기계학습에 필요한 선형 대수

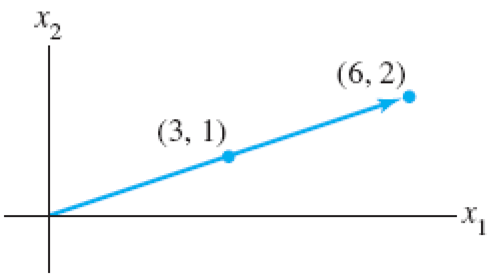
Contents

a $\underline{a}$, $\vec{a}$, $\boldsymbol{a}$, $\mathbf{a}$ $\boldsymbol{A}$, A , $\mathbf{A}$

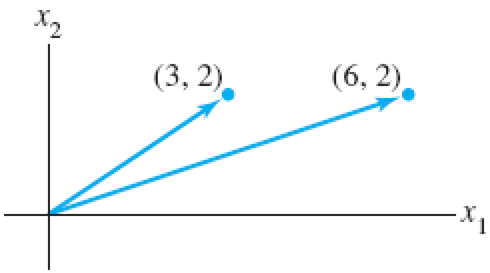
벡 터
(vector)

벡 터

- 벡터는 크기와 방향을 가진다.

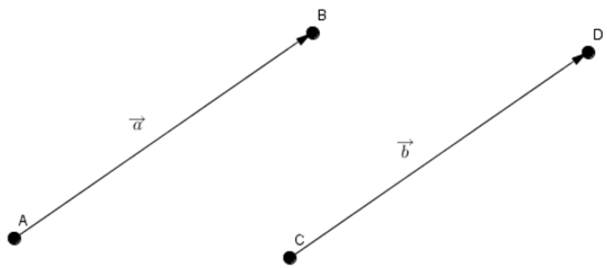


Linearly dependent



Linearly independent

서로 같은 벡터라 함은 '크기'와 '방향'이 둘다 같은 벡터임을 의미합니다.



벡 터

IRIS Data

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa

Iris Versicolor **Iris Setosa** **Iris Virginica**

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

벡 터

벡터

- 샘플을 특징 벡터로feature vector 표현
- 예) Iris 데이터에서 꽃받침의 길이, 꽃받침의 너비, 꽃잎의 길이, 꽃잎의 너비라는 4개의 특징이 각각 5.1, 3.5, 1.4, 0.2인 샘플. $\mathbf{x} \in \boldsymbol{R}^4$.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}$$

- 여러 개의 특징 벡터를 첨자로 구분

$$\mathbf{x}_1 = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 4.9 \\ 3.0 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 4.7 \\ 3.2 \\ 1.3 \\ 0.2 \end{pmatrix}, \cdots, \mathbf{x}_{150} = \begin{pmatrix} 5.9 \\ 3.0 \\ 5.1 \\ 1.8 \end{pmatrix}$$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행 렬

행렬

- 여러 개의 벡터를 담음
- 훈련집합을 담은 행렬을 설계행렬(design matrix)이라 부름
- 예) Iris 데이터에 있는 150개의 샘플을 설계 행렬 $\mathbf{X}$로 표현

$$\mathbf{X} = \begin{pmatrix} 5.1 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 3.2 & 1.3 & 0.2 \\ 4.6 & 3.1 & 1.5 & 0.2 \\ \vdots & \vdots & \vdots & \vdots \\ 6.2 & 3.4 & 5.4 & 2.3 \\ 5.9 & 3.0 & 5.1 & 1.8 \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \\ \vdots & \vdots & \vdots & \vdots \\ x_{149,1} & x_{149,2} & x_{149,3} & x_{149,4} \\ x_{150,1} & x_{150,2} & x_{150,3} & x_{150,4} \end{pmatrix}$$

행row

열column

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

텐 서 (tensor)

- 텐서(tensor)
  - 3차원 이상의 구조를 가진 숫자 배열
  - 예) 3차원 구조의 RGB 컬러 영상: 2차원 행렬이 세 장 있음
  - 모든 차원을 포괄하는 표현.
  - 스칼라는 0차원 텐서, 벡터는 1차원 텐서, 행렬은 2차원 텐서
  - 6*6*3의 텐서 **A**

$$\mathbf{A} = \begin{array}{cccccccc}
 & & 4 & 1 & 0 & 3 & 2 & 2\\
 & 2 & 0 & 2 & 2 & 3 & 1 & 6\\
3 & 0 & 1 & 2 & 6 & 7 & 6 & 3\\
3 & 1 & 2 & 3 & 5 & 6 & 3 & 0\\
1 & 2 & 2 & 2 & 2 & 3 & 0 & 3\\
3 & 0 & 0 & 1 & 1 & 0 & 3 & 1\\
5 & 4 & 1 & 3 & 3 & 3 & 1 & \\
2 & 2 & 1 & 2 & 2 & 1 & &
\end{array}$$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

벡터의 곱

▷ **두 벡터의 곱**

$\boldsymbol{a} = (a_1, a_2, \ldots, a_n)'$, $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)'$

$\boldsymbol{a}$와 $\boldsymbol{b}$의 곱 : $\boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a} = \sum_{k=1}^{n} a_k b_k$: $(= \boldsymbol{a} \cdot \boldsymbol{b})$

$\boldsymbol{a} \cdot \boldsymbol{b}$로도 표현하고 **내적**(inner product or dot product)이라 부르기도 한다.

eg. $\boldsymbol{a}'\boldsymbol{a} = \sum a_i^2 = \|\boldsymbol{a}\|^2 \geq 0$

: $\|\cdot\|$는 벡터의 크기를 나타낸다(놈, norm).

위의 식에서 등호가 성립하는 경우는 $\boldsymbol{a} = \boldsymbol{0}$인 경우뿐이다.

※참고

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad : n \times 1 \text{ 열벡터} \Rightarrow$$

$$\boldsymbol{u}\boldsymbol{u}' = \begin{pmatrix} u_1^2 & u_1 u_2 & \cdots & u_1 u_n \\ u_2 u_1 & u_2^2 & \cdots & u_2 u_n \\ \vdots & \vdots & \vdots & \vdots \\ u_n u_1 & u_n u_2 & \cdots & u_n^2 \end{pmatrix} : n \times n \text{ 행렬},$$

$$\boldsymbol{u}'\boldsymbol{u} = \sum_{i=1}^{n} u_i^2 \quad : 1 \times 1 \text{ 스칼라}$$

벡터의 크기

벡터의 크기를 놈(norm)으로 측정

- 벡터의 p차 놈

$$p\text{차 놈: } \|\mathbf{x}\|_p = \left(\sum_{i=1,d} |x_i|^p\right)^{\frac{1}{p}} \tag{2.3}$$

$$\text{최대 놈: } \|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|, \cdots, |x_d|) \tag{2.4}$$

  - 예) $\mathbf{x} = (3\ \ -4\ \ 1)$ 일 때, 2차 놈은 $\|\mathbf{x}\|_2 = (3^2 + (-4)^2 + 1^2)^{1/2} = 5.099$
  - 2차놈: 유클리디언 놈(Euclidean norm)

- 단위벡터(unit vector): 길이가 1인 벡터 : $\frac{\mathbf{x}}{\|\mathbf{x}\|_2}$

예) $\underline{\mathbf{x}} = \binom{2}{1}$ 이면, $\|\underline{\mathbf{x}}\|_2 = \sqrt{2^2 + 1^2} = \sqrt{5}$ 이므로 길이가 1인 단위벡터(정규화된 벡터)

는 $\underline{\mathbf{e}} = \frac{1}{\sqrt{5}}\binom{2}{1} = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 0.8944 \\ 0.4472 \end{pmatrix}$ 이다.

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행렬의 크기

- **행렬**의 놈: 프로베니우스 놈 (Frobenius norm): 요소들의 제곱합의 제곱근

프로베니우스 놈: $\|\mathbf{A}\|_F = \left(\sum_{i=1,n} \sum_{j=1,m} a_{ij}^2 \right)^{\frac{1}{2}}$ (2.6)

예를 들어, $\left\| \begin{pmatrix} 2 & 1 \\ 6 & 4 \end{pmatrix} \right\|_F = \sqrt{2^2 + 1^2 + 6^2 + 4^2} = 7.550$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

벡터의 선형독립

벡터공간 $\mathbf{X}$ 에 속한 원소 $\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n} \in \mathbf{X}$ 의 선형결합(linear combination)은

$$c_1\mathbf{x_1} + c_2\mathbf{x_2} + \cdots + c_n\mathbf{x_n} = \mathbf{c}^T\mathbf{X}$$

라고 표현 가능하다. 영벡터가 아닌 $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$ 에 대해 선형결합이 0이 될 때, $\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}$ 는 선형종속이라고 하며, 오직 $\mathbf{c} = \mathbf{0}$ 일 때만 선형결합이 0이 된다면, 이를 선형독립이라고 한다.

