Linear Algebra for Statistics

Chapter 1

Instructor: Seokho Lee (이석호)

Hankuk University of Foreign Studies

Chapter 1 미분(Derivative)

1.1 사상(Mapping)

사상(mapping)이란 어떤 그룹 안의 원소와 다른 그룹의 원소를 서로 연결해 주는 관계를 의미

- **일대일 사상 (one-to-one mapping)** : X 원소 하나에 Y 원소 한 개가 대응
- **일대다 사상 (one-to-many mapping)** : X 원소 하나에 Y 원소 여러 개가 대응
- **다대일 사상 (many-to-one mapping)** : X 원소 여러 개에 Y 원소 한 개가 대응
- **다대다 사상 (many-to-many mapping)** : X 원소 여러 개에 Y 원소 여러 개가 대응

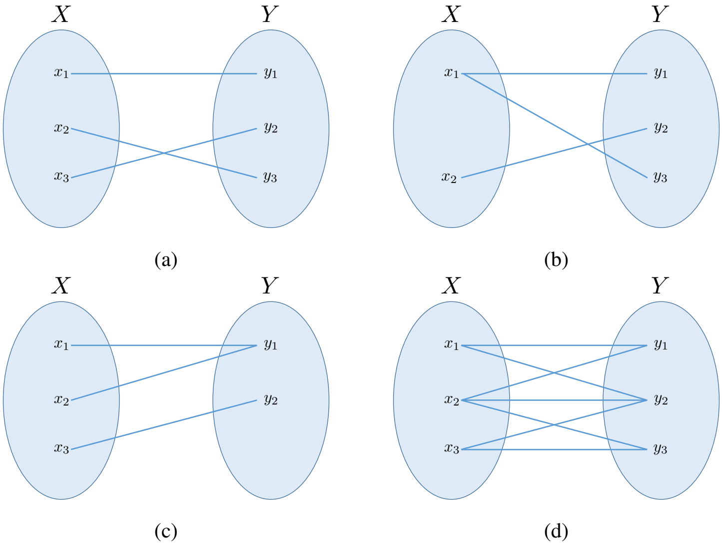


Figure 1.1: (a) 일대일 사상 (b) 일대다 사상 (c) 다대일 사상 (d) 다대다 사상

1.2 함수(Function)

한 그룹(X)에서 다른 그룹(Y)으로 방향성이 주어진 사상 중에서 X의 한 개의 값에 대응되는 Y의 값이 오직 1개 뿐인 사상을 함수(function)라고 한다.

- 사상 M을 $M : X \to Y$로 표현
- X를 정의역(domain), Y를 변역(range)이라 부름
- 함수는 일대일 혹은 다대일 사상의 일종임

함수의 예:

- $f(x) = 1 + 2x$
- $g(y) = y + y^2$
- $h(x) = \dfrac{x}{x+2}$ $(x > 0)$
- $f(t) = \sin(2\pi t)$

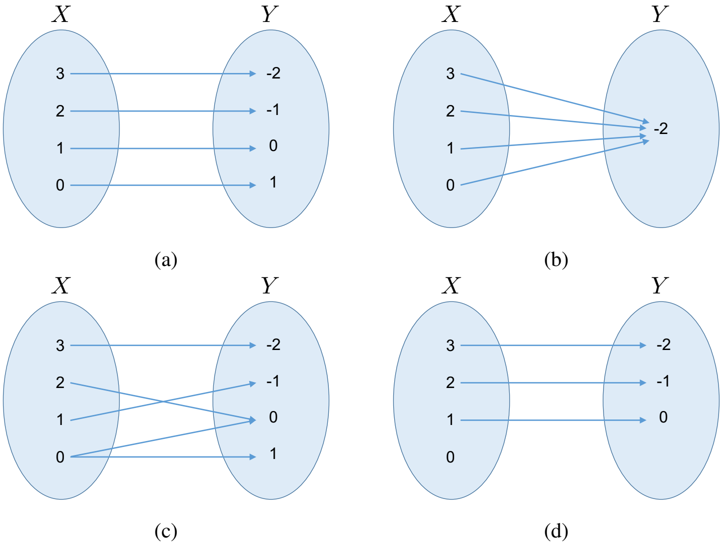


Figure 1.2: (a) 함수 (b) 함수 (c) 비함수 (d) 비함수

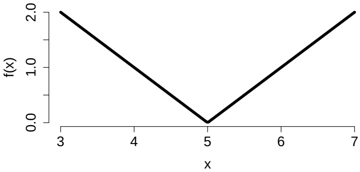
Example

아래에 주어진 $f : X \to Y$가 함수인지 아닌지 판별하라.

