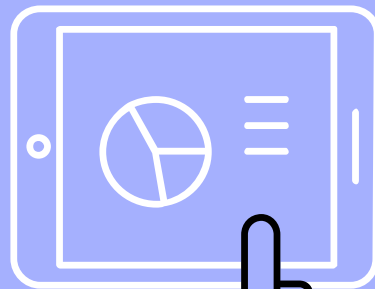
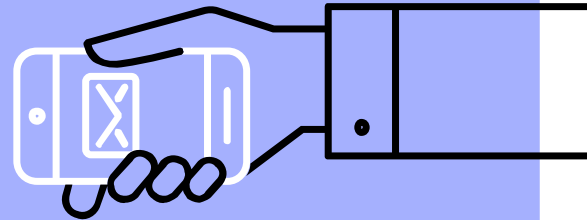


이론과 기술기



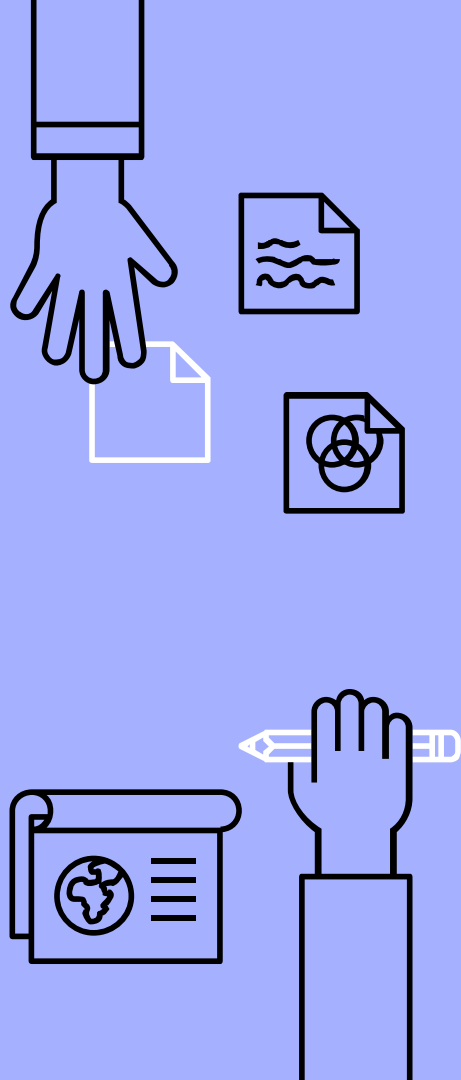
미분에 대해서 알아보자!!

해석미분

- 고등학생 때 배운 미분
- 이론에 입각해 미분 수행
- 오차를 포함하지 않는
"진정한 미분" 값을 구해줌

수치미분

- 프로그램적 계산을 통해
미분 수행
- 약간의 오차 있음
- 아주 작은 차분(함수들의 값
차이)으로 미분하는 것
- 근사치로 계산하는 방법



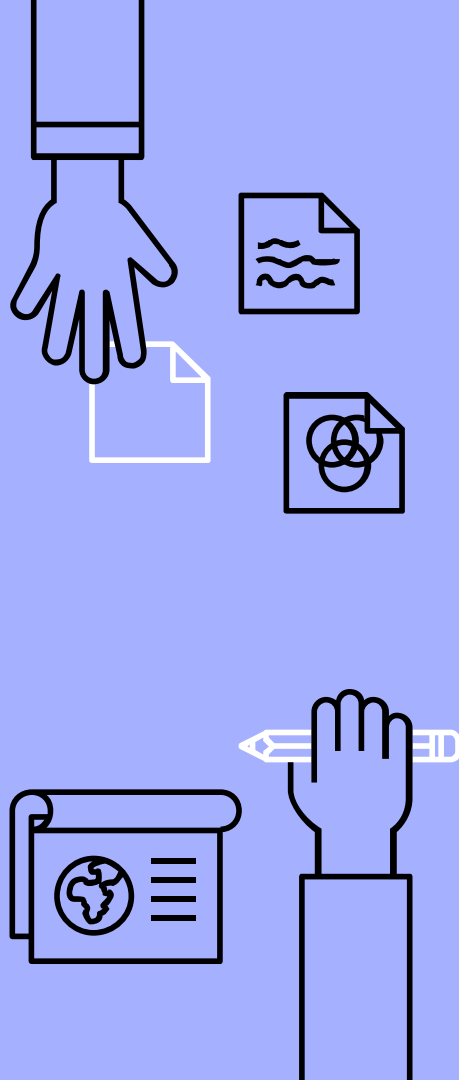
도함수 & 미분법?

