미분 : 머신러닝을 위한 수학

김종혜

시작하면서...

cost function, 경사하강법, backward를 이해하려면 편미분, 체인룰에 대한 선이해가 필요합니다.

오래간만에 미분을 보시는 분들을 위해 간단히 정리를 해 볼까 합니다.

도함수

- 함수 $f(x)$를 미분한다. == 함수 $f(x)$의 도함수 $f'(x)$를 구하는 것입니다.

- 어떤 함수안에 포함된 값 각각이 0에 한없이 가까워지는 극함값(미분계수)를 구하는 함수입니다.

$$f'(x) = \frac{df}{dx}$$

$$= \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$$

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

도함수

- 예를 들어,

$f(x) = 3x$일때의 도함수 $f'(x) = 3$

$f(x) = x^2$ -> $f'(x) = 2x$

- 그렇다면, (x : 변수, c : 상수)

$c' = 0, x' = 1, (x^2)' = 2x, (e^x)' = e^x$

알아야 하는 미분 성질

(1) $c' = 0$ (c : 정수)

(2) $(cx)' = c$ (c : 정수)

(3) { f(x) + g(x) }' = f '(x) +g '(x)

(4){ cf(x) }' = c f '(x)

- (예) $c = (2-y)^2$ => $c' = (4-4y+y^2)' = 4' - (4y)' + (y^2)' = 0 - 4 + 2y = 2y - 4$
- (예) $f(x) = 2x^2 + 3x + 1$ => $f'(x) = 4x + 3$

- (예) f(x) = 1+e^(-x)
  f '(x) = (1+e^(-x))' = 1' + (e^(-x))' = -e(^(-x))

(5) $f(x) = x^n$ => $f'(x) = nx^(n-1)$

(6){ f(x)g(x) }' = f '(x)g(x) + f(x)g '(x)

