Linear Algebra for Statistics

Chapter 3

Instructor: Seokho Lee (이석호)

Hankuk University of Foreign Studies

Chapter 3 스칼라, 벡터, 그리고 행렬

3.1 유클리드 공간(Euclidean Space)

모든 실수의 집합을 $\mathbb{R}$ 이라 할 때, 이를 기하학적으로 유클리드 1차공간(Euclidea 1-space) 라고 부름

x-y 좌표공간은 유클리드 2차공간(Euclidean 2-space)라 부르며 이를 $\mathbb{R}^2$ 로 표기

이를 확장하여 축이 n 개인 공간을 유클리드 n차공간이라 하며 $\mathbb{R}^n$ 으로 표기

어떤 원소 $\mathbf{x}$ 가 n차원 유클리드 공간에서 정의된다면 이를 $\mathbf{x} \in \mathbb{R}^n$ 로 표현

3.2 스칼라와 벡터(Scalar and Vector)

Definition

스칼라와 벡터는 아래와 같이 정의한다.

- 스칼라(scalar) : 크기를 나타내는 값으로 영문소문자료 표시 (예: a, b, x, y)
- 벡터(vector) : 크기와 방향을 나타내는 값으로 아래와 같이 다양한 표기법을 사용

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \underset{\sim}{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \ldots$$

Remark: 여기서는 **a** 표기법을 사용하고 특별한 언급이 없는 한 벡터는 열벡터(column vector)를 의미한다.

벡터는 원소의 개수에 따라 크기가 결정됨. $\mathbf{x} \in \mathbb{R}^2$은 3×1벡터(three-by-one vector)이며 $\mathbf{x} \in \mathbb{R}^2$는 2×1벡터임

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 10 \end{pmatrix}, \qquad \mathbf{y}^T = (1, 2)$$

열벡터를 기본으로 사용할 때, 행벡터는 $\mathbf{y}^T$ 와 같이 T(전치, transpose)를 위 첨자로 붙여줌.

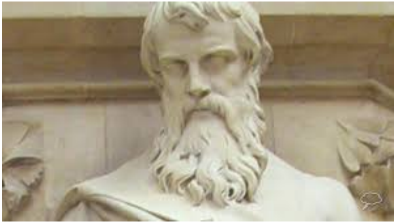

Figure 3.1: 유클리드

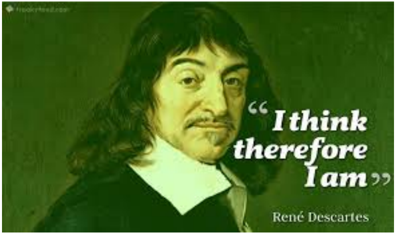

Figure 3.2: 데카르트

3.3 데카르트 좌표계(Cartesian Coordinate)

공간을 직교하는 축으로 표현하고 각 축에 대한 좌표를 이용하여 공간 상의 벡터를 표현하는 표기법을 데카르트 좌표계(Cartesian coordinate system) 혹은 유클리드 좌표계(Euclidean coordinate system)이라 한다.

예를 들어 x-y 좌표평면 상의 원점에서 시작하여 점 (x_1, y_1) 에 이르는 벡터를 **a** 라고 하면, 이 벡터를 좌표계에서는

$$\mathbf{a} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad \text{혹은} \quad \mathbf{a} = (x_1, y_1)^T \tag{3.1}$$

으로 표기한다.

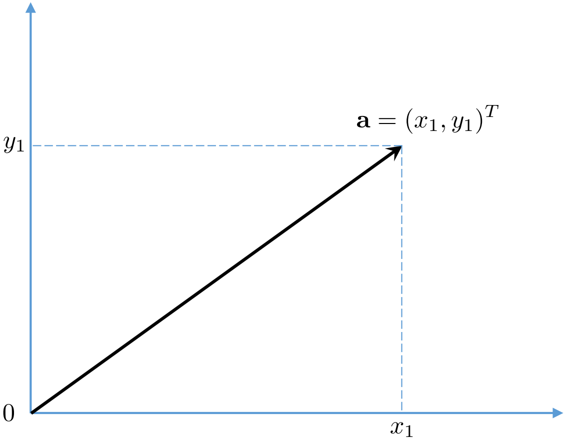

Figure 3.3: 데카르트 좌표계에 표시된 벡터 $\mathbf{a} = (x_1, y_1)^T$

3.4 스칼라와 벡터의 크기

스칼라의 크기는 절대값으로 표현

$$|3| = 3, \quad |-10| = 10$$

벡터의 크기는 데카르트 좌표계에서 정의되는 **원점으로부터의 거리**로서 $\|\mathbf{a}\|$ 라고 표기. 이는 아래와 같이 계산되며 벡터 $\mathbf{a}$의 **노옴(norm)**이라 부른다.

