통계 분석

2019-2020년 겨울 계절 학기

강봉주

미분

미분

[미분 대한 이해가 필요한 이유]

- 비용 함수의 미분을 통한 가중값의 근사
- 기울기 하강법(gradient descent)에 의한 최저점 찾기
- $C(\beta_1, \ldots, \beta_p)$ 으로 구성된 스칼라 함수에 대한 미분의 필요, 즉, 편미분 벡터에 대한 값 구하기

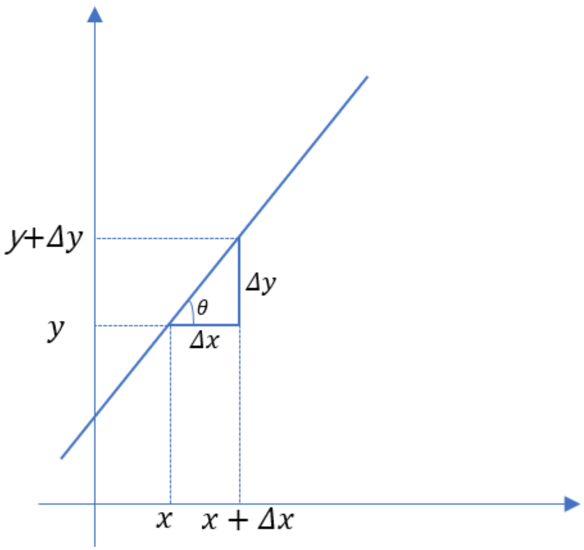
$$\beta_1 := \beta_1 - \lambda \frac{\partial C}{\partial \beta_1}$$

미분

[개요]

- 미분(differentiation)은 도함수(derivative)를 얻는 행위
- 도함수는 입력변수의 특정한 값에서의 출력변수 값의 변화율(rate of change)을 정의한 함수
- 기울기(slope)의 의미: 구간의 변화율
- $y = f(x) = ax + b$

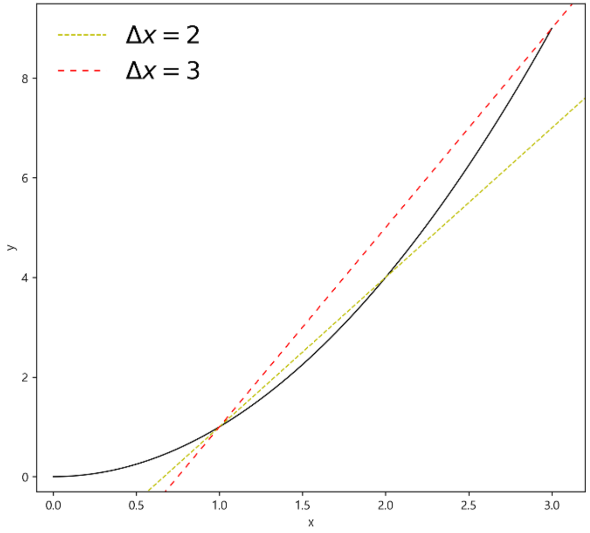
$$a = \frac{y\text{의 변화}}{x\text{의 변화}} = \frac{y + \Delta y \ - y}{x + \Delta x - x} = \frac{\Delta y}{\Delta x}$$



미분

[개요]

- 직선인 경우 Δx가 아무리 작은 값이라 하더라도 기울기는 변함 없음
- 직선이 아닌 경우에는?
- x의 증분(increment)에 따라 기울기가 다름