도함수

- 어떤 함수의 정의역(함수의 X 값)속 각 점에서 독립 변수의 변화량과 함수값의 변화량의 비율의 극한으로 구성된 집합으로 치역(함수의 X 값에 대한 Y 값)이 구성된 함수

미분법

- 미분하는 작업
- 도함수를 구하는 일
- 함수에 대해 특정 순간의 변화량, x 의 변화가 함수 $f(x)$ 를 얼마나 변화시켰는지 나타냄



미분 공식

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

-> '진정한 미분': x 위치의 함수의 기울기

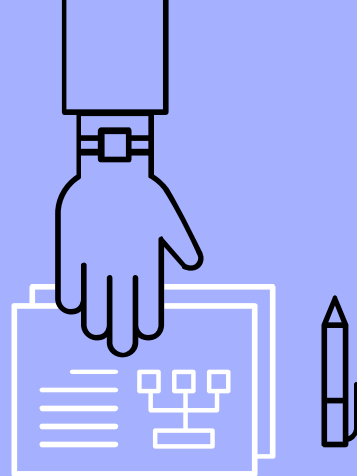
```
# 수치 미분 파이썬으로 구현
def numerical_diff(f,x):
    h = 10e-50
    return (f(x+h)-f(x))/h
```

-> x+h와 x 사이의 함수의 차분 계산, 오차가 있음

```
import numpy as np
print(np.float32(1e-50))
```

```
h = 1e-4
print(h)
```

-> 결과 : 0.0 / 0.0001



차분과 중심차분(중앙 차분)

차분

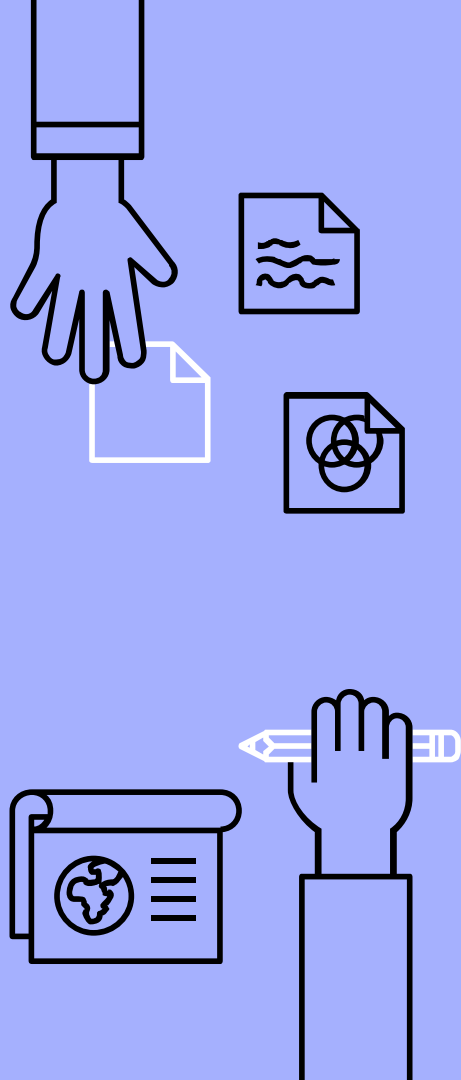
- 함수의 값에서 그 하나 앞의 함수의 값을 뺀 것

중앙 차분

- $f(x+h)$ 값에서 $f(x-h)$ 를 뺀 값
- x 를 기준으로 그 전 후 차분을 계산
- 오차를 줄이기 위해 사용

중앙차분을 적용한 식

```
def numerical_diff(f, x):  
    h = 1e-4  
    return (f(x+h)-f(x-h)) / (2*h)
```



편미분

- 입력변수(독립변수)가 2개 이상인 다변수 함수에서 미분하고자 하는 변수를 제외한 나머지 변수들을 상수취급해서 미분 / 3차원 그래프가 그려짐
- 편미분 계산 시 변수가 여러이므로 어떤 변수에 대한 미분인지를 정해줘야 함

문제 1 : $x_0 = 3, x_1 = 4$ 일 때, x_0 에 대한 편미분 $\frac{\partial f}{\partial x_0}$ 를 구하라.

```
>>> def function_tmp1(x0):  
...     return x0*x0 + 4.0**2.0  
...  
>>> numerical_diff(function_tmp1, 3.0)  
6.0000000000000378
```