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \longrightarrow \quad \|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2} \tag{3.2}$$

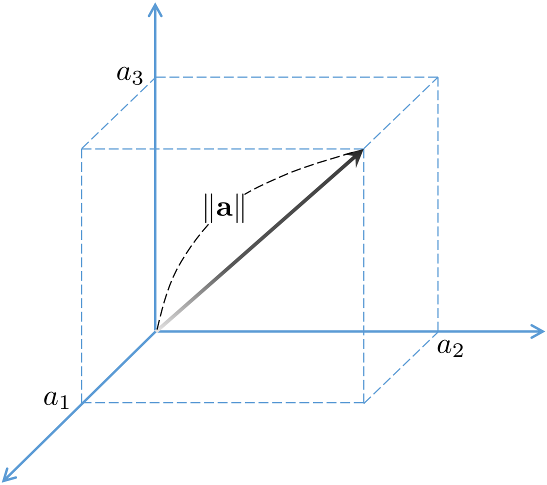

Figure 3.4: 데카르트 좌표계에 표시된 벡터 $\mathbf{a} = (a_1, a_2, a_3)^T$의 크기 $\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$

Example

아래 스칼라, 혹은 벡터를 유클리드 공간에 표시하고, 그 크기를 구하라.

1. $\mathbf{x} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

2. $y = -1$

3. $\mathbf{x}^T = (1, -1, 0)$

3.5 벡터의 연산작용

Definition

스칼라 a 와 벡터 $\mathbf{x} \in \mathbb{R}^n$ 간의 곱은 다음과 같이 정의한다.

$$a \times \mathbf{x} = a\mathbf{x} = a \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_i \\ \vdots \\ ax_n \end{pmatrix}.$$

$a\mathbf{x}$ 는 벡터의 길이를 a배 하며, a 가 음수인 경우 $\mathbf{x}$ 의 방향이 바뀌게 된다.

표기법을 간편하게 하기 위해 벡터 $\mathbf{x}$ 를 $\mathbf{x} = (x_i)_{i=1,\ldots,n}$ 으로 표기한다. 이 경우

$$a\mathbf{x} = a(x_i)_{i=1,\ldots,n} = (ax_i)_{i=1,\ldots,n}$$

으로 표기할 수 있다.

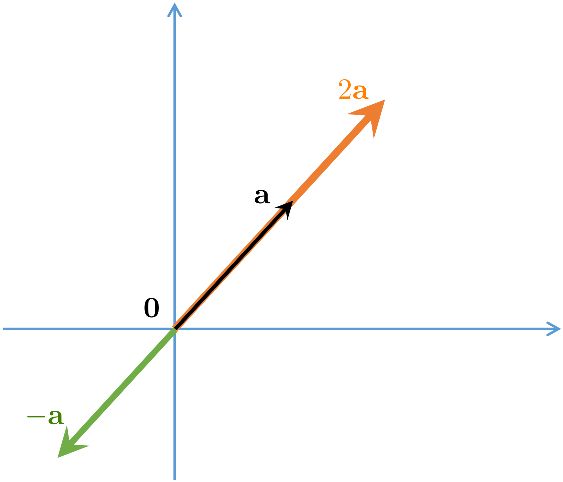

Figure 3.5: 2차원 유클리드 공간에서 정의된 벡터 **a** 의 스칼라 곱 2**a**, −**a** 의 모습

Definition

같은 공간상에서 정의되는 벡터 $\mathbf{x} \in \mathbb{R}^n$ 와 벡터 $\mathbf{y} \in \mathbb{R}^n$ 간의 덧셈 및 뺄셈의 정의는 다음과 같다.

$$\mathbf{x} \pm \mathbf{y} = (x_i \pm y_i)_{i=1,\ldots,n}$$

벡터의 덧셈은 두 벡터가 이루는 평행사변형의 사잇각을 통과하는 대각꼭지점까지의 벡터로 표현 가능

벡터의 뺄셈 $\mathbf{x} - \mathbf{y}$ 는 $\mathbf{x} + (-\mathbf{y})$ 로 표현할 수 있음

Remark: 영벡터는 벡터의 모든 원소가 0으로 이루어진 벡터이다.