미분

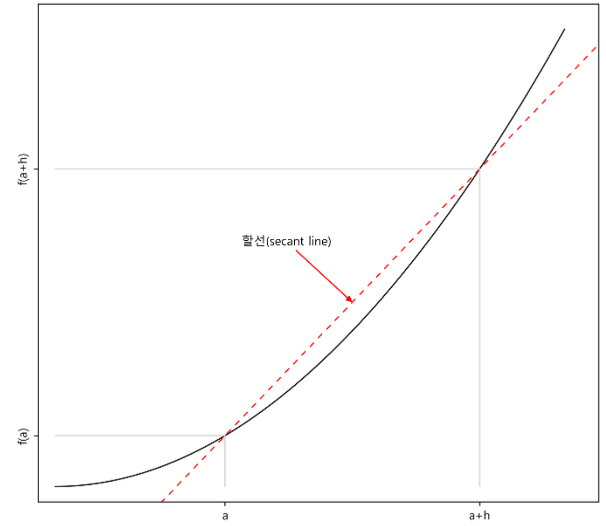
[표기법]

- 라이브니츠(Leibniz) 표기법
- x의 극소의 (infinitesimal) 변화를 dx
- x에 따른 y의 도함수를 $\frac{dy}{dx}$
- 라그랑지 표기법: $f'(x)$

미분

[정의]

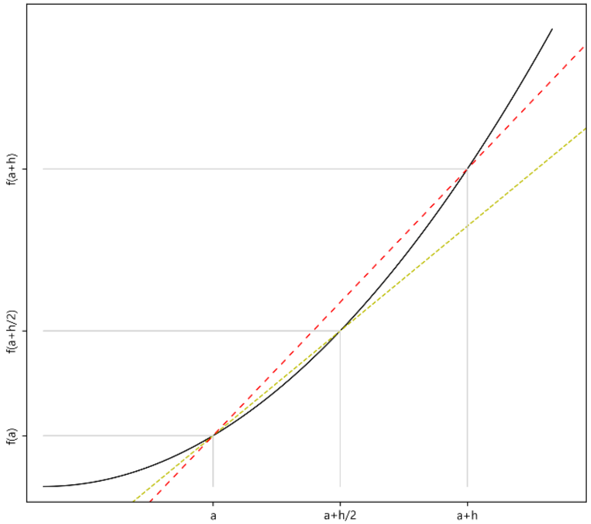
- 한 점 $x = a$ 에서의 미분값
- $\left.\frac{df}{dx}\right|_{x=a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$
- 극한값이 존재하면 함수 f는 $x = a$에서 미분가능(differentiable)
- $\frac{f(a+h)-f(a)}{h}$: 차분 몫(difference quotient)

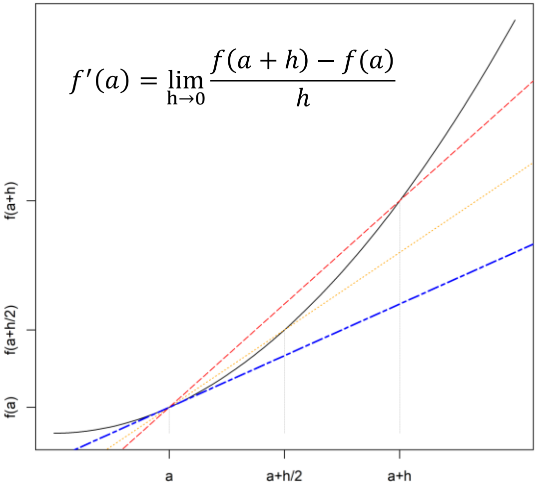


미분

[정의]

- 한 점 $x = a$ 에서의 미분값
- $\frac{f(a+h)-f(a)}{h}$: 차분 몫(difference quotient)
- h의 값에 따라 아주 많은 할선들이 존재, 이러한 할선들의 극한

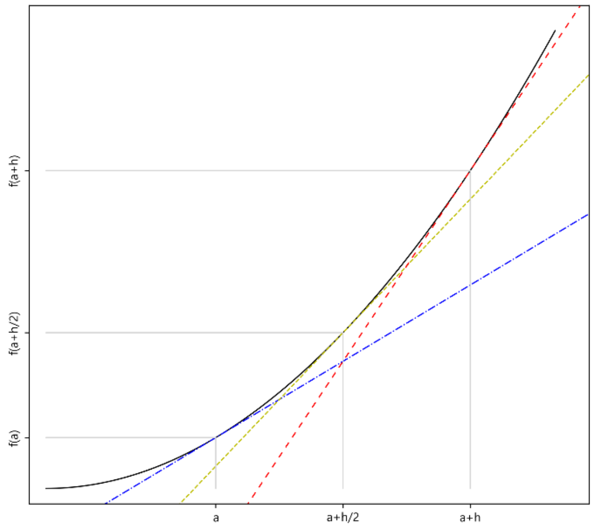




미분

[도함수]