- $\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n} \in \mathbf{X}$ 에 대하여,

$$c_1\mathbf{x_1} + c_2\mathbf{x_2} + \cdots + c_n\mathbf{x_n} = \mathbf{c}^T\mathbf{X} = \mathbf{0}$$

  을 만족하는 $\mathbf{c}$ 가 오직 $\mathbf{c} = \begin{pmatrix} c_1 & c_2 & \cdots & c_n \end{pmatrix}^T = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}^T$ 일 때, $\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}$ 는 선형독립(linearly independent)이라고 하고, $\mathbf{c} \neq \mathbf{0}$ 일 때 이를 선형종속(linearly dependent)이라고 한다.

벡터의 선형독립

Example: 아래의 벡터는 서로 선형 독립이다.

$$\mathbf{x_1} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{x_2} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{x_3} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

왜냐하면,

$$c_1\mathbf{x_1} + c_2\mathbf{x_2} + c_3\mathbf{x_3} = \begin{pmatrix} c_3 \\ c_2 - 2c_3 \\ c_1 - c_2 + c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

에서

$$\mathbf{c} = (c_1 \quad c_2 \quad c_3)^T = (0 \quad 0 \quad 0)^T$$

일 수 밖에 없기 때문이다.

벡터의 선형종속

선형종속의 개념을 자세히 살펴보면,

- 만일 식

$$c_1\mathbf{x_1} + c_2\mathbf{x_2} + \cdots + c_n\mathbf{x_n} = \mathbf{c}^{\mathbf{T}}\mathbf{X} = \mathbf{0}$$

의 $c_1, c_2, \cdots, c_n$ 중 적어도 한 개가 0 이 아니라고 가정하자.(i번째 계수인 c_i 가 0 이 아니라고 하자). 이 때,

$$\mathbf{x_i} = -\frac{c_1}{c_i}\mathbf{x_1} - \frac{c_2}{c_i}\mathbf{x_2} - \cdots - \frac{c_{i-1}}{c_i}\mathbf{x_{i\text{-}1}} - \frac{c_{i+1}}{c_i}\mathbf{x_{i+1}} - \cdots - \frac{c_n}{c_i}\mathbf{x_n}$$

라고 표현 가능하게 되어서 결국 $\mathbf{x_i}$ 가 나머지 원소들의 선형결합으로 표현되는 '종속관계'를 의미하게 된다.

벡터의 선형독립

◎ 예제 5.5 on p.150

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}$$

(1) $2\boldsymbol{x}_1 + \boldsymbol{x}_2 - \boldsymbol{x}_3 = \mathbf{0} \quad \Rightarrow \quad \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$: 선형종속

(2) $\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 = \mathbf{0}$이기 위해서는

$\alpha_1 - \alpha_2 = 0, \quad 2\alpha_1 = 0, \quad -2\alpha_1 + \alpha_3 = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = 0$이어야 하므로

$\Rightarrow \quad \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$: 선형독립

마찬가지로 $\{\boldsymbol{x}_1, \boldsymbol{x}_3\}$ $\{\boldsymbol{x}_2, \boldsymbol{x}_3\}$: 선형독립

◎ # 5.5 on p.193

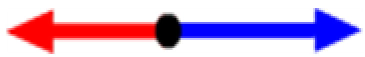
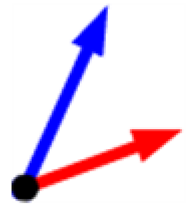
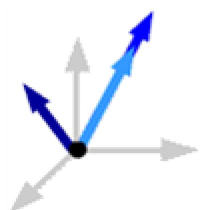
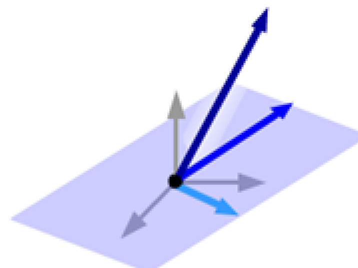
(1) $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix}$ (2) $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} 6 \\ 6 \\ 5 \end{pmatrix}$

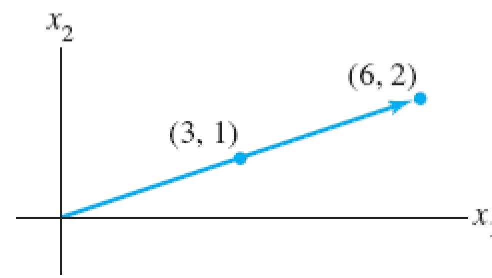
(1) $\boldsymbol{x}_3 = 2\boldsymbol{x}_1 + 2\boldsymbol{x}_2$이므로 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$는 선형종속이다.

(2) 선형연립방정식 $\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \alpha_3\boldsymbol{x}_2 = \mathbf{0}$을 풀면 $\alpha_1 = \alpha_2 = \alpha_3 = 0$

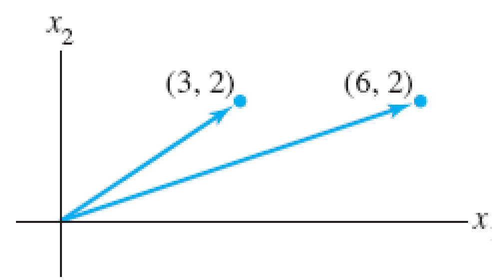
$\Rightarrow \quad \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$: 선형독립

벡터의 선형독립

	선형종속	선형독립
2차원		
3차원		



Linearly dependent



Linearly independent

참고: 선형독립인 벡터의 수가 rank임.
선형독립인 벡터들은 기저(basis)를 이루지만 반드시 직교하는 것은 아님.
벡터공간을 구성하는 기본 요소인 기저(basis)가 직교정규벡터일 때 이를 직교정규기저(orthonormal basis)라고 함.

선형결합과 벡터공간

벡터

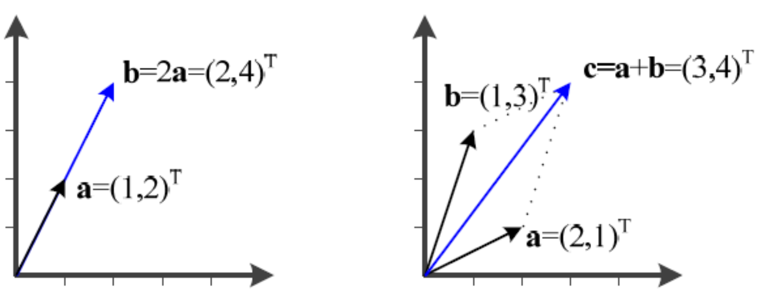
- 공간상의 한 점으로 화살표 끝이 벡터의 좌표에 해당

선형결합이 만드는 벡터공간(vector space)

- 기저벡터(basis vector) **a**와 **b**의 선형결합

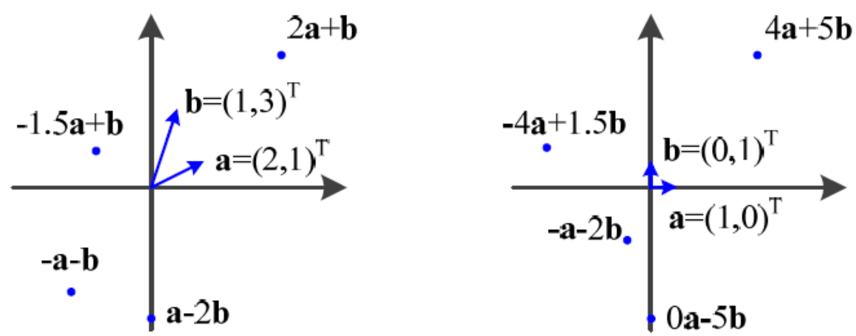
$$\mathbf{c} = \alpha_1 \mathbf{a} + \alpha_2 \mathbf{b} \tag{2.12}$$