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (0)_{i=1,\ldots,n} \in \mathbb{R}^n$$

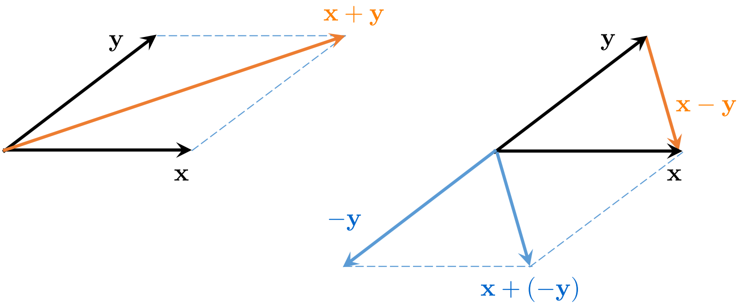

Figure 3.6: $\mathbb{R}^2$에서 정의된 두 벡터 $\mathbf{x}, \mathbf{y}$ 의 덧셈 $\mathbf{x}+\mathbf{y}$ 과 뺄셈 $\mathbf{x}-\mathbf{y}$

Proposition (3.1)

벡터 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ 에 대하여 다음이 성립한다.

1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
3. $\mathbf{x} + \mathbf{0}_n = \mathbf{x}$
4. $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}_n$
5. $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$, c는 스칼라
6. $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$, c, d는 스칼라
7. $1\mathbf{x} = \mathbf{x}$

Example

벡터 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ 가 아래와 같을 때, 다음을 계산하라

$$\mathbf{x} = (3, 4, 0)^T, \quad \mathbf{y} = (2, -1, 2)^T. \tag{3.3}$$

1. $\mathbf{x} - \mathbf{y}$
2. $\sqrt{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2}$
3. $\|\mathbf{x} + 2\mathbf{y}\|$

3.6 벡터의 내적(Inner Product)

Definition

벡터의 전치(transpose of vector)는 벡터를 구성하는 원소의 행을 열로, 열을 행으로 바꾸는 과정을 말하며 윗첨자 T를 사용해 표기한다.

Example

1. $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \longrightarrow \quad \mathbf{x}^T = (2, 1, 3)$
2. $\mathbf{y}^T = (1, 2) \quad \longrightarrow \quad (\mathbf{y}^T)^T = \mathbf{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Definition

같은 공간에 정의된 두 벡터 $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^n$ 의 내적은 다음과 같이 정의된다.

$$\mathbf{x}^T\mathbf{y} = (x_1, x_2, \ldots, x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{3.4}$$

$$= x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

$$= \sum_{i=1}^{n} x_iy_i$$

Example

$$\mathbf{x} = (1,1)^T, \quad \mathbf{y} = (1,-1)^T, \quad \mathbf{z} = (0,1)^T$$

1. $\mathbf{x}^T\mathbf{y} = (1)(1) + (1)(-1) = 0$
2. $\mathbf{x}^T\mathbf{z} = (1)(0) + (1)(1) = 1$
3. $\mathbf{y}^T\mathbf{y} = (1)(1) + (-1)(-1) = 2$

Theorem (정리 3.2)

두 벡터 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 간의 사잇각을 θ 라고 하면, 내적 $\mathbf{x}^T\mathbf{y}$ 는 아래의 성질을 만족한다.

$$\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta. \tag{3.5}$$

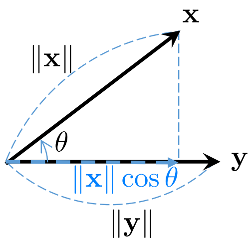

Figure 3.7: 벡터 $\mathbf{x}$ 와 벡터 $\mathbf{y}$ 의 내적

Proof.