- 모든 x에 대하여 미분 값과 대응되는 함수를 만들어 낼 수 있는데 이를 도함수(derivative function)



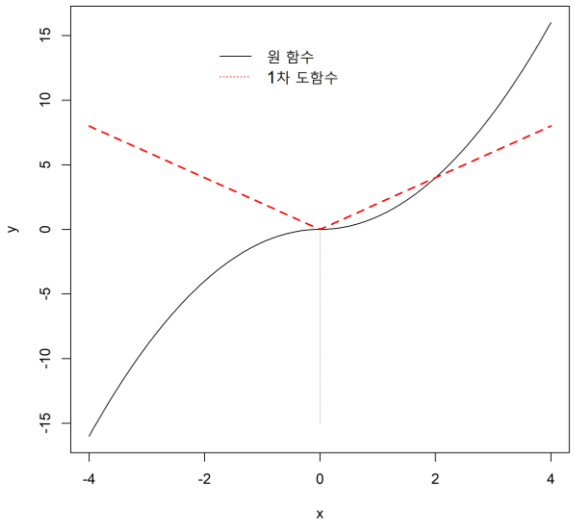
미분

[고차 도함수]

- 고차 도함수: 2차 이상 미분
- $\frac{d^2f}{dx^2}, f'', \frac{d^3f}{dx^3}, \dots$
- 저차가 존재한다고 해서 그 이상의 고차 도함수가 존재하지 않음

$$f(x) = \begin{cases} x^2, if\ x \geq 0 \\ -x^2, if\ x \leq 0 \end{cases}$$

$$f'(x) = \begin{cases} 2x, if\ x \geq 0 \\ -2x, if\ x \leq 0 \end{cases}$$



미분

[고차 도함수]

- 어떤 함수가 무한히 미분가능한 경우, 즉, 모든 차수의 미분이 존재하는 경우에 해당 함수를 무한 미분가능 (infinitely differentiable, smooth)이라고 함

미분

[고차 도함수]

- 어떤 함수가 무한히 미분가능한 경우, 즉, 모든 차수의 미분이 존재하는 경우에 해당 함수를 무한 미분가능 (infinitely differentiable, smooth)이라고 함

$$x^n, e^x, a^x, \ln(x), \log_a(x), \sin(x), \ \ldots.$$

미분

[도함수의 계산]

- $\frac{d}{dx}(7x^3)$
- $\frac{d}{dx}(ax^n) = anx^{n-1}$

In	# 다항 함수의 미분
	n = sympy.Symbol('n')
	a = sympy.Symbol('a')
	x = sympy.Symbol('x', real=True)
	expr = a*x**n
	expr	

Out		ax^n

In	expr.diff(x)	
Out		$\frac{anx^n}{x}$

미분

[도함수의 계산]

- $\frac{d}{dx}(7x^3)$
- $\frac{d}{dx}(ax^n) = anx^{n-1}$

In	`# 다항 함수의 미분 예`
	`x = sympy.Symbol('x', real=True)`
	`expr = 7*x**3`
	`expr`

Out	$7x^3$

In	`expr.diff(x)`
Out	$21x^2$

미분

[도함수의 계산]

- $\frac{d}{dx}(e^x) = e^x$
- $\frac{d}{dx}(a^x) = a^x \ln(a)$

In	`# 오일러 수 또는 자연로그 밑 상수의 정의`
	`n = sympy.Symbol('n', integer=True)`
	`expr = sympy.Sum(1/sympy.factorial(n),(n, 0, sympy.oo))`
	`expr`