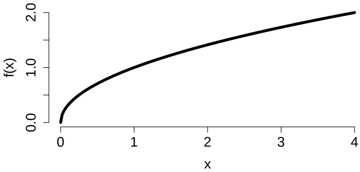
1. $f(x) = |x - 5|$
2. $f(x) = \sqrt{x}, \quad x > 0$
3. $f(x) = \pm\sqrt{x + 2}, \quad x > -2$
4. $f(x) = \ln x$
5. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Solution

1. $f(x) = |x - 5|$

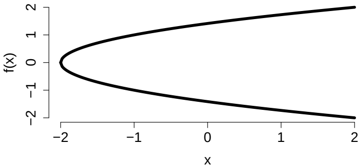


2. $f(x) = \sqrt{x}, \quad x > 0$

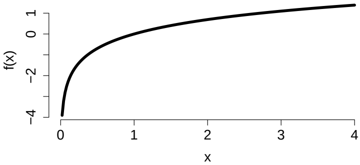


Solution

3. $f(x) = \pm\sqrt{x+2}, \quad x > -2$

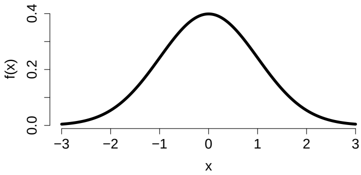


4. $f(x) = \ln x$



Solution

5. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$



1.3 다변량함수(Multivariate Function)

다변량함수(multivariate function)는 정의역이 2개 이상인 함수를 의미한다.

다변량함수의 예:

- $h(x,y) = 1 + 2x + 3y$
- $g(t_1, t_2) = \frac{t_1 + t_2}{t_1 t_2}$
- $f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{(x-\mu_x)^2 - 2\rho(x-\mu_x)(y-\mu_y) + (y-\mu_y)^2}{2(1-\rho^2)}\right)$

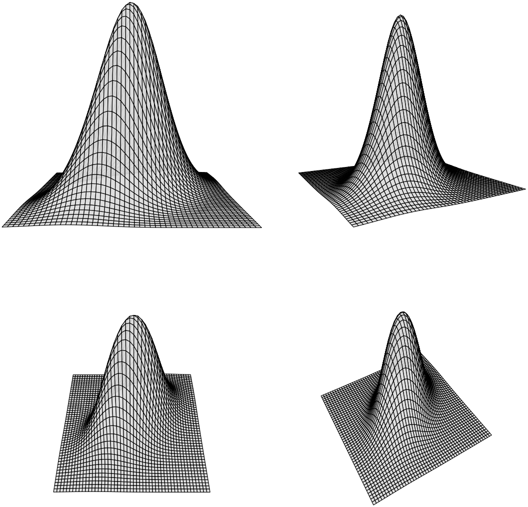


Figure 1.3: 이변량정규분포의 확률밀도함수의 예시

1.4 미분함수(Derivative)

x에서 $h(>0)$ 만큼 변화할 때, 함수 $f(x)$의 변화율은 다음과 같이 정의된다

$$\Delta f(x) = \frac{f(x+h) - f(x)}{h} \tag{1.1}$$

점 x에서 순간변화율은 위의 식에서 h를 0에 매우 가깝에 함으로써 얻을 수 있으며, 이러한 순간변화율을 $f(x)$**의 미분**이라고 한다.