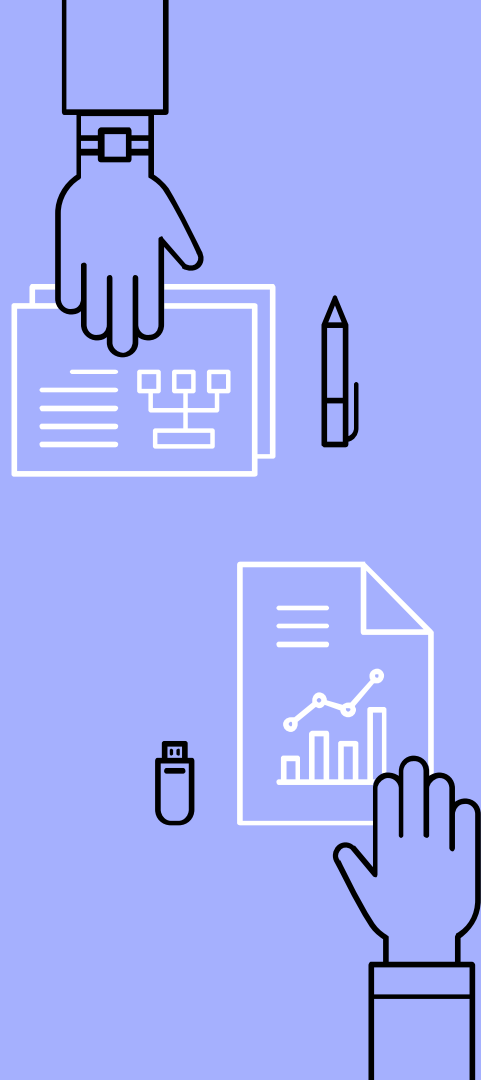
문제 2 : $x_0 = 3, x_1 = 4$ 일 때, x_1 에 대한 편미분 $\frac{\partial f}{\partial x_1}$ 를 구하라.

```
>>> def function_tmp2(x1):  
...     return 3.0**2.0 + x1*x1  
...  
>>> numerical_diff(function_tmp2, 4.0)  
7.9999999999999119
```



기울기

- 모든 변수의 편미분을 벡터로 정리한 것
 - 각 지점에서 낮아지는 방향
 - 각 장소에서 함수의 출력 값을 가장 크게 줄이는 방향
-
- 기울기 < 0 : 오른쪽으로 이동할수록 값이 작아짐
 - 기울기 > 0 : 왼쪽으로 이동할수록 값이 작아짐



경사법

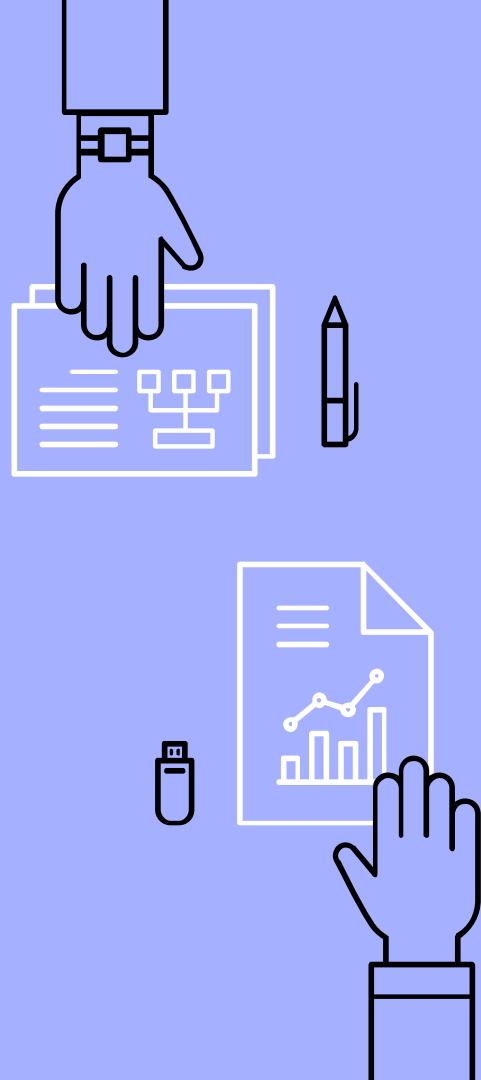
- 기울기를 잘 이용해 손실 함수의 값(나뭇)을 최소화 하는 방법을 찾으려는 것

-> 경사하강법(최솟값 찾음)

-> 경사상승법(최댓값 찾음)

※ 함수가 극솟값, 최솟값, 안장점이 되는 장소에서는
기울기가 0이다.

※ 안장점 : 어느 부분에서 보면 극대값이지만 다른
부분에서 보면 극소값이 되는 점



경사법 in Deep Learning

LR, learning Rate

- 학습률이 너무 크거나 작으면 좋은 결과를 얻을 수 없다.
- 학습률이 너무 크면 큰 값으로 발산해 버리고 작으면 갱신되지 않은 채 끝난다.