- α_1과 α_2 를 변화시키면서 두 벡터 **a** 와 **b**의 선형결합으로 만들어지는 공간을 벡터공간이라 부름



(a) 벡터에 스칼라 곱 (b) 두 벡터의 덧셈

그림 2-6 벡터의 연산



(a) 기저 벡터와 벡터공간 (b) 정규직교 기저 벡터

그림 2-7 벡터공간

정규직교 (orthonormal) 기저벡터: 길이가 1이고, 서로 수직인 기저벡터

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행렬의 계수(rank)

- 행렬의 계수: 선형독립인 열(행)벡터의 수

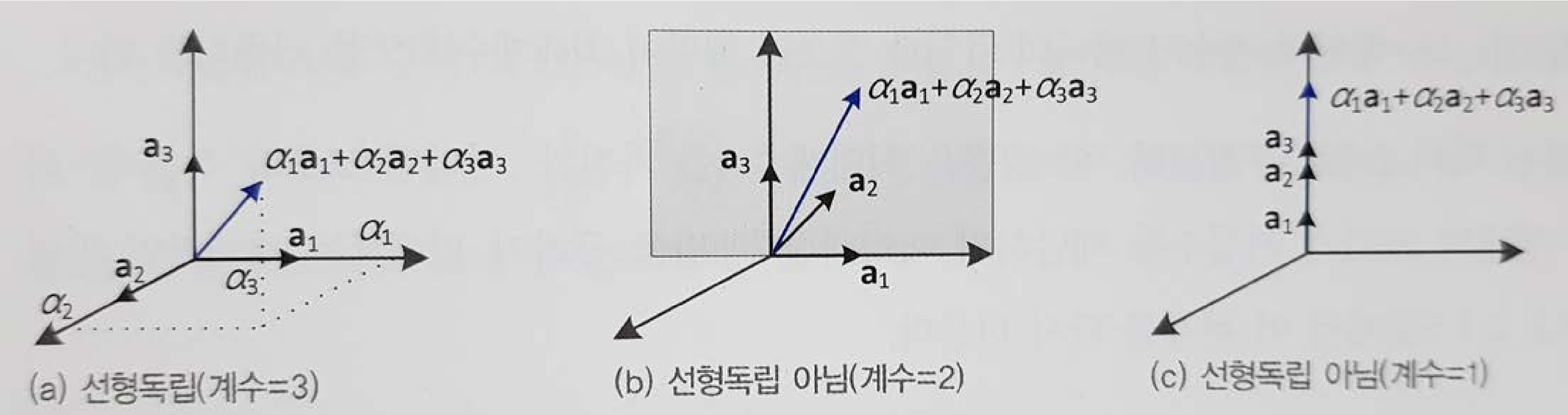


그림 2-8 3차원 벡터공간

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행렬의 계수(rank)

- 행렬 $\mathbf{A}$ 의 계수(rank) :

행 또는 열의 선형 독립인 벡터의 개수

0 이 아닌 행 또는 열의 숫자

예)

$$A = \begin{bmatrix} 2 & 4 & 1 & 3 \\ -1 & -2 & 1 & 0 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 2 & 5 \end{bmatrix}$$

첫 번째 열과 세 번째 열은 선형독립이지만, 두 번째 열은 첫 번째 열의 두 배와 같고 네 번째 열은 첫 번째 열과 세 번째 열의 합과 같으므로 $\mathbf{A}$ 의 계수는 2이다.

가우스 소거법을 통해 다음과 같이 나타낼 수 있다.

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

0이 아닌 행이 두 개임 => 계수가 2임.

출처: 위키피디아(ko.Wikipedia.org)

벡터의 유사도

유사도

- 벡터를 기하학적으로 해석

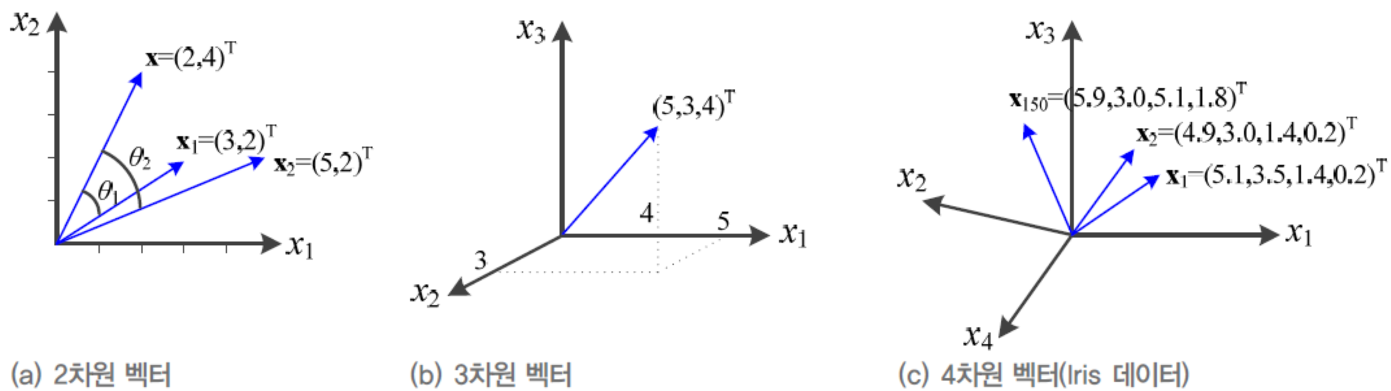


그림 2-2 벡터를 기하학적으로 해석

- 두 벡터의 유사도 측정을 위해서 두 벡터의 내적을 사용하면 된다. 하지만 벡터가 길기만 하면 유사도가 커지는 문제가 있다 따라서, 단위벡터 내적 사용.
- 코사인 유사도(cosine similarity): 단위벡터의 내적. 두 문서의 유사도 계산에 주로 사용.

$$cosine\_similarity(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}}{\|\mathbf{a}\|}\cdot\frac{\mathbf{b}}{\|\mathbf{b}\|} = cos(\theta) \tag{2.7}$$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행　　렬
(matrix)

행렬의 기초

a $\underline{a}$, $\vec{a}$, $\boldsymbol{a}$, $\mathbf{a}$ $\boldsymbol{A}$, A , $\mathbf{A}$

행렬의 표현

- 행렬은 여러 개의 원소들을 사각형 안에 배치시켜 놓은 것. 사각형이 행(row)을 m개, 열(column)을 n개 가질 때 원소의 총 개수는 mn개가 된다. 또 i 번째 행과 j 번째 열에 해당되는 원소를 a_{ij} 라 하고 이로 이루어지는 행렬을 A라 하면 이는 다음과 같이 표시된다.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$A = \{a_{ij}\}, \quad (i = 1, 2, \cdots, m; \; j = 1, 2, \cdots, n)$$

벡터와 행렬

▷ **벡터(vector)** : 행렬의 하나의 특별한 경우라 할 수 있는데,

- $n=1$인 경우의 행렬, $m\times1$행렬을 **열벡터(column vector)**라 하고,
- $m=1$인 경우의 행렬, $1\times n$행렬을 **행벡터(row vector)**라 한다.

[Note] 행벡터를 전치하면 열벡터를 얻고, 열벡터를 전치하면 행벡터를 얻는다.

▷ **스칼라(scalar)** : $m=n=1$ 경우의 원소가 하나인 행렬, **eg.** 실수

[Note] 스칼라를 전치하면 변화가 없이 같은 스칼라를 얻는다.

[Note] 보통 벡터를 스칼라와 구분하기 위해서 진하게(bold) 표현하고, 앞으로 특별히 다른 표시가 없으면 벡터는 열벡터를 의미한다.