Out	$\sum_{n=0}^{\infty} \frac{1}{n!}$

In	`expr.doit()`
Out	e

In	`expr.evalf(20)`
Out	2.7182818284590452354

미분

[도함수의 계산]

- $\frac{d}{dx}(e^x) = e^x$
- $\frac{d}{dx}(a^x) = a^x \ln(a)$

In	`# 지수 함수의 미분`
	`x = sympy.Symbol('x', real=True)`
	`expr = sympy.exp(x)`
	`expr`

Out	e^x

In	`expr.diff(x)`
Out	e^x

미분

[도함수의 계산]

- $\frac{d}{dx}(e^x) = e^x$
- $\frac{d}{dx}(a^x) = a^x \ln(a)$

In	`x = sympy.Symbol('x', real=True)`
	`a = sympy.Symbol('a')`
	`expr = a ** x`
	`expr`

Out	a^x

In	`expr.diff(x)`
Out	$a^x \log(a)$

미분

[도함수의 계산]

- $\frac{d}{dx}\ln(x) = \frac{1}{x}, x > 0$
- $\frac{d}{dx}\log_{\mathrm{a}} x = \frac{1}{x\ln(a)}$

In	`# 로그 함수의 미분`
	`x = sympy.Symbol('x', real=True)`
	`expr = sympy.ln(x)`
	`expr`

Out	$\log(x)$

In	`expr.diff(x)`
Out	$\frac{1}{x}$

미분

[도함수의 계산]

- $\frac{d}{dx}\ln(x) = \frac{1}{x}, x > 0$
- $\frac{d}{dx}\log_{\mathrm{a}} x = \frac{1}{x\ln(a)}$

In	`x = sympy.Symbol('x', real=True)`
	`a = sympy.Symbol('a')`
	`expr = sympy.log(x,a)`
	`expr`

Out	$\frac{\log(x)}{\log(a)}$

In	`expr.diff(x)`
Out	$\frac{1}{x\log(a)}$

미분

[도함수의 계산]

- $\frac{d}{dx}\sin(x) = \cos(x)$
- $\frac{d}{dx}\cos(x) = -\sin(x)$
- $\frac{d}{dx}\tan(x) = \sec^2(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x)$

In	`# 삼각 함수의 미분`
	`x = sympy.Symbol('x', real=True)`
	`expr = sympy.sin(x)`

Out	$\cos(x)$

In	`expr = sympy.cos(x)` `expr.diff(x)`
Out	$-\sin(x)$

In	`expr = sympy.tan(x)` `expr.diff(x)`
Out	$\tan^2(x) + 1$

미분

[도함수의 계산]

- $\frac{d}{dx}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}, -1 < x < 1$
- $\frac{d}{dx}\arccos(x) = -\frac{1}{\sqrt{1-x^2}}, -1 < x < 1$
- $\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}$

In

```
# 삼각함수의 역함수의 미분
expr = sympy.asin(x)
expr.diff(x)
```

Out

$$\frac{1}{\sqrt{1-x^2}}$$

In

```
expr = sympy.acos(x)
expr.diff(x)
```

Out

$$-\frac{1}{\sqrt{1-x^2}}$$

미분

[도함수의 계산]

- 합 규칙
- $\frac{d}{dx}(\alpha f + \beta g) = \alpha \frac{d}{dx} f + \beta \frac{d}{dx} g$

In

```
# 합 규칙
x = sympy.Symbol('x', real=True)
f = sympy.Function('f')(x)
g = sympy.Function('g')(x)
a ,b = sympy.symbols("a, b", real=True)
expr = a*f + b*g
expr.diff(x)
```

Out

$$a \frac{d}{dx} f(x) + b \frac{d}{dx} g(x)$$

[도함수의 계산]

- 곱 규칙
- $\frac{d}{dx}(fg) = \left(\frac{d}{dx}f\right)g + f\left(\frac{d}{dx}g\right)$