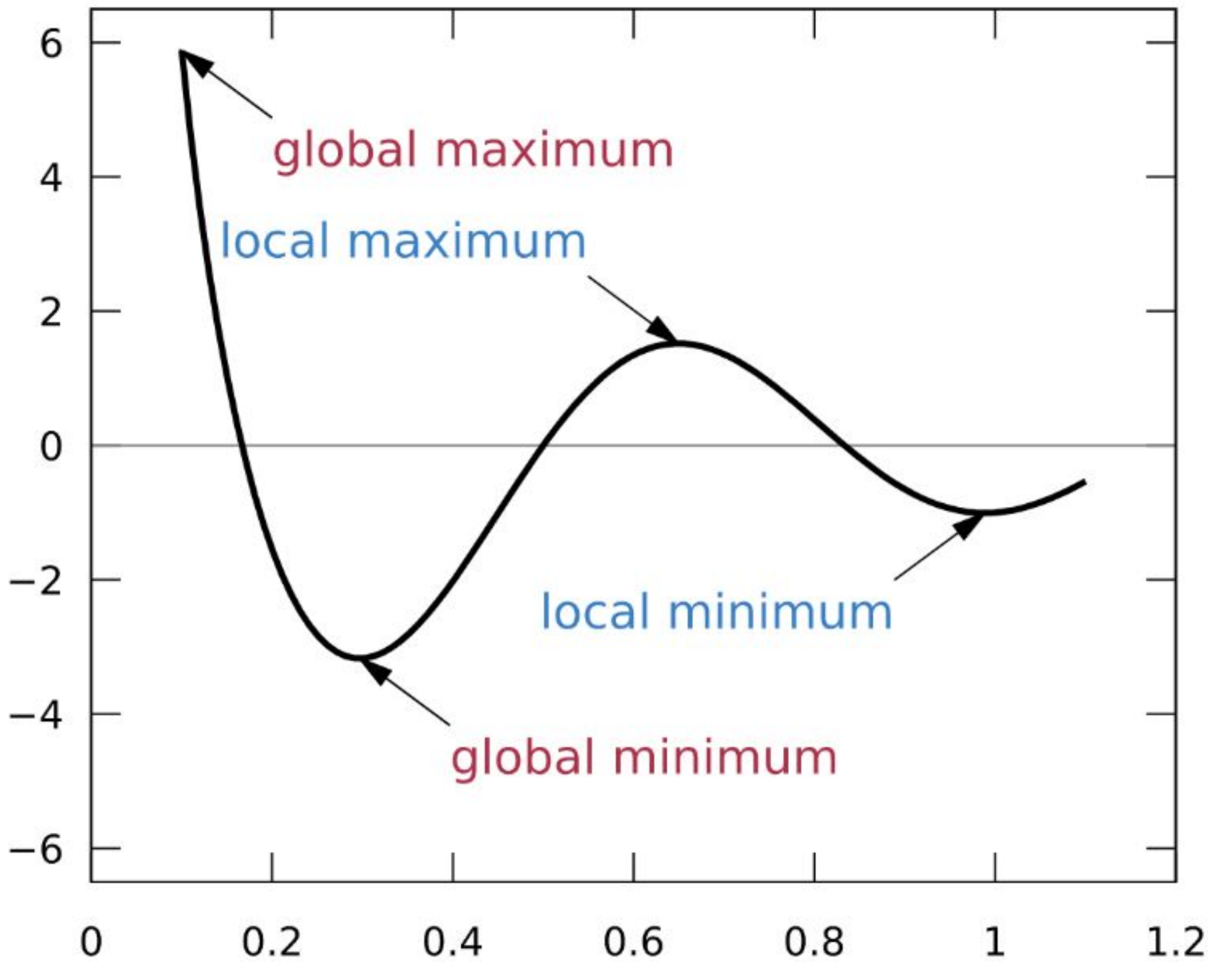
최솟값의 필요성

최솟값(Global minimum), 극소값(local minimum)의 공통점은 미분값이 0입니다.

최솟값과 극소값을 해결해야 합니다.

- 함수$f(x)$가 $x = a$일때 최소값이라면 $f'(a) = 0$
$x = a$이고, $f(x)$가 최소값일때 그 점에서 접선의 기울기(도함수의 값)은 0

$f(a) = 0$은 함수 $f(x)$가 $x = a$에서 최솟값이 되기 위한 필요조건입니다.
BUT!! 접선의 기울기가 0이더라도 꼭 최솟값(global minimum)이라는 보장은 없습니다.
접선의 기울기=0 이지만 극소값(local minimum)일 수 있습니다.

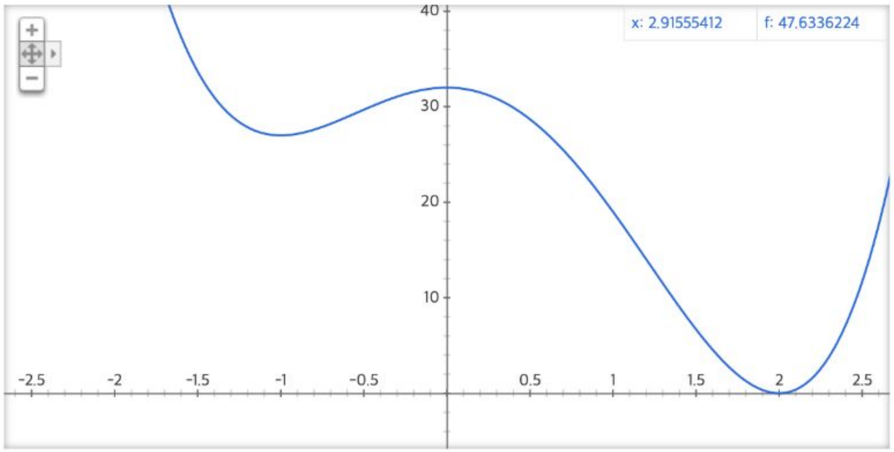


경사하강법은 시작점부터 가장 아래로 내려가는 쪽으로 움직입니다.

그렇기 때문에 시작점을 어디에 하느냐에 따라 local minimum, 즉 주변보다는 낮지만 가장 낮은 값은 아닌 지점에 빠질 위험이 있습니다.

(예) 3차함수 $f(x) = 3x^4 - 4x^3 - 12x^2 + 32$의 최소값을 구한다면 ...

3*x^4-4*x^3-12*x^2+32 그래프



=>$f'(x) = 12x^3 - 12x^2 - 24x = 12x(x^2 + 1)(x - 2)$

x	..	-1	..	0	..	2	..
f '(x)	-	0	+	0	-	0	+
f(x)	하강중	27	상승중	32	하강중	0	상승중
		극소값 (local minimum)		극대값 (local maximum)		최소값 (global minimum)	

3차함수 극소값, 극대값, 최소값

(예) 2차함수 $f(x) = 2x^2 - 4x + 3$ => $f'(x) = 4x - 4$ -> 최소값(global minimum), 즉 f '(x)=0은 "1" 입니다.