행렬을 이루는 벡터

▷ **다양한 벡터**

- 영벡터 : 모든 원소가 0인 벡터, $\mathbf{0}$,
- 일벡터 : 모든 원소가 1인 벡터, $\mathbf{1}$,
- 항등행렬의 i번째 열벡터 : i번째 원소만 1이고 나머지 원소는 모두 0인 벡터, $\boldsymbol{e}_i$,

eg. $$\mathbf{0}=\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix},\quad \mathbf{1}=\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix},\quad \boldsymbol{e}_2=\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix},\quad I_5=(\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \boldsymbol{e}_3,\ \boldsymbol{e}_4,\ \boldsymbol{e}_5)$$

▷ **행렬을 이루는 벡터 :**

일반적으로 행렬 A의 i-번째 행벡터 : $\boldsymbol{a}_{i\cdot}' = (a_{i1},\ a_{i2},\ \ldots,\ \ a_{in})$,

j-번째 열벡터 : $\boldsymbol{a}_{\cdot j} = \begin{pmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{pmatrix}$

전치행렬(transpose matrix)

- 모든 (i, j) 쌍에 대해 a_{ij} 원소를 a_{ji} 로 전치하여 놓은 행렬A 의 전치행렬(transpose matrix, A^T 또는 A')이라 한다.

$$A = \{a_{ij}\}_{mn} \ , \ A^T = \{a_{ji}\}_{nm}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

전치행렬(transpose matrix)

eg.

$$A = \begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix} \Rightarrow A' = \begin{pmatrix} 2 & -1 \\ 0 & 2 \\ 3 & 0 \end{pmatrix}$$

[Note] 여기서 보는 바와 같이 A의 첫 번째 행벡터 $\boldsymbol{a}_{1\cdot}' = (2, 0, 3)$를 전치시키면

A'의 첫 번째 열벡터 $\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$를 얻을 수 있다.

◎ **예제**

$$A = (\boldsymbol{a}_{\cdot 1}, \ \boldsymbol{a}_{\cdot 2}, \ \boldsymbol{a}_{\cdot 3}, \ \boldsymbol{a}_{\cdot 4}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 1 & 2 & 3 \end{pmatrix} \Rightarrow A' = \begin{pmatrix} \boldsymbol{a}_{\cdot 1}' \\ \boldsymbol{a}_{\cdot 2}' \\ \boldsymbol{a}_{\cdot 3}' \\ \boldsymbol{a}_{\cdot 4}' \end{pmatrix} = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 6 & 1 \\ 3 & 7 & 2 \\ 4 & 8 & 3 \end{pmatrix}$$

전치행렬(transpose matrix)

행row

$$\mathbf{X} = \begin{pmatrix} 5.1 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 3.2 & 1.3 & 0.2 \\ 4.6 & 3.1 & 1.5 & 0.2 \\ \vdots & \vdots & \vdots & \vdots \\ 6.2 & 3.4 & 5.4 & 2.3 \\ 5.9 & 3.0 & 5.1 & 1.8 \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \\ \vdots & \vdots & \vdots & \vdots \\ x_{149,1} & x_{149,2} & x_{149,3} & x_{149,4} \\ x_{150,1} & x_{150,2} & x_{150,3} & x_{150,4} \end{pmatrix}$$

열column

- Iris의 설계 행렬을 전치행렬 표기에 따라 표현하면,

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_{150}^{\mathrm{T}} \end{pmatrix}$$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

대칭행렬(symmetric matrix)

- 행렬에서 행과 열의 차원이 같을 때($m = n$) 정방행렬 (정사각 행렬, square matrix)이라 한다. 행렬 A 가 정방행렬이면서

$$A = A^T$$

가 성립하는 행렬을 대칭행렬(symmetric matrix)이라 한다.

$$A = \begin{bmatrix} 1 & -4 & -5 \\ -4 & 2 & -6 \\ -5 & -6 & 3 \end{bmatrix} = A^T$$

대각행렬(diagonal matrix)

- 행렬이 대각요소만으로 이루어진 행렬을 대각행렬(diagonal matrix)이라 하고 대각행렬에서 모든 대각요소가 1인 행렬을 단위행렬(unit matrix 또는 identity matrix)이라 한다.

$$D = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 5 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

다항식의 행렬 표현

- 행렬을 이용하면 수학을 간결하게 표현할 수 있음
  - 예) 다항식의 행렬 표현

$$
\begin{aligned}
f(\mathbf{x}) &= f(x_1, x_2, x_3) \\
&= 2x_1x_1 - 4x_1x_2 + 3x_1x_3 + x_2x_1 + 2x_2x_2 + 6x_2x_3 - 2x_3x_1 + 3x_3x_2 + 2x_3x_3 + 2x_1 + 3x_2 - 4x_3 + 5 \\
&= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 2 & -4 & 3 \\ 1 & 2 & 6 \\ -2 & 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 2 & 3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + 5 \\
&= \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c
\end{aligned}
$$

- 특수한 행렬들

정사각행렬 $\begin{pmatrix} 2 & 0 & 1 \\ 1 & 21 & 5 \\ 4 & 5 & 12 \end{pmatrix}$, 대각행렬 $\begin{pmatrix} 50 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8 \end{pmatrix}$,

단위행렬 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, 대칭행렬 $\begin{pmatrix} 1 & 2 & 11 \\ 2 & 21 & 5 \\ 11 & 5 & 1 \end{pmatrix}$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

행렬의 덧셈

▷ **두 행렬의 합** : 두 행렬의 차수가 같아야 한다.

$A = \{a_{ij}\}$, $B = \{b_{ij}\}$ $m \times n$ 행렬

두 행렬의 합의 (i, j)-번째 원소 : $(A+B)_{i,j} = a_{ij} + b_{ij}$

eg. $A+B = \begin{pmatrix} 4 & 3 \\ -1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ -6 & 4 \end{pmatrix}$

▷ **행렬과 스칼라의 곱** : 모든 원소에 스칼라를 곱해서 얻을 수 있다.

$(\alpha A)_{ij} = \alpha a_{ij}$

eg. $3\begin{pmatrix} 4 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 12 & 9 \\ -3 & 6 \end{pmatrix}$

행렬의 곱셈

- 행렬의 곱셈에서는 피승수 행렬의 차원이 $m \times p$ 이면 승수 행렬의 행의 차수는 p 가 되어야 한다. 즉 피승수 행렬의 열의 수는 승수 행렬의 행의 수와 일치되어야 한다.

$$A_{2\times 3} \cdot B_{3\times 2} = C_{2\times 2} \qquad A_{2\times 3} \cdot B_{2\times 3} \neq C$$

- 2×3 행렬 A 와 3×2 행렬 B 를 곱한 결과 얻어진 행렬 C 의 차원은 2×2 가 되며, 곱셈 순서는 다음과 같다.

$$\begin{array}{ccccc} A & & B & & C \\ \begin{bmatrix} \boxed{a_{11} \; a_{12} \; a_{13}} \\ a_{21} \; a_{22} \; a_{23} \end{bmatrix} & \cdot & \begin{bmatrix} \boxed{\begin{matrix} b_{11} \\ b_{21} \\ b_{31} \end{matrix}} & \begin{matrix} b_{12} \\ b_{22} \\ b_{32} \end{matrix} \end{bmatrix} & = & \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \end{array}$$

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$$

행렬의 곱셈

- A 행렬의 첫째 행과 B 행렬의 첫째 열을 대응되는 것끼리 곱하여 합하면 C 행렬의 첫째 행과 첫째 열의 원소 c_{11}이 얻어진다. 이와 같은 방법으로 C 행렬을 구하면 다음과 같다.

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}$$

- 일반적으로 $m \times p$ 행렬 A 와 $p \times n$ 행렬 B 의 곱 $C = A \cdot B$ 의 차원은 $m \times n$ 이고, 각 원소는 다음의 식에 의해 계산된다.

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \quad , \quad \text{for all} \quad i, j$$

행렬의 곱셈

▷ **두 행렬의 곱** : A와 B를 AB로 곱하려면 : A의 열의 수 = B의 행의 수

A: $m \times p$행렬, B: $p \times n$행렬 $\Rightarrow$ AB는 $m \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{pmatrix}$$

- AB의 (i,j)-번째 원소 : $(AB)_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$

- 행렬에 곱셈에서는 다음과 같은 관계가 성립한다.