Definition

함수 $f(x)$의 미분함수(derivative function, 도함수)는 $f'(x)$라고 표기하며 아래와 같이 정의된다

$$f'(x) = \frac{d}{dx} f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1.2}$$

미분함수의 예:

- $f(x) = 1 + 2x + 3x^2 \quad \Longrightarrow \quad f'(x) = 2 + 6x$
- $f(x) = \sin x \quad \Longrightarrow \quad f'(x) = \cos x$
- $f(x) = \ln x \quad \Longrightarrow \quad f'(x) = \dfrac{1}{x}$
- $f(x) = \log_{10} x \quad \Longrightarrow \quad f'(x) = \dfrac{1}{x \ln 10}$
- $f(x) = \exp(x) \quad \Longrightarrow \quad f'(x) = \exp(x)$
- $f(x) = x^a \quad \Longrightarrow \quad f'(x) = ax^{a-1}$
- $f(x) = a^x,\ (a > 0) \quad \Longrightarrow \quad f'(x) = a^x \ln a$
- $f(x) = g(x)m(x) \quad \Longrightarrow \quad f'(x) = g(x)'m(x) + g(x)m'(x)$
- $f(x) = \dfrac{g(x)}{m(x)} \quad \Longrightarrow \quad f'(x) = \dfrac{g'(x)m(x) - g(x)m'(x)}{m^2(x)}$

1.5 편미분함수(Partial Derivative)

Definition

함수 $f(x_1, x_2)$의 점 (x_1, x_2)에서의 편미분함수는

$$\frac{\partial f}{\partial x_1} = \frac{\partial}{\partial x_1} f(x_1, x_2) = \lim_{h \to 0} \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h}$$
$$\frac{\partial f}{\partial x_2} = \frac{\partial}{\partial x_2} f(x_1, x_2) = \lim_{h \to 0} \frac{f(x_1, x_2 + h) - f(x_1, x_2)}{h}$$

라고 정의하며, 이를 벡터로 표시한

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} \tag{1.3}$$

를 $f(x_1, x_2)$의 스코어벡터(score vector) 혹은 기울기벡터(gradient vector)라고 한다.

편미분함수의 예:

- $f(x_1, x_2) = x_1^2 + 3x_1x_2 + x_2 - 1 \quad \Longrightarrow \quad \frac{\partial f}{\partial x_1} = 2x_1 + 3x_2, \quad \frac{\partial f}{\partial x_2} = 3x_1 + 1$
- $f(x, y) = \ln(x + y) \quad \Longrightarrow \quad \frac{\partial f}{\partial x} = \frac{1}{x+y}, \quad \frac{\partial f}{\partial y} = \frac{1}{x+y}$

1.6 헤시안행렬(Hessian Matrix)

함수 $f(x)$를 x에 대해 두번 미분한 함수를 2차 미분함수(second derivative)라고 하며 아래와 같이 정의한다

Definition

함수 $f'(x)$가 $f(x)$의 미분함수일 때, $f''(x)$는 $f'(x)$의 미분함수이며

$$f''(x) = \frac{d}{dx} f'(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h} \tag{1.4}$$

라고 정의된다.

다변량함수 $f(x,y)$에 대해 2차 편미분함수(second partial derivative)가 정의되며, 아래와 같이 표기된다.

$$\frac{\partial^2 f}{\partial x^2} := \frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial x}\right) \qquad \frac{\partial^2 f}{\partial x \partial y} := \frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial y}\right)$$

$$\frac{\partial^2 f}{\partial y \partial x} := \frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial x}\right) \qquad \frac{\partial^2 f}{\partial y^2} := \frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial y}\right)$$

이를 행렬로 정리하면

$$H(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \tag{1.5}$$

이며, 이를 함수 $f(x,y)$의 헤시안행렬(hessian matrix)라고 한다.

Example

다음 다변량함수들의 헤시안행렬을 구하라

1. $f(x, y) = x^2 - 4x + y^2 - y - xy$

2. $f(x, y) = \dfrac{x - y}{x^2 + y^2 + 2}$

3. $f(x, y) = \exp(-\frac{x^2+y^2}{2})$

1.7 합성함수(Composite Function)

$f : X \to Y$이고 $g : Y \to Z$일 때, 합성함수 h를 $h : X \to Z$로써 $g(f(x))$로 정의하며 $h(x) = g \circ f(x)$라 표기한다. 함수 h를 합성함수(composite function)이라 한다