증명을 간단히 하기 위해, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ 라 가정하자. $\mathbf{x}$ 및 $\mathbf{y}$ 가 수평축과 이루는 각을 각각 θ_1, θ_2 라 하면 사잇각은 $\theta = \theta_1 - \theta_2$ 이다. 증명을 위해

$$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 = \|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta_1 - \theta_2) \tag{3.6}$$

를 보이면 충분하다. 삼각함수의 성질에 의해

$$\cos(\theta_1 - \theta_2) = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 \tag{3.7}$$

이고

$$\cos\theta_1 = \frac{x_1}{\|\mathbf{x}\|}, \quad \cos\theta_2 = \frac{y_1}{\|\mathbf{y}\|}, \quad \sin\theta_1 = \frac{x_2}{\|\mathbf{x}\|}, \quad \sin\theta_2 = \frac{y_2}{\|\mathbf{y}\|} \tag{3.8}$$

를 이용하면

$$\|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta_1 - \theta_2) = \|\mathbf{x}\|\|\mathbf{y}\|\left(\frac{x_1}{\|\mathbf{x}\|}\frac{y_1}{\|\mathbf{y}\|} + \frac{x_2}{\|\mathbf{x}\|}\frac{y_2}{\|\mathbf{y}\|}\right) = x_1y_1 + x_2y_2 = \mathbf{x}^T\mathbf{y} \tag{3.9}$$

가 성립함을 알 수 있다. □

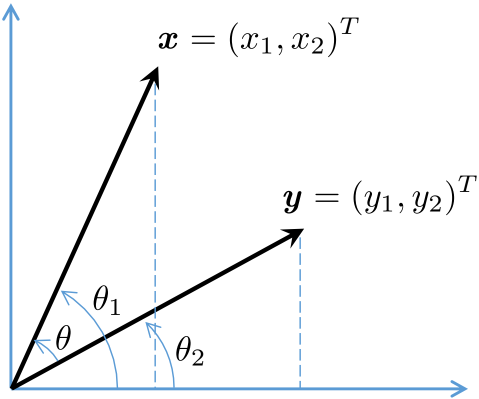

Figure 3.8: 벡터 $\mathbf{x}$ 와 벡터 $\mathbf{y}$ 의 내적 $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$ 의 증명과정

Proposition (3.3)

벡터 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ 에 대하여 다음이 성립한다.

1. $\mathbf{x}^T\mathbf{y}$ *는 스칼라이다.*
2. $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$
3. $\mathbf{x}^T(\mathbf{y} + \mathbf{z}) = \mathbf{x}^T\mathbf{y} + \mathbf{x}^T\mathbf{z}$
4. $\mathbf{0}_n^T\mathbf{x} = 0$
5. $\mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$

Proof.

내적의 정의 $\mathbf{x}^T\mathbf{y} = \sum_{i=1}^n x_i y_i$ 를 이용하면 손쉽게 증명된다. □

3.7 직교벡터(Orthogonal Vectors)

두 벡터 간의 사잇각이 직각(90°)을 이룬다면 **두 벡터는 서로 직교한다(orthogonal)**고 말하고 이를 직교벡터라 부른다.

직교벡터의 내적은 0이며 역도 성립한다.

$$\mathbf{a}^T\mathbf{b} = \|\mathbf{a}\|\|\mathbf{b}\| \cos\left(\frac{\pi}{2}\right) = \|\mathbf{a}\|\|\mathbf{b}\| \times 0 = 0.$$

Definition

$\mathbf{0}$ 이 아닌 두 벡터 $\mathbf{a}, \mathbf{b}$ 간의 사잇각이 직각이면 두 벡터는 직교벡터(orthognal vectors)라고 한다.

Definition

$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 에 대하여, $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ 이고 $\mathbf{x}^T\mathbf{y} = 0$ 이면, $\mathbf{x}$ 와 $\mathbf{y}$ 는 직교정규벡터 (northonomal vectors)라 한다.