$A \cdot B \neq B \cdot A$

일반적으로 $(A \cdot B)^T = B^T \cdot A^T$

모든 정방행렬에 대하여 $A \cdot I = I \cdot A = A$ 가 성립한다

$(A \cdot B) \cdot C = A \cdot (B \cdot C)$

$C \cdot (A + B) = C \cdot A + C \cdot B$

행렬의 곱셈

◎ 예제4.2 on p.107

$$A=\begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix},\quad B=\begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{pmatrix},\quad C=\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$$AB=\begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{pmatrix}=\begin{pmatrix} 11 & 3 \\ -1 & 4 \end{pmatrix}$$

$$ABC=\begin{pmatrix} 11 & 3 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}=\begin{pmatrix} 11 & 6 \\ -1 & 8 \end{pmatrix}$$

$$3(AB+C)=3\left(\begin{pmatrix} 11 & 3 \\ -1 & 4 \end{pmatrix}+\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right)=3\begin{pmatrix} 12 & 3 \\ -1 & 6 \end{pmatrix}=\begin{pmatrix} 36 & 9 \\ -3 & 18 \end{pmatrix}$$

eg. 앞의 2×3 행렬 A에 대하여 $A'A$와 AA'를 구해보면 다음과 같이 전혀 다른 행렬을 얻게 된다.

$$A=\begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix},\quad A'=\begin{pmatrix} 2 & -1 \\ 0 & 2 \\ 3 & 0 \end{pmatrix}$$

$$A'A=\begin{pmatrix} 2 & -1 \\ 0 & 2 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix}=\begin{pmatrix} 5 & -2 & 6 \\ -2 & 4 & 0 \\ 6 & 0 & 9 \end{pmatrix},\qquad AA'=\begin{pmatrix} 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 0 & 2 \\ 3 & 0 \end{pmatrix}=\begin{pmatrix} 13 & -2 \\ -2 & 5 \end{pmatrix}$$

[Note] 행렬 A가 $n\times n$ 정사각행렬일 때 AA는 A^2으로 표현한다.

$$(A^2)_{ij}=\sum_{k=1}^{n}a_{ik}a_{kj}$$: A의 i-번째 행벡터와 A의 j-번째 열벡터의 곱

행렬식(determinant)

- $n \times n$ 정방행렬 $A = \{a_{ij}\}$ 의 행렬식(determinant)은 $|A|$ (또는 $\det(A)$)로 표시하며, 하나의 실수 값으로 표현된다. 3차 행렬까지의 행렬식은 다음과 같은 방법으로 구할 수 있다.

$$|a_{11}| = a_{11}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21}$$

$$- a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{32}a_{23}$$

행렬식이란 정사각행렬에 어떤 특정한 방법으로 하나의 수를 대응시키는 함수이다. 행렬식은 역행렬이 존재하는지를 판단하는 근거가 되며, 문자의 수와 식의 수가 같은 연립일차방정식의 근이 유일하게 존재하는지를 결정하는 데 중요한 역할을 한다.

행렬식의 영어 표현인 determinant는 determine에서 나온 것으로서 '결정'한다는 뜻을 내포한다.

출처: 네이버 지식백과 http://terms.naver.com/entry.nhn?docId=3338357&cid=47324&categoryId=47324

행렬식(determinant)

행렬식에 대해서는 다음과 같은 관계가 성립한다.

- 행렬 A와 B가 모두 정방행렬이면 $|AB| = |BA| = |A||B|$ 이다.
- $|A| = |A^T|$ 이다.
- 행렬 A가 대각행렬이면 $|A| = a_{11} \times a_{22} \times \cdots \times a_{nn}$ 이다.
- 행렬 A의 행벡터나 열벡터의 원소가 모두 0일 때, $|A| = 0$ 이다.
- 행렬 A가 $n \times n$ 정방행렬이면 상수 k 에 대하여 $|kA| = k^n|A|$ 이다.
- $A = \begin{pmatrix} 2 & -1 \\ 21 & 2 \end{pmatrix}$ => $|A| = 4 - (-21) = 25$
- $\left|\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = \left|\begin{pmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{pmatrix}\right| = (5-\lambda)^2 - 4 = (\lambda - 3)(\lambda - 7)$

역행렬(inverse matrix)

- $n \times n$ 정방행렬 $A = \{a_{ij}\}$ 의 행렬식이 0 이 아닐 때 (즉, $|A| \neq 0$이면), 만약 $AB = BA = I_n$을 만족시키는 행렬 B 가 존재한다면, 행렬 B를 행렬 A의 역행렬(inverse matrix)이라 하고 A^{-1} 로 표현한다.

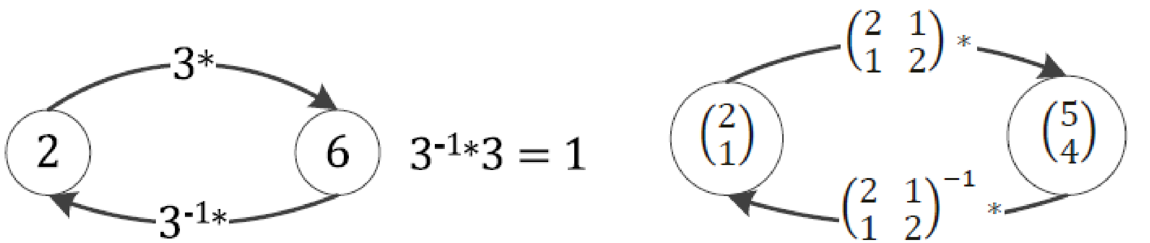
A의 역행렬 $A^{-1} \leftrightarrow AA^{-1} = A^{-1}A = I$

eg. $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad \Rightarrow \quad |A| = -2, \quad A^{-1} = \frac{1}{(-2)}\begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1.5 \\ 1 & -0.5 \end{pmatrix}$$

역행렬(inverse matrix)

역행렬의 원리



(a) 역수의 원리 (b) 역행렬의 원리

$$\begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}^{-1} * \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \mathbf{I}$$

그림 2-9 **역행렬**

- 정사각행렬 $\mathbf{A}$의 역행렬 $\mathbf{A}^{-1}$

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

- 예를 들어, $\begin{pmatrix}2 & 1\\6 & 4\end{pmatrix}$의 역행렬은 $\begin{pmatrix}2 & -0.5\\-3 & 1\end{pmatrix}$

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

정칙행렬(nonsingular matrix)

- $n \times n$ 정방행렬 $A = \{a_{ij}\}$ 에 대해 $AB = I$ 를 만족시키는 행렬 B가 유일하게 존재하면 행렬 B는 A^{-1}이 되고, 이 때 행렬 A를 정칙행렬 (nonsingular matrix) 이라 한다. 그렇지 않으면 행렬 A를 비정칙행렬 (특이행렬, singular matrix)이라 한다.

- A 가 정칙행렬이 되기 위한 필요 충분 조건은 $|A| \neq 0$ 이다.
- 또 A 가 정칙행렬이면, $|A^{-1}| = |A|^{-1}$ 이다.

- **정칙(nonsingular)행렬** : 행렬식이 0이 아닌 정사각행렬
- **특이(singular, 비정칙)행렬** : 행렬식이 0인 정사각행렬

▷ **역행렬** A : 정칙행렬일 때 역행렬이 존재한다. 즉, 어떤 행렬의 역행렬이 존재하기 위해서는 그 행렬이 정칙행렬이어야 한다.

직교행렬(orthogonal matrix)

- $n \times n$ 정방행렬 $A = \{a_{ij}\}$ 의 역행렬 A^{-1} 이 전치행렬 A^T 와 같으면, 행렬 A를 직교행렬(orthogonal matrix)이라 한다.

  즉, 직교행렬 $A \quad \leftrightarrow \quad A^{-1} = A^T$

- A 가 직교행렬이면 다음이 성립한다.

$$AA^T = A^T A = I$$

- $n \times n$ 정방행렬 $A = \{a_{ij}\}$의 각 열벡터를 $\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \dots \underline{\boldsymbol{a}}_n$ 이라 하면 행렬 A 가 직교행렬이기 위한 필요충분조건은 다음과 같다.

$$\underline{\boldsymbol{a}}_i^T \underline{\boldsymbol{a}}_j = \begin{cases} 0 & for \quad i \neq j \\ 1 & for \quad i = j \end{cases}$$

즉, 각 열벡터는 길이가 1이고 다른 열벡터와 직교한다. 이러한 벡터를 정규직교(직교정규, orthonormal) 벡터라고 한다.

