

# 2017-06-28 (2주차-1)

## 1. 파이썬으로 공부하는 미적분과 최적화

### a. 함수

- 변수와 함수
- 예측 문제와 함수

예측(prediction) 문제는 독립 변수  $x$  를 입력으로 하여 원하는 종속 변수  $y$ 와 가장 비슷한 값을 만드는 함수  $f$ 를 찾는 문제라고 할 수 있다.

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

- 함수의 그래프
- 기울기

$y$ 의 변화량 /  $x$ 의 변화량  $\rightarrow$   $x$ 의 변화량이 얼마인가에 따라 기울기가 달라짐. 이를 해결하기 위해  $x$ 의 변화량이 0으로 근접할때의 변화율을 기울기(slope)라 정의

- 데이터분석에서 많이 쓰이는 함수  
다항식 함수, 지수 함수  
로그함수의 성질
  1. 0~1사이의 값을 증폭시켜 준다. 0~1사이의 작은값의 차이를 크게 해줌
  2.  $x$ 을  $+x$ 로 변환시켜준다.
  3.  $\log$ 를 취해도 최대, 최소는 바뀌지 않는다.

### b. SymPy를 사용한 함수 미분

- 데이터 분석에서 미분의 필요성  
데이터 분석의 목표는 최적의 예측모형을 구하는 것. 예측모형은 입력 데이터외에도 모수(parameter)를 입력을 가짐  
모수를 입력으로 하고 모델의 성능을 출력으로 하는 함수를 생각할 수 있다. 예측모델 함수와 다른 이 함수의 성능을 최대화 하는 것이 최적화의 목표
- 미분이란 함수로부터 새로운 함수를 도출하는 변환의 일종  
미분을 통해만들어진 함수를 도함수라고 함(도출된 함수)
- 미분공식
- 선형조합
- 곱셈법칙
- 연쇄법칙
- 2차도함수  
최저점을 GD로 찾아갈때 빠르게 이동가능하도록 한다.
- 편미분  
만약 함수가 두 개 이상의 독립변수를 가지는 다변수 함수인 경우에도 미분 즉, 기울기는 하나의 변수에 대해서만 구할 수 있다. 이를 편미분(partial differentiation)이라 함
- SymPy  
<https://datascienceschool.net/view-notebook/5f5f4a966f9042efa43d711008df9c0b/>

### c. 행렬의 미분

스칼라값 뿐만 아니라 함수도 행렬의 원소가 될 수 있음. 이는 함수 종속변수  $y$ 가 행렬인 경우로 볼 수있음  
반면, 여러개의 독립변수를 갖는 다변수 함수는 함수의 독립변수가 벡터인 경우로 볼 수 있음  
이러한 행렬을 입력이나 출력으로 가지는 함수를 미분하는 것을 행렬 미분(편미분)

- 스칼라를 벡터로 미분하는 경우
- 행렬 미분 법칙