Example

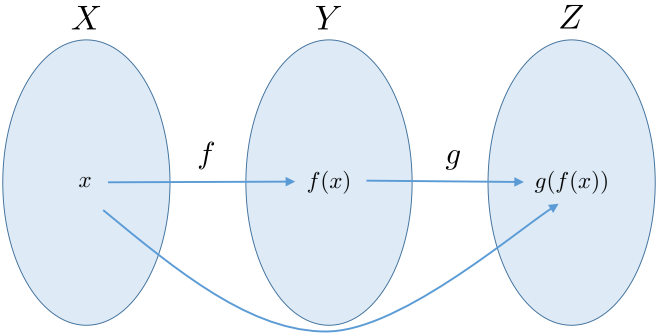
아래 합성함수 $z = h(x) = g \circ f(x)$에 대하여 해당되는 $z = g(y), y = f(x)$를 각각 적어라

1. $h(x) = (x^2 + 3x - 1)^{10}$

   (풀이) $f(x) = x^2 + 3x - 1$, $g(y) = y^{10}$

2. $h(x) = \ln\left(\frac{x-1}{x+1}\right)$

   (풀이) $f(x) = \frac{x-1}{x+1}$, $g(y) = \ln(y)$



$$h(x) = g \circ f(x) = g(f(x))$$

Figure 1.4: 합성합수 $h : X \to Z$의 구조

1.8 연쇄법칙(Chained Rule)

Theorem (정리 1.1 Chained Rule:단변량)

$z = g(y)$이고, $y = f(x)$라면, z를 x로 미분할 경우,

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$$

가 된다.

Example

연쇄법칙에 의해 아래 함수의 미분함수를 구하여라.

1. $h(x) = (x^2 + 3x - 1)^{10}$

   (풀이): $f(x) = x^2 + 3x - 1 = y$, $g(y) = y^{10}$이라 하면 $df/dx = 2x + 3$, $dg/dy = 10y^9$ 이므로,

   $$\frac{dz}{dy}\frac{dy}{dx} = 10y^9(2x+3) = 10(2x+3)(x^2+3x-1)^9$$

2. $h(x) = \ln\left(\frac{x-1}{x+1}\right)$

3. $h(x) = \sin(x^2 + e^x)$

1.9 다변량함수의 연쇄법칙

Theorem (정리 1.2 Chained Rule:다변량)

다변량 함수 $w = f(x, y)$*가* $x = g_1(t)$, $y = g_2(t)$*를 만족할 때,* $w = f(x, y)$*를* t*에 대해 편미분하면*

$$\begin{aligned} \frac{\partial w}{\partial t} &= \frac{\partial f(x,y)}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f(x,y)}{\partial y} \cdot \frac{\partial y}{\partial t} \\ &= \frac{\partial f(x,y)}{\partial x} \cdot \frac{\partial g_1(t)}{\partial t} + \frac{\partial f(x,y)}{\partial y} \cdot \frac{\partial g_2(t)}{\partial t} \qquad (1.6) \end{aligned}$$

Example

아래 합성함수 z에 대하여 미분함수 $\partial z/\partial u$, $\partial z/\partial v$를 각각 구하여라

1. $z = 4e^x \ln y$, 여기서 $x = \ln(u \cos v)$ 그리고 $y = u \sin v$

   (풀이)

$$
\begin{aligned}
\partial z/\partial u &= (\partial z/\partial x)(\partial x/\partial u) + (\partial z/\partial y)(\partial y/\partial u) = (4e^x \ln y)(1/u) + (4e^x/y)(\sin v) \\
\partial z/\partial v &= (\partial z/\partial x)(\partial x/\partial v) + (\partial z/\partial y)(\partial y/\partial v) \\
&= (4e^x \ln y)(\sin v/\cos v) + (4e^x/y)(u \cos v)
\end{aligned}
$$

2. $z = ab + bc + ac$, 여기서 $a = u + v$, $b = u - v$, 그리고 $c = uv$

1.10 최적점과 안장점(Optimal and Saddle Point)

Definition

단변량함수 $f(x)$에 대하여, $x = x_0$에서

$$\frac{df(x_0)}{dx} = 0 \text{ 이고 } \quad \frac{d^2f(x_0)}{dx^2} < 0 \text{ 이면,} \tag{1.7}$$

$f(x_0)$는 극대값이며, 한편

$$\frac{df(x_0)}{dx} = 0 \text{ 이고 } \quad \frac{d^2f(x_0)}{dx^2} > 0 \text{ 이면,} \tag{1.8}$$

$f(x_0)$는 극소값이 된다.