고유값과 고유벡터

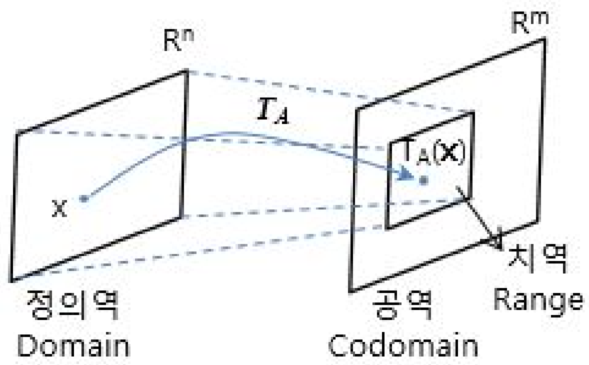
행렬 변환(벡터의 선형변환)

행렬 변환의 공간적 표현

$$T_A : R^n \rightarrow R^m$$

$$\mathbf{x} \xrightarrow{T_A} A\,\mathbf{x}$$

$$T_A(\mathbf{x}) = A\,\mathbf{x}$$



- n차원 벡터공간 R^n 안의 임의 벡터를 m차원 벡터공간 R^m 안의 임의 벡터로 보내는 변환
  - R^n 안의 임의 벡터 $\mathbf{x}$를 R^m 안의 벡터 $T_A(\mathbf{x}) = A\ \mathbf{x}$ 로의 대응 규칙
    - . T_A : 행렬변환
    - . A : m x n 행렬
    - . R^n : 변환 T_A의 정의역
    - . R^m : 변환 T_A의 공역
    - . 치역은, 정의역 모든 원소들의 상으로 만 이루어진 공역의 부분집합

출처: http://www.ktword.co.kr/abbr_view.php?m_temp1=4655

행렬 변환(벡터의 선형변환)

비례변환

T 는 R^2 안의 벡터 $\mathbf{x} = (x_1, x_2)$를 R^2 안의 벡터 $2\mathbf{x} = (2x_1, 2x_2)$로 사상하는 변환

$$(x_1, x_2) \xrightarrow{T} (2x_1, 2x_2) \qquad \text{또는} \quad T(\mathbf{x}) = 2\mathbf{x}$$

$$\mathbf{x} \xrightarrow{T} 2\mathbf{x}$$

If $\mathbf{x} = (-1, 3)$, $T(\mathbf{x}) = (-2, 6)$ ← $\mathbf{x}$ = (–1, 3)의 상(像)

행렬곱셈의 변환

- T_A는 R^2 안의 2×1 벡터 $\mathbf{x}$를 R^3 안의 3×1 벡터 $A\mathbf{x}$로 보내는 변환

$$T_A(\mathbf{x}) = A\mathbf{x} \qquad A\mathbf{x} = \begin{bmatrix} 1 & -1 \\ 2 & 5 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

$$\mathbf{x} \xrightarrow{T_A} A\mathbf{x}$$

$$\text{또는} \quad T_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

$$T_A\left(\begin{bmatrix} -1 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} -4 \\ 13 \\ 9 \end{bmatrix} \leftarrow \mathbf{x} = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \text{ 의 상(像)}$$

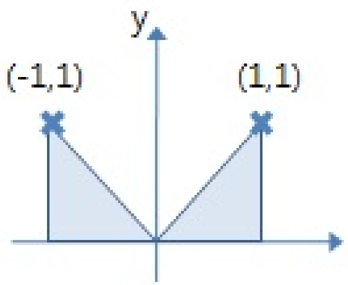
출처: 선형변환, 전북대학교 전자공학부 송상섭 강의노트

기하 변환(벡터의 선형변환)

반사변환

- y축 반사

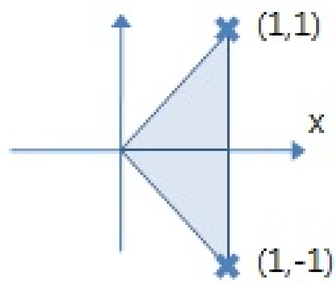
$$A=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$



$$T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

- x축 반사

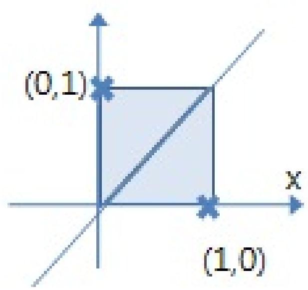
$$A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$



$$T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

- 직선 y=x 반사

$$A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$



$$T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

출처: http://www.ktword.co.kr/abbr_view.php?m_temp1=5768

1. 벡터의 선형변환

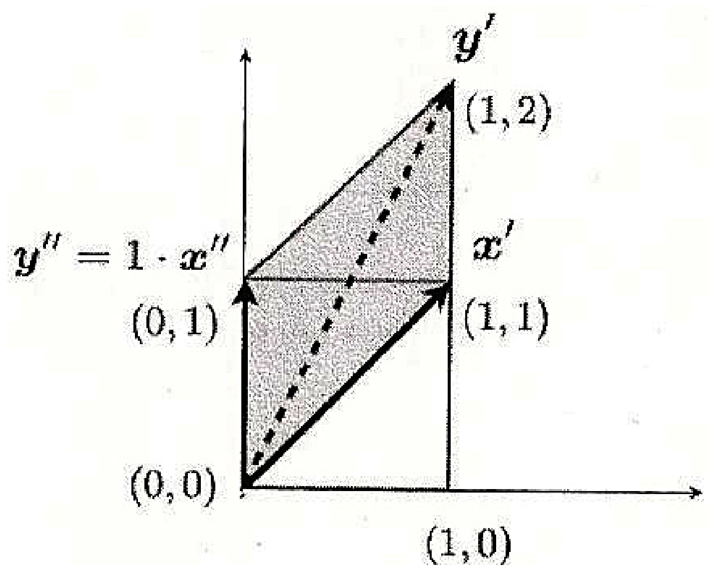
- 아래 행렬 $\mathbf{A}$ 는 $\mathrm{R}^n \to \mathrm{R}^n$ 로 변환시키는 선형변환 $\mathbf{y} = \mathbf{A}\mathbf{x}$ 에 사용된 행렬이며 아래와 같다.

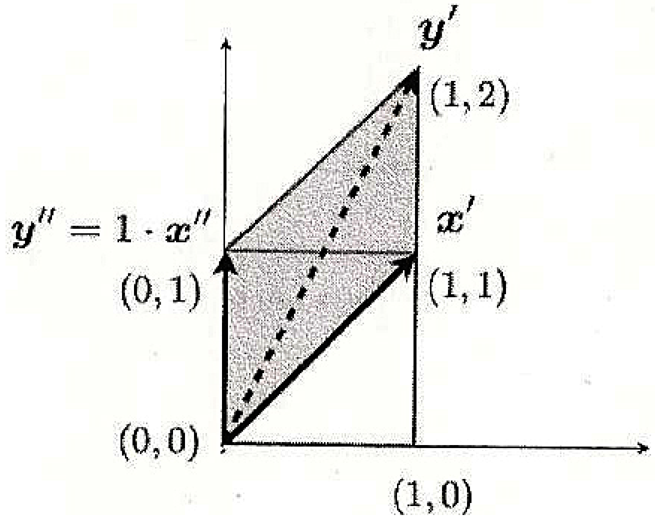
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- 위 선형변환은 아래 그림과 같이 정사각형을 평행사변형으로 변형시키는 역할을 한다.

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$





- 예를 들어, 점 $(1\,,\,1)$ 에 해당되는 벡터 $\mathbf{x}'$ 의 경우 점 $(1\,,\,2)$ 에 이르는 벡터 $\mathbf{y}'$ 으로 변환된다.

$$\mathbf{y}' = \mathbf{A}\mathbf{x}'$$

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- 그러나 점 $(0\,,\,1)$에 해당되는 $\mathbf{x}''$의 경우는 $\mathbf{A}\mathbf{x} = 1 \cdot \mathbf{x}''$ 를 만족하게 되며, 이 변환된 벡터는 변환 이전 벡터에 상수배 만큼만 변하게 된다. 이 경우, $\lambda = 1$ 이 $\mathbf{A}$ 의 고유값이 되고 $\mathbf{x}''$ 가 $\lambda = 1$ 에 대한 $\mathbf{A}$ 의 고유벡터가 된다.

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

- 정방행렬 $\mathbf{A}_{n\times n}$이 $R^n \to R^n$ 의 선형변환행렬 이라고 할 때, $\mathbf{y} = \mathbf{Ax}$라는 선형변환이 어떤 벡터 $\mathbf{x}$ 에서는 아래와 같이 방향은 같고 크기만 바뀌는 변환이 되기도 한다.