Example

다음 세 벡터는 직교정규벡터이다.

$$\mathbf{e}_1 = (1, 0, 0)^T, \quad \mathbf{e}_2 = (0, 1, 0)^T, \quad \mathbf{e}_3 = (0, 0, 1)^T$$

Remark 임의의 벡터 $\mathbf{x} \in \mathbb{R}^n$ 에 대하여

$$\mathbf{x}^* = \frac{1}{||\mathbf{x}||}\mathbf{x} \tag{3.10}$$

은 항상 길이가 1인 단위벡터(unit vector)가 되며, $\mathbf{x}^*$ 를 $\mathbf{x}$ **의 정규화벡터(normalized vector)** 라 하며, 이 과정을 **벡터의 정규화(vector normalization)**라고 한다.

Example

아래 벡터가 직교정규벡터가 되기 위한 c 값을 구하라.

$$\mathbf{x}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} -c \\ \frac{1}{2} \\ c \end{pmatrix}.$$

3.8 벡터와 미분

$\mathbf{x}(t) \in \mathbb{R}^2$ 가 시점 t 에서의 위치를 나타내며 $x_1(t)$ 는 경도, $x_2(t)$ 는 위도를 나타낸다고 하자.

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \tag{3.11}$$

시점 $t+h$ 에서 위치 $\mathbf{x}(t+h)$ 는

$$\mathbf{x}(t+h) = \begin{pmatrix} x_1(t+h) \\ x_2(t+h) \end{pmatrix} \tag{3.12}$$

이다. 따라서 시점 t 에서 점의 속도는 다음과 같이 미분으로 주어진다.

$$\mathbf{v}(t) = \frac{d}{dt}\mathbf{x}(t) = \begin{pmatrix} \lim_{h\to 0} \frac{x_1(x+h)-x_1(x)}{h} \\ \lim_{h\to 0} \frac{x_2(x+h)-x_2(x)}{h} \end{pmatrix} = \begin{pmatrix} \frac{d}{dt}x_1(t) \\ \frac{d}{dt}x_2(t) \end{pmatrix} \tag{3.13}$$

속도의 크기, 혹은 속력은 $\|\mathbf{v}(t)\|$ 로 구해진다.

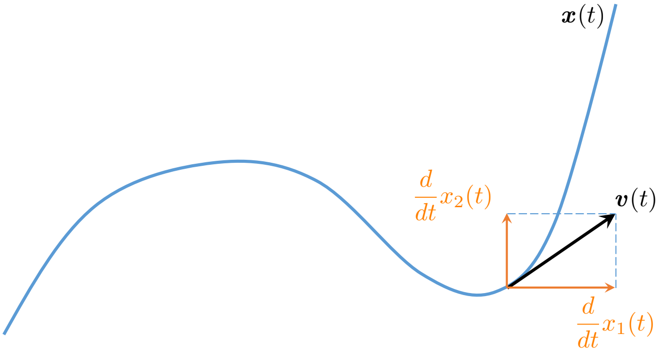

Figure 3.9: 위치벡터 $\mathbf{x}(t)$ 에서 각 수평-수직방향 속도

Definition

벡터 $\mathbf{x} \in \mathbb{R}^2$ 가 t 의 함수일 때, 벡터 $\mathbf{x}$ 를 스칼라 t 로 미분한 벡터는

$$\frac{d}{dt}\mathbf{x}(t) = \begin{pmatrix} \frac{d}{dt}x_1(t) \\ \frac{d}{dt}x_2(t) \\ \vdots \\ \frac{d}{dt}x_n(t) \end{pmatrix} \tag{3.14}$$

이다.

Example

움직이는 자동차의 위치 $\mathbf{x}$ 가 시간 t 가 변함에 따라 GPS 상의 (경도, 위도)로 아래와 같이 변한다고 한다.

$$\mathbf{x} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}. \tag{3.15}$$

1. 이 자동차는 어떤 움직임을 하고 있는지 적어라.
2. 시간이 $\pi/2$ 일 때, 이 자동차는 어느 방향으로 진행하고 있는지 동서남북으로 표현하라.

이전 예와 같이 벡터를 스칼라로 미분하는 경우 외에도 스칼라함수를 벡터로 미분할 수 있다.

