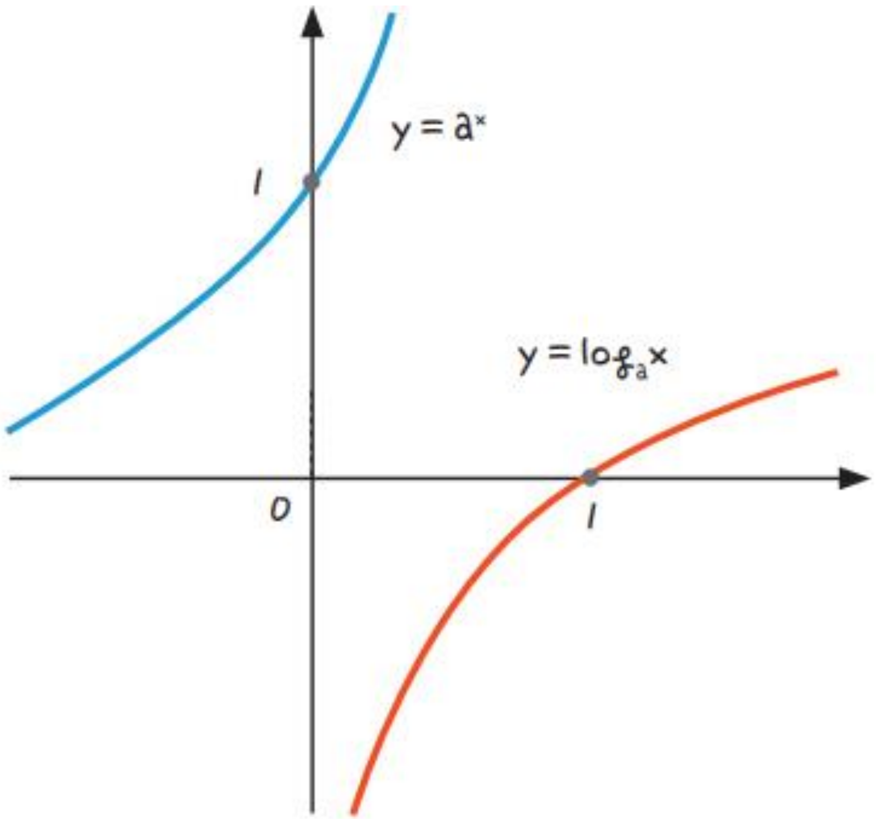
- 지수 함수 $y = a^x$ $(a \neq 1, a > 0)$는 로그 정의를 따라 $x = \log_a y$로 바꿀 수 있음
- 역함수를 만들기 위해 x와 y를 서로 바꾸어 주면 됨
- 다음 식이 바로 로그 함수의 형태

$$y = \log_a x$$

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 로그 함수와 지수 함수의 관계 (2/2)

- 역함수의 그래프는 $y = x$에 대해 대칭인 선으로 나타남
- 아래 그림은 지수 함수 $y = a^x$의 그래프를 $y = x$에 대칭으로 이동시킨 로그 함수 $y = \log_a x$의 그래프를 보여 줌



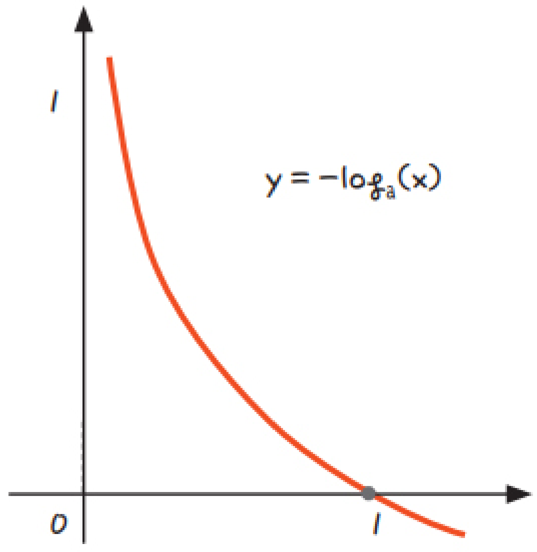
[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

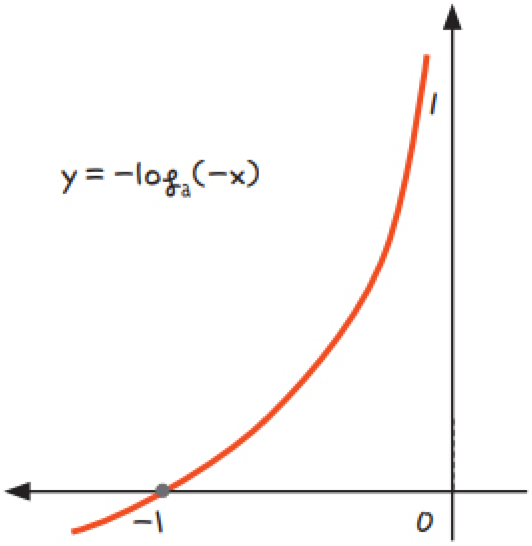
❖ 로그 함수 (1/2)

- 향후 로지스틱 회귀를 배울 때, 우리는 x가 1에 가까워지거나 0에 가까워질수록 오차가 커지는 그래프가 필요함
- 이러한 그래프를 만들기 위해 $y = \log_a x$를 x축 또는 y축으로 대칭 이동하거나 알맞게 평행 이동하면 다음과 같음