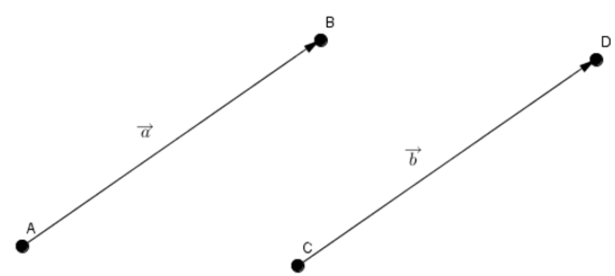
$$\mathbf{Ax} = \lambda\mathbf{x}, \qquad (\mathbf{x} \neq \mathbf{0})$$

- 이 때, 스칼라 λ를 "행렬 $\mathbf{A}$의 고유값(eigenvalue)"이라 하고, 벡터 $\mathbf{x}$를 "고유값 λ에 대한 고유벡터(eigenvector)" 라고 말한다.

- 정방행렬 $\mathbf{A}$에 대하여, $\mathbf{Ax} = \lambda\mathbf{x}$가 성립할 때, λ 를 고유값(eigenvalue), $\mathbf{x}$ 를 λ 에 따른 고유벡터(eigenvector)라고 한다. 이 때, $\mathbf{x}$ 는 영이 아닌 벡터이다.

- 참고: 벡터는 크기와 방향을 가진다.

서로 같은 벡터라 함은 '크기'와 '방향'이 둘다 같은 벡터임을 의미합니다.

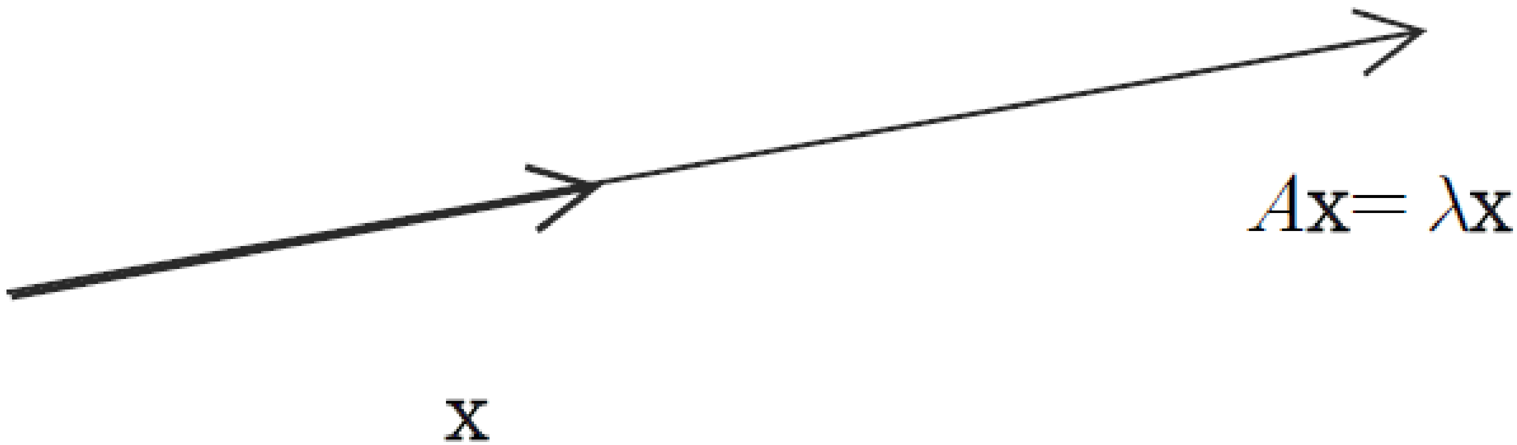


2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

A : n차의 정사각행렬

다음을 만족할 때 λ를 A의 고유값이라 하고 x를 λ에 대응하는 A 의 고유벡터라한다.

$$\Rightarrow Ax=\lambda x$$

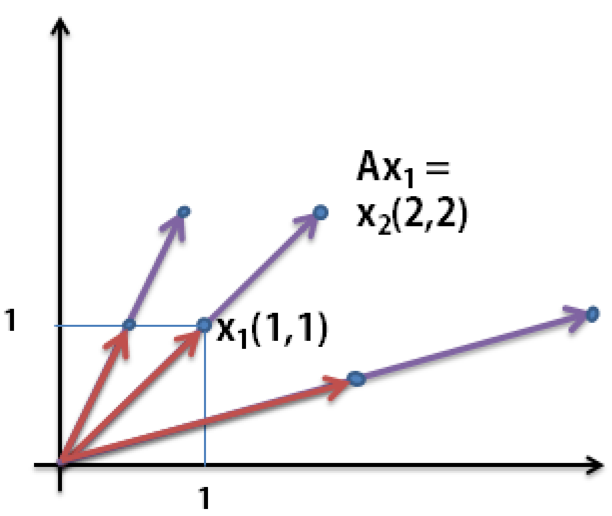


2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

- 또 다른 예) 점 (2, 2)에 해당되는 $\mathbf{x}_2$ 의 경우는 $\mathbf{A}\mathbf{x}_1 = 2 \cdot \mathbf{x}_1 = \mathbf{x}_2$ 를 만족하게 되며, 이 변환된 벡터는 변환 이전 벡터에 상수배 만큼만 변하게 된다.
  이 경우, $\lambda = 2$ 가 $\mathbf{A}$ 의 고유값이 되고 $\mathbf{x}_1$ 이 $\lambda = 2$ 에 대한 $\mathbf{A}$ 의 고유벡터가 된다.

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Linear Transformation



2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

정방행렬 A에 대하여 다음이 성립하는 0이 아닌 벡터 x가 존재할 때

$Ax = \lambda x$ (상수 λ)

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

상수 **λ**를 행렬 A의 **고유값 (eigenvalue)**,

x를 이에 대응하는 **고유벡터 (eigenvector)** 라고 함

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

예제:

정방행렬 A (square matrix A)	$\begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}$	
고유값 λ (eigenvalue)	λ = 7	λ = 2
고유벡터 x (eigenvector)	$\begin{pmatrix} 2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

행렬 A에 대한 고유값, 고유벡터

$$A\,x = \lambda\,x$$

$$\begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 7\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4\times 2 + 2\times 3 \\ 3\times 2 + 5\times 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 21 \end{pmatrix} = 7\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\underbrace{A}_{\text{square matrix}}\underbrace{x}_{\text{eigenvector}} = \underbrace{\lambda}_{\text{eigenvalue}}\,\underbrace{x}_{\text{eigenvector}}$$