두개의 변수 x 및 y 에 의존하는 함수 $f(x, y)$ 를 고려하면 이는 3차 곡면으로 표현 가능하다. 이를 두 변수로 편미분으로 표현하는 벡터를 기울기벡터(gradient vector)라 한다.

$$\mathbf{g} = \nabla f = \begin{pmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{pmatrix} \tag{3.16}$$

기울기벡터는 각 변수 방향의 증가율을 나타낸다.

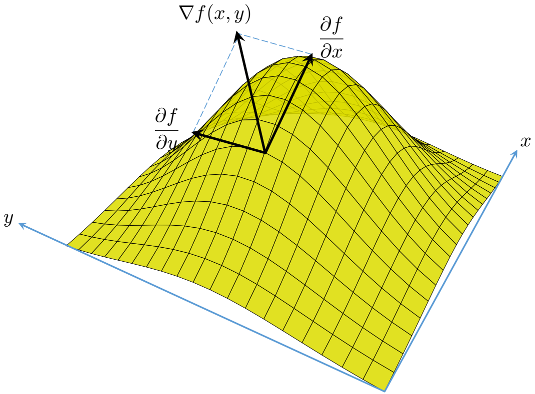

Figure 3.10: 이변량함수 $f(x,y)$ 의 기울기벡터 $\nabla f(x,y)$

Definition

스칼라 함수 $f(x_1, x_2, \ldots, x_n)$ 가 벡터 $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ 의 함수 일 때, 함수 $f(x_1, x_2, \ldots, x_n)$ 의 기울기벡터는 $f(x_1, x_2, \ldots, x_n)$ 을 벡터 $\mathbf{x}$ 로 미분함으로써 얻어지며

$$\nabla f(x_1, x_2, \ldots, x_n) = \frac{d}{d\mathbf{x}} f(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(\mathbf{x}) \\ \frac{\partial}{\partial x_2} f(\mathbf{x}) \\ \vdots \\ \frac{\partial}{\partial x_n} f(\mathbf{x}) \end{pmatrix} \tag{3.17}$$

라고 정의된다. 이 때 각 원소는 각 축 방향의 기울기이다.

Example

점 $(0,1,1)$ 에서 함수 $f(x,y,z) = 3xy - z^2$ 의 각 축 방향의 기울기를 구하라

(풀이)

$$\frac{\partial}{\partial x} f(x,y,z) = 3y$$
$$\frac{\partial}{\partial y} f(x,y,z) = 3x$$
$$\frac{\partial}{\partial z} f(x,y,z) = -2z$$

이고 따라서 기울기벡터는 $\nabla f = (3y, 3x, -2z)^T$ 이다.

3.9 코쉬-슈바르츠(Cauchy-Schwartz) 부등식

Theorem (정리 3.9 코쉬-슈바르츠)

실수 x_i 와 y_i $(i = 1, 2, \ldots, n)$ 에 대하여, 아래 부등식이 항상 성립하며 이를 ***코쉬-슈바르츠 부등식(Cauchy-Schwartz Inequality)****이라고 부른다.*

$$(x_1y_1 + x_2y_2 + \ldots + x_ny_n)^2 \leq (x_1^2 + x_2^2 + \ldots + x_n^2)(y_1^2 + y_2^2 + \ldots + y_n^2)$$

혹은

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \tag{3.18}$$

Proof.

$\mathbf{x} = (x_1, \ldots, x_n)^T$ 및 $\mathbf{y} = (y_1, \ldots, y_n)^T$ 라 하면,

$$\begin{aligned}
\sum_{i=1}^{n} x_i y_i &= \mathbf{x}^T \mathbf{y} \\
\sum_{i=1}^{n} x_i^2 &= \left(\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}\right)^2 = \|\mathbf{x}\|^2 \\
\sum_{i=1}^{n} y_i^2 &= \left(\sqrt{y_1^2 + y_2^2 + \ldots + y_n^2}\right)^2 = \|\mathbf{y}\|^2
\end{aligned}$$

이다. 그런데,

$$-1 \leq \cos\theta = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$$

이므로

$$0 \leq (\cos\theta)^2 = \frac{(\mathbf{x}^T \mathbf{y})^2}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2} \leq 1$$

이다. 따라서

$$(\mathbf{x}^T \mathbf{y})^2 = \left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) = \|\mathbf{x}\|^2\|\mathbf{y}\|^2.$$