In	`# 곱 규칙`
	`expr = f*g`
	`expr.diff(x)`

Out	$f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x)$

미분

[도함수의 계산]

- 몫 규칙
- $\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{\left(\frac{d}{dx}f\right)g - f\left(\frac{d}{dx}g\right)}{g^2}$

In	`# 몫 규칙`
	`expr = f/g`
	`expr.diff(x)`

Out	$-\frac{f(x)\frac{d}{dx}g(x)}{g^2(x)} + \frac{\frac{d}{dx}f(x)}{g(x)}$

[도함수의 계산]

- 연쇄 규칙
- $\frac{d}{dx}\left(f\left(g(x)\right)\right) = \frac{d}{dg(x)} f\left(g(x)\right) \frac{d}{dx} g(x)$

In	`# 연쇄 규칙`
	`x = sympy.Symbol('x', real=True)`
	`g = sympy.Function('g')`
	`f = sympy.Function('f')`
	`expr = f(g(x))`
	`expr.diff(x)`

Out	$\frac{d}{dg(x)} f(g(x)) \frac{d}{dx} g(x)$

미분

[도함수의 계산]

- 과제 2
- 다음을 sympy 패키지를 이용하여 구하세요.

1) $\frac{d}{dx}(x) = ?$

2) $\frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = ?$

3) $\frac{d}{dx}(\tanh(x)) = \frac{d}{dx}\left(\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}\right) = ?$

4) $\frac{d}{dx}(\ln(1+e^{x})) = ?$

5) $\frac{d}{dx}\left(e^{-x^2}\right) = ?$

6) $\frac{d}{dx}(\max(0,x)) = ?$

7) $\frac{d}{dx}(-x\ln x) = ?$

[다변수 함수의 미분]

- 편미분: 다변수로 정의된 함수에 대한 미분
- $f(x,y) = x^2 + xy + y^2$

In	`# 다변수 함수`
	`x, y = sympy.symbols('x, y', real=True)`
	`sympy.Lambda((x, y), x**2 + x*y + y**2)`

Out	$\left((x,\ y) \mapsto x^2 + xy + y^2\right)$

[다변수 함수의 미분]

- 편미분: 다변수로 정의된 함수에 대한 미분
- $f_x(y) = x^2 + xy + y^2$: x가 고정된 경우에는 y 만의 함수

In	`# x가 고정된 경우: y 만의 함수`
	`x = sympy.Float(2)`
	`sympy.Lambda(y, x**2 + x*y + y**2)`

Out	$(y \mapsto y^2 + 2.0y + 4.0)$

미분

[다변수 함수의 미분]

- 편미분: 다변수로 정의된 함수에 대한 미분
- $f_x(y) = x^2 + xy + y^2$: x가 고정된 경우에는 y 만의 함수
- $\frac{\partial f}{\partial y} = x + 2y$

In
```python
# 편미분
x, y = sympy.symbols('x, y', real=True)
f = sympy.Function('f')(x, y)
f.diff(y)
```
Out

$$\frac{\partial}{\partial y} f(x, y)$$

In
```python
f = x**2 + x*y + y**2
f.diff(y)
```
Out

$$x + 2y$$

[다변수 함수의 미분]

- 그래디언트: 모든 변수에 대한 편미분 벡터
- $\nabla f(a_1, \dots, a_n) = \left(\frac{\partial f}{\partial x_1}(a_1, \dots, a_n), \dots, \frac{\partial f}{\partial x_n}(a_1, \dots, a_n)\right)$

In

```
# 그래디언트(기울기, 경사도)
x, y = sympy.symbols('x, y', real=True)
f = sympy.Function('f')(x, y)
G = sympy.Matrix([f.diff(x), f.diff(y)])
G
```