Definition

함수 $f(x,y)$가 (x_0,y_0)에서, 두 조건

$$\frac{\partial f(x_0,y_0)}{\partial x} = \frac{\partial f(x_0,y_0)}{\partial y} = 0 \tag{1.9}$$

$$\left(\frac{\partial^2 f(x_0,y_0)}{\partial x^2}\right)\left(\frac{\partial^2 f(x_0,y_0)}{\partial y^2}\right) - \left(\frac{\partial^2 f(x_0,y_0)}{\partial x \partial y}\right)^2 > 0 \tag{1.10}$$

을 만족시킬 때, 만일

$$\frac{\partial^2 f(x_0,y_0)}{\partial x^2} + \frac{\partial^2 f(x_0,y_0)}{\partial y^2} < 0 \tag{1.11}$$

이면 $f(x_0,y_0)$는 극대값이고, 만일

$$\frac{\partial^2 f(x_0,y_0)}{\partial x^2} + \frac{\partial^2 f(x_0,y_0)}{\partial y^2} > 0 \tag{1.12}$$

이면 $f(x_0,y_0)$는 극소값 이다.

Definition

함수 $f(x, y)$가 (x_0, y_0)에서, 두 조건

$$\frac{\partial f(x_0, y_0)}{\partial x} = \frac{\partial f(x_0, y_0)}{\partial y} = 0 \tag{1.13}$$

$$\left(\frac{\partial^2 f(x_0, y_0)}{\partial x^2}\right)\left(\frac{\partial^2 f(x_0, y_0)}{\partial y^2}\right) - \left(\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y}\right)^2 < 0 \tag{1.14}$$

을 만족시키면 $f(x_0, y_0)$는 안장점(saddle point)이다.

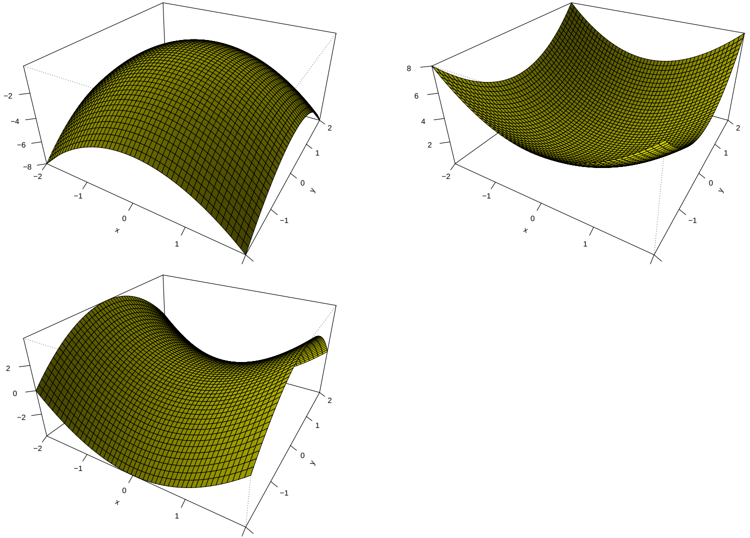


Figure 1.5: 함수의 헤시안행렬이 양정치행렬이어서 함수가 극대값(왼쪽 위)과 극소값(오른쪽 위)을 가질 때의 모습, 그리고 헤시안행렬이 음정치행렬이어서 함수가 안장점(아래)을 갖을 때의 모습