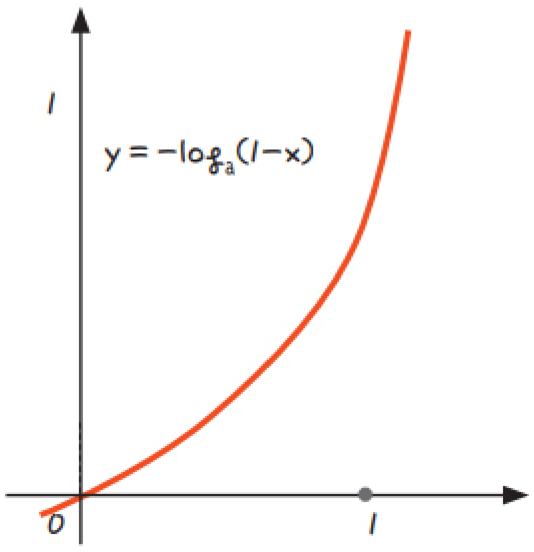
①: x축에 대해 대칭 이동



②: x축과 y축에 대해 대칭 이동



③: ②의 그래프를 x축 오른쪽 방향으로 1만큼 평행 이동

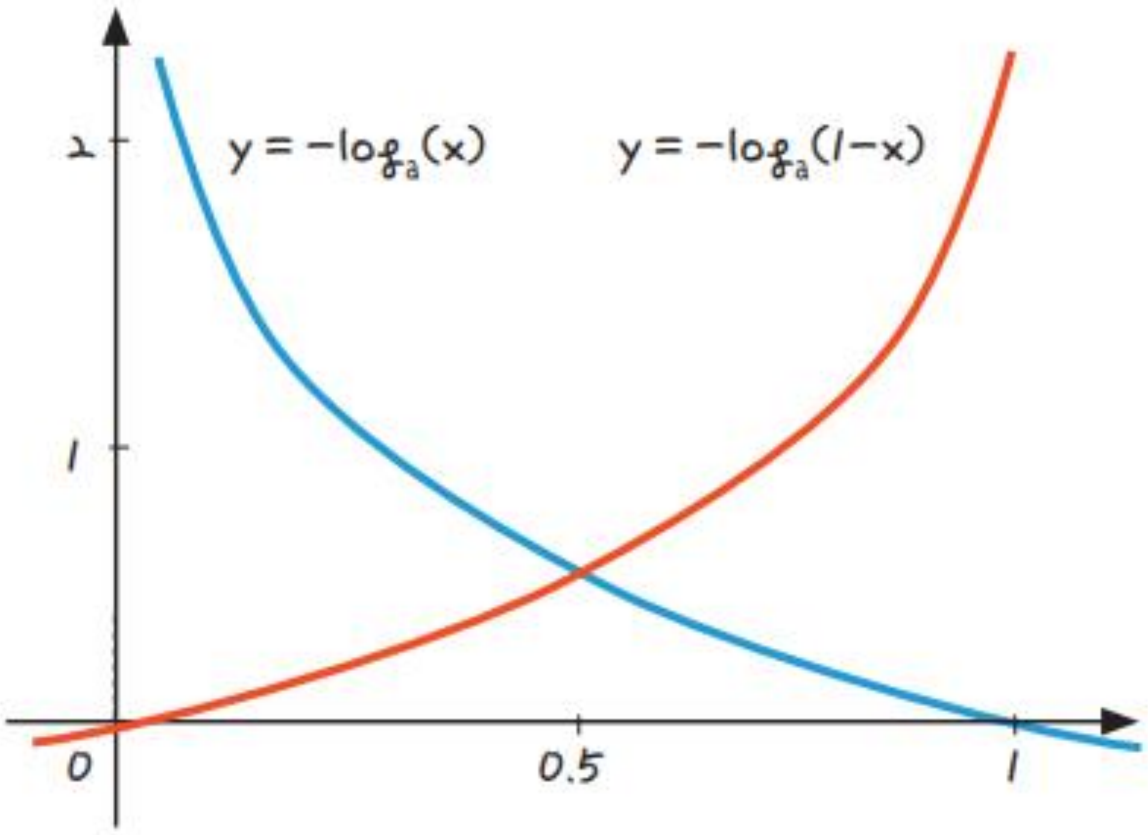


[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ **로그 함수 (2/2)**

- ①과 ③을 함께 나타낸 그래프



[사진출처] 모두의 딥러닝 (출판사: 길벗, 저자: 조태호)

Copyright 2023. Dongduk Women's University all rights reserved.

❖ 마무리하기 전에

- 지금까지 설명한 일차 함수, 이차 함수, 미분, 편미분, 지수 함수, 시그모이드 함수 그리고 로그 함수 이렇게 일곱 가지를 알고 있으면 앞으로 배울 내용을 모두 이해할 수 있음
- 여기에 합을 표현하기 위해 만들어진 ∑(시그마) 기호가 종종 나옴
- $\sum_{i=1}^{n} F(i)$라고 하면 i를 1부터 n까지 $F(i)$에 대입해 더하라는 뜻
- 즉, $F(1) + F(2) + F(3) + \cdots + F(n)$이 됨
- 위 내용까지 숙지하고 있으면 앞으로 다룰 모든 딥러닝 예제를 이해하고 실행하는 데는 문제없음

Copyright 2023. Dongduk Women's University all rights reserved.

끝맺음

- ❖ 01. 일차 함수, 기울기와 y-절편
- ❖ 02. 이차 함수와 최솟값
- ❖ 03. 미분, 순간 변화율과 기울기
- ❖ 04. 편미분
- ❖ 05. 지수와 지수 함수
- ❖ 06. 시그모이드 함수
- ❖ 07. 로그와 로그 함수

Copyright 2023. Dongduk Women's University all rights reserved.

THANK YOU!
Q & A

- Name: 권범
- Office: 동덕여자대학교 인문관 B821호
- Phone: 02-940-4752
- E-mail: bkwon@dongduk.ac.kr

Copyright 2023. Dongduk Women's University all rights reserved.