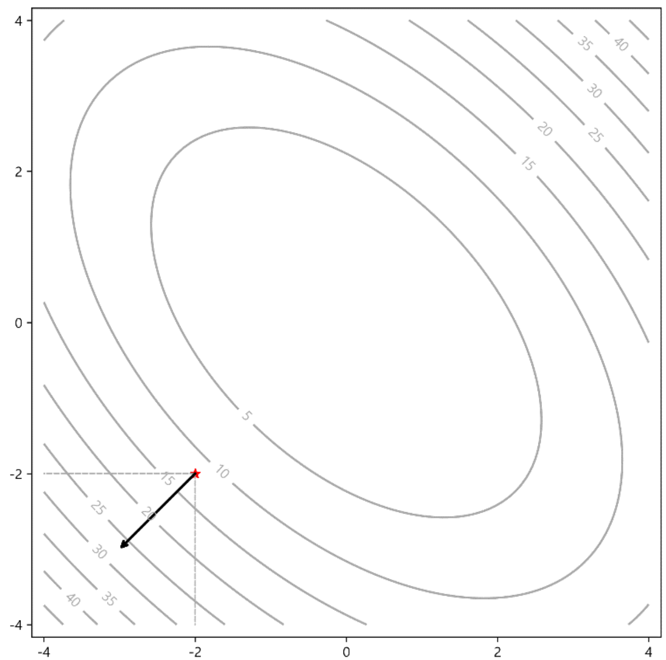
Out

$$\begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix}$$

미분

[다변수 함수의 미분]

- 그래디언트의 기하학적인 의미
- $f(x, y) = x^2 + xy + y^2$
- 그래디언트 벡터는 타원의 접선과 수직이며 최대값 방향을 표시



미분

[행렬 미분]

- 행렬 미분
- 다변수에 대한 미분을 특별한 기호로 단순화 한 것
- $y = (y_1, \dots, y_n)^T, x = (x_1, \dots, x_p)^T$

야코비안(Jacobian) 행렬

$$\frac{\partial y}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_p} \end{pmatrix}_{n \times p}$$

$$\left(\frac{\partial y}{\partial x}\right)_{ij} = \frac{\partial y_i}{\partial x_j}$$

[행렬 미분]

- 행렬 미분
- $\mathbf{y} = \mathbf{A}\mathbf{x} \Rightarrow \dfrac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}$

In	`# 행렬 미분`
	`a, b, c, d = sympy.symbols("a, b, c, d")`
	`A = sympy.Matrix([[a, b], [c, d]])`
	`A`

Out	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$
In	`x = sympy.Matrix(sympy.symbols("x_1, x_2"))` `x`
Out	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
In	`expr = A*x` `expr`
Out	$\begin{bmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{bmatrix}$
In	`expr.diff(x[0])`
Out	$\begin{bmatrix} a \\ c \end{bmatrix}$
In	`expr.diff(x).reshape(2, 2).transpose()`
Out	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

미분

[행렬 미분]

- 행렬 미분
- $\boldsymbol{\alpha} = \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} \Rightarrow \frac{\partial\boldsymbol{\alpha}}{\partial\mathbf{x}} = \mathbf{x}^{\mathbf{T}}\left(\mathbf{A} + \mathbf{A}^{\mathbf{T}}\right)$

In	`# y=x^TAx`
	`expr = x.transpose() * A * x`
	`expr`

Out	$\left[x_1(ax_1 + cx_2) + x_2(bx_1 + dx_2)\right]$
In	`expr.diff(x).reshape(2, 1)`
Out	$\begin{bmatrix} 2ax_1 + bx_2 + cx_2 \\ bx_1 + cx_1 + 2dx_2 \end{bmatrix}$

In	`expr2 = x.transpose()*(A+A.transpose())` `expr2.reshape(2, 1)`
Out	$\begin{bmatrix} 2ax_1 + bx_2 + cx_2 \\ bx_1 + cx_1 + 2dx_2 \end{bmatrix}$

미분

[행렬 미분]

- 과제 3
- $X_{n\times p}, \beta_{p\times 1}, y_{n\times 1}$ 에 대하여 $\frac{\partial}{\partial\beta}(y - X\beta)^T(y - X\beta)$을 구하세요.