Example

다음을 증명하여라

1. 함수 $f(x,y)=x^2+y^2$는 점 (0,0)에서 극소점을 갖는다.
   (증명) $\partial f(0,0)/\partial x = \partial f(x,y)/\partial x|_{(x,y)=(0,0)} = 2x|_{(x,y)=(0,0)} = 2\cdot 0 = 0$, $\partial f(0,0)/\partial y = 2\cdot 0 = 0$. 그리고 $\partial^2 f(0,0)/\partial x^2 = 2$, $\partial^2 f(0,0)/\partial y^2 = 0$, $\partial^2 f(0,0)/\partial x\partial y = 0$. 그러므로 $(\partial^2 f(0,0)/\partial x^2)(\partial^2 f(0,0)/\partial y^2) - (\partial^2 f(0,0)/\partial x\partial y)^2 = (2)(2) - (0)^2 > 0$. 그리고 $\partial^2 f(0,0)/\partial x^2 + \partial^2 f(0,0)/\partial y^2 = 2+2 > 0$. 따라서 $f(0,0)$은 극소점이다.
2. 함수 $f(x,y)=y^2-x^2$은 점 (0,0)에서 안장점을 갖는다.

1.11 제한된 조건에서 최적점 찾기

Theorem (정리 1.3 라그랑지 승수법)

제약조건 $g(x,y,z) = c$ *하에서의 함수* $f(x,y,z)$ *의 극대, 극소값은 다음 조건을 만족시킨다.* $Q(x,y,z) = f(x,y,z) - \lambda(g(x,y,z) - c)$ *에서*

$$\frac{\partial Q}{\partial x} = \frac{\partial Q}{\partial y} = \frac{\partial Q}{\partial z} = \frac{\partial Q}{\partial \lambda} = 0. \tag{1.15}$$

여기서 Q*를 라그랑지 함수(Lagrange function),* λ*를 라그랑지 승수(Lagrange multiplier)라고 일컫는다.*

Example

제한조건 $x^2+y^2=1$에서, 함수 $f(x,y)=x+2y$의 극대값이 $\sqrt{5}$임을 라그랑지 승수법을 사용하여 증명하라.

(증명) 라그랑지 함수 $Q(x,y)=x+2y-\lambda(x^2+y^2-1)$에 대하여,

$$\frac{\partial Q}{\partial x}=1-2\lambda x=0$$

$$\frac{\partial Q}{\partial y}=2-2\lambda y=0$$

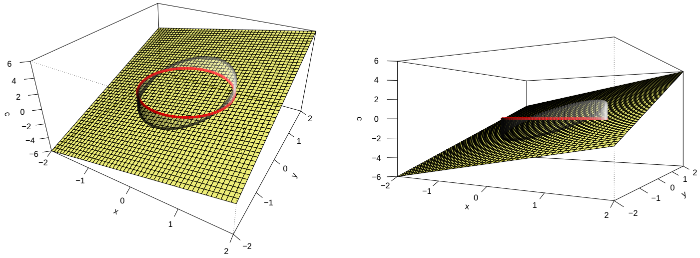
$$\frac{\partial Q}{\partial \lambda}=x^2+y^2-1=0$$

를 만족하는 해를 찾아보면, 처음 두 식으로 부터 $x=1/2\lambda$ 및 $y=1/\lambda$ 임을 알 수 있고, 이를 세번째 식에 대입하면 $\lambda^2=5/4$, 즉 $\lambda=\pm\sqrt{5}/2$이다. 따라서, 극값은 $(x,y)=\pm(1/\sqrt{5},2/\sqrt{5})$에서 주어지며 해당 극값은 $f(1/\sqrt{5},2/\sqrt{5})=\sqrt{5}$ 및 $f(-1/\sqrt{5},-2/\sqrt{5})=-\sqrt{5}$으로 주어진다. 따라서 극대값은 $\sqrt{5}$이며 극소값은 $-\sqrt{5}$이다.

Example

$x^2 + y^2 = 1$와 $x + 2y = c$의 그래프를 그려보고, 두 그래프가 만나는 조건 하에서 c를 최대로 하는 점 (x, y)를 구하라.

(풀이)



1.12 평균값 정리(Mean Value Theorem)

Theorem (정리 1.4 평균값 정리)

함수 $f(x)$가 구간 $[a, b]$에서 정의되고, (a, b)에서 미분가능할 경우,

$$f'(c) = \frac{f(b) - f(a)}{b - a}, \quad a < c < b \tag{1.16}$$

되는 점 c가 적어도 한개 존재한다.

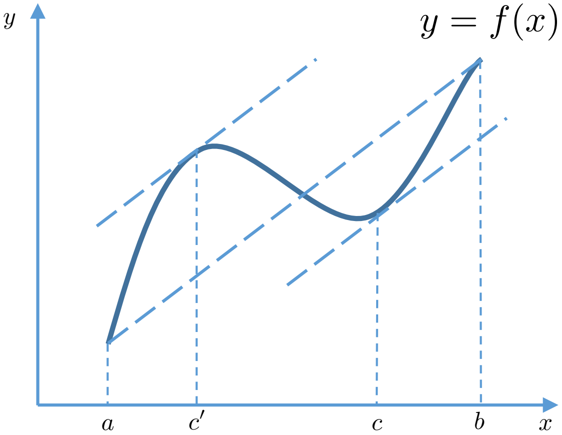


Figure 1.6: 평균값의 정리: 이 그림에서 구간 $[a, b]$ 사이에서 평균기울기와 일치하는 순간기울기는 두 곳, $x = c$ 및 $x = c'$에서 발생하고 있다.

Example

함수 $f(x) = e^x$를 사용하여 구간 $[0, 1]$에 대한 위의 평균값정리를 다시 적어라

(풀이) 지수함수는 전영역에서 미분가능하므로 열린구간 $(0, 1)$에서 미분 가능하며, $f'(x) = e^x$이다. 따라서 $e^c = e - 1$을 만족하는 점 c가 $(0, 1)$ 사이에 적어도 하나 존재한다.

Example

평균값정리를 사용하여 아래 극한값을 구하라.

$$\lim_{x \to 0+} \frac{e^x - 1}{x}$$

(풀이) 평균값 정리에 의하여 양수 x에 대해 $e^c = \frac{e^x - 1}{x}$를 만족하는 c가 구간 $(0, x)$에서 존재한다. 만일 $x \to 0+$이면 $c \to 0+$이므로

$$\lim_{x \to 0+} \frac{e^x - 1}{x} = \lim_{c \to 0+} e^c = 1$$

임을 알 수 있다.

1.13 테일러급수(Taylor's Series)

무한번 미분가능한 모든 함수가 다항식의 무한개 합(infinite sum)으로 표현 가능. 이는 곧 유한개 합(finite sum)에 의해 그 함수가 근사(approximate)될 수 있음을 의미한다.

Theorem (정리 1.5 테일러전개)

무한번 미분가능한 함수 $f(x)$는 점 $x = a$에서 아래와 같은 급수(series)로 전개 가능하다

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots \tag{1.17}$$

이를 테일러전개(Taylor expansion)이라고 한다.

테일러전개를 1차다항식과 나머지로 표현한다면,

$$f(x) = f(a) + f'(a)(x-a) + R_2 \tag{1.18}$$

라고 표현할 수 있으며 이때 나머지항 R_2는

$$R_2 = \frac{f^{(2)}(c)}{2!}(x-c)^2, \quad (c\text{는 } x\text{와 } a \text{ 사이에 위치한 점}) \tag{1.19}$$

으로 표현 가능한다. 만일 나머지 항을 생략해서 근사식으로 표현하면

$$f(x) \approx f(a) + f'(a)(x-a) \tag{1.20}$$

이며, 이를 함수 $f(x)$의 점 $x = a$에서의 1차 테일러 근사식(the 1st order Taylor's approximation)이라 한다.

Example

아래 함수를 $x = 1$에서 2차 다항식으로 근사하고자 한다. 테일러전개를 사용하여 근사시켜라

1. $f(x) = \ln x$

   (풀이) $f'(x) = 1/x$ 이고 $f''(x) = -1/x^2$ 이므로

$$\begin{aligned} f(x) &\approx f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 \\ &= 0 + (1)(x-1) + (-1/2)(x-1)^2 \; = \; -x^2/2 + 2x - 3/2 \end{aligned}$$

2. $f(x) = e^x/x$

Example

계산기에 $\ln x$를 계산할 수 있는 기능이 없다고 가정했을 때, 1차 테일러근사를 이용하여 $\ln 2$의 근사값을 구하라

(풀이) 이전 예제에서 $\ln x$의 $x = 1$에서 1차 테일러근사는 $\ln x \approx \ln 1 + (1)(x-1) = x - 1$이다. 따라서 $\ln 2 \approx 2 - 1 = 1$.