$$A\,x = \lambda\,x$$

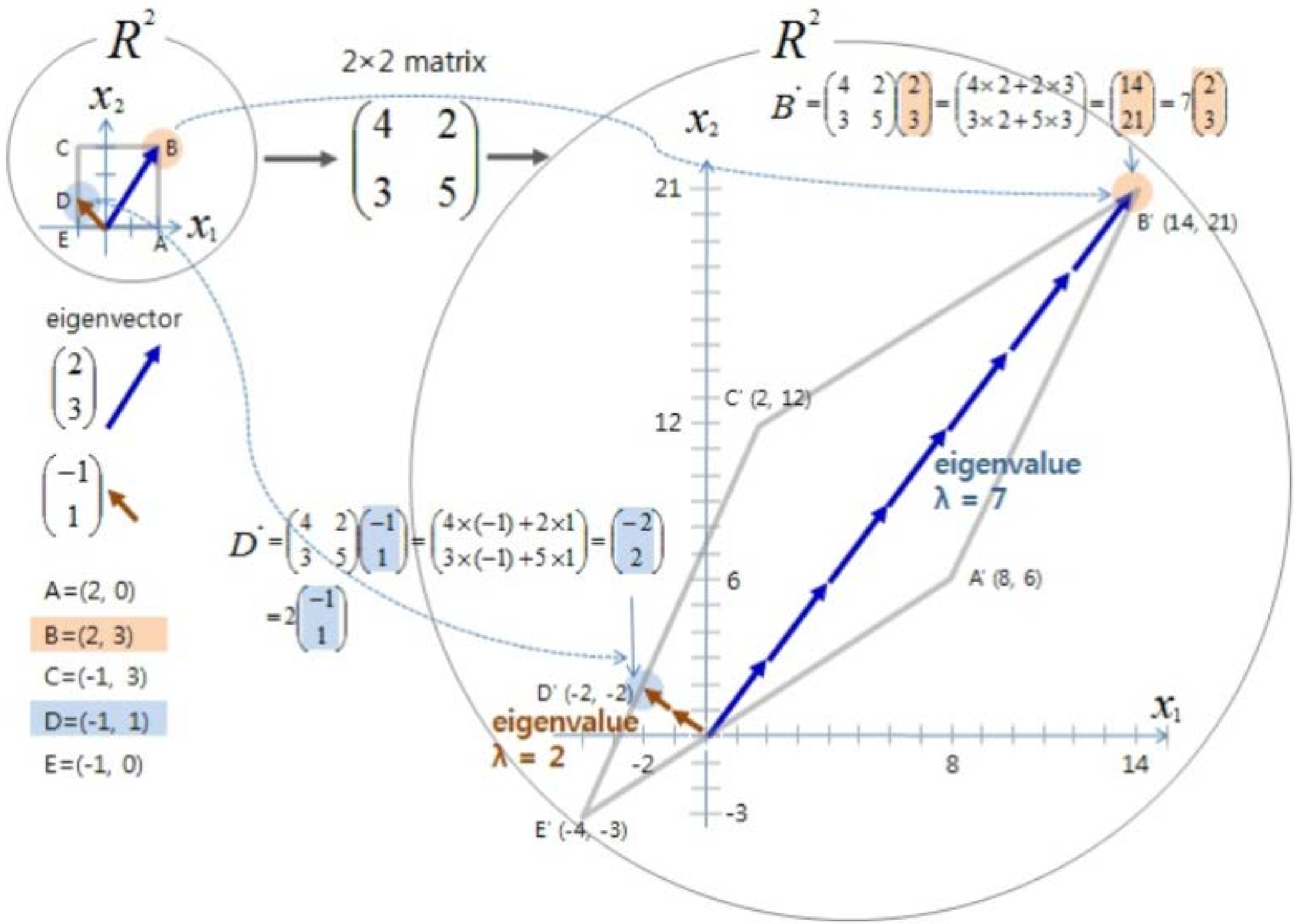
$$\begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4\times(-1) + 2\times 1 \\ 3\times(-1) + 5\times 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

2. 고유값(Eigenvalue)과 고유벡터(Eigenvector)

왼쪽의 A, B, C, D, E 의 좌표점들이 오른쪽에는 A', B', C', D', E' 로 정방행렬 A=(4, 3 2, 5)에 의해 변환되었습니다. 계산 예시로 B (2, 3) -> B' (14, 21)와 D (-1, 1) -> D' (-2, 2) 만 아래 그래프 위에 겹쳐서 제시해보았습니다.



출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

3. 고유값과 고유벡터 구하는 방법

◆ 행렬의 고유값을 구하기 위한 행렬식(determinant)의 사용법

- $\mathbf{Ax} = \lambda\mathbf{x}$ 이므로 정리하면 $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ 가 된다.

 이 때, 고유벡터 $\mathbf{x}$ 는 위 방정식을 만족하는 0이 아닌 벡터를 의미한다.

 위 방정식이 0 이 아닌 해를 갖기 위한 조건은 행렬 $\mathbf{A} - \lambda\mathbf{I}$ 가 비정칙행렬(singular) 임을 의미하며 이는 행렬식(determinant) $|\mathbf{A} - \lambda\mathbf{I}| = 0$ 임을 의미한다.

왜냐하면, $|(\mathbf{A} - \lambda\mathbf{I})| \neq 0$ 이라면 $(\mathbf{A} - \lambda\mathbf{I})^{-1}$ 가 존재하므로, $\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\,0 = 0$ 가 되어 $\mathbf{x}$ 가 영 아닌 벡터라는 가정에 모순된다.

그러므로 $|(\mathbf{A} - \lambda\mathbf{I})| = 0$ 을 만족시키는 λ 가 행렬 $\mathbf{A}$ 의 고유값이 될 것이다.

3. 고유값과 고유벡터 구하는 방법

◆ 행렬식(determinant)의 특징

- 행렬 A 의 고유값은

$$p(\lambda) = \left|(\mathbf{A} - \lambda\mathbf{I})\right| = 0$$

을 만족하는 λ 를 구하면서 얻게 된다. 여기서 $p(\lambda) = \left|(\mathbf{A} - \lambda\mathbf{I})\right| = 0$ 을 이 행렬 $\mathbf{A}$ 의 특성화방정식(characteristic equation)이라고 한다.

3. 고유값과 고유벡터 구하는 방법

◆ 고유벡터를 구하는 방법

- 고유값이 주어졌을 때, 고유벡터를 구하는 방법은

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

즉,

$$\underbrace{\left(\mathbf{A} - \lambda\mathbf{I}_n\right)}_{M}\mathbf{x} = 0$$

을 만족하는 $\mathbf{x}$ 를 구하는 방법과 일치한다.따라서, 위 방정식 $M\mathbf{x} = 0$ 의 해를 구하기 위해 확장행렬 $\left(M \mid \mathbf{0}\right)$에 대한 가우스-조단 알고리즘을 사용하여 해를 구한다.

- 행렬 $\mathbf{A}$ 의 고유값이 λ 일 때, 고유벡터는 $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ 을 만족하는 $\mathbf{x}$ 를 구함으로써 얻어진다.

3. 고유값과 고유벡터 구하는 방법

- 일반적으로 고유벡터는 단위길이를 갖도록 정규화된 고유벡터

$\mathbf{e_i} = \mathbf{x_i} / \sqrt{\mathbf{x_i}'\mathbf{x_i}}$ 로 표시된다.

행렬 $\mathbf{A}$ 가 $n \times n$ 정방행렬이면, 행렬 $\mathbf{A}$ 에서 n개 고유값과 정규화된 고유벡터 짝을 구할 수 있다.

$$((\lambda_1, \mathbf{e_1}),\ (\lambda_2, \mathbf{e_2}), \cdots, (\lambda_n, \mathbf{e_n}))$$

이 때 $\mathbf{A}$ 가 정방행렬이면서 **대칭행렬**이면, n개의 정규화된 고유벡터들은 단위벡터를 갖고 서로 직교를 이룬다. 이와 같은 성질을 벡터의 내적으로 표시하면 다음과 같다.

$$\mathbf{e_1}'\mathbf{e_1} = \mathbf{e_2}'\mathbf{e_2} = \cdots, = \mathbf{e_n}'\mathbf{e_n} = 1$$

$$\mathbf{e_i}'\mathbf{e_j} = \mathbf{0} \ \ (\mathbf{i} \neq \mathbf{j})$$

3. 고유값과 고유벡터 구하는 방법

Example 1 :

- 행렬 $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ 의 고유값과 고유 벡터는? (행렬 $\boldsymbol{A}$ 는 대칭행렬)

Sol) 특성화 방정식을 구하면 다음과 같다.

$$p(\lambda) = |(\mathbf{A} - \lambda \mathbf{I})| = \left| \left(\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \right| = \left| \begin{pmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{pmatrix} \right| = 0$$

이는

$$p(\lambda) = (1-\lambda)(-2-\lambda) - 2*2 = -2 - \lambda + 2\lambda + \lambda^2 - 4 = \lambda^2 + \lambda - 6$$

$$=> p(\lambda) = 0$$

$$=> \lambda^2 + \lambda - 6 = 0$$

$$=> (\lambda - 2)(\lambda + 3) = 0$$

$$=> \lambda = 2 \quad \text{또는} \quad \lambda = -3$$

3. 고유값과 고유벡터 구하는 방법

Example 1 (계속) :

- 행렬 $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ 의 고유값과 고유 벡터는? (행렬 $\boldsymbol{A}$ 는 대칭행렬)

Sol) $\lambda_1 = 2, \quad \lambda_2 = -3$

(1) $\lambda_1 = 2$ 에 대응하는 고유벡터 $\underline{\mathbf{x}}_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$ 는

$$\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = 2\begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} \rightarrow \begin{pmatrix} x_{11} + 2x_{12} \\ 2x_{11} - 2x_{12} \end{pmatrix} = \begin{pmatrix} 2x_{11} \\ 2x_{12} \end{pmatrix} \rightarrow \begin{matrix} x_{11} - 2x_{12} = 0 \\ 2x_{11} - 4x_{12} = 0 \end{matrix}$$

$\rightarrow x_{11} = 2x_{12}$ ($x_{11} = 0$ 이 아닌 상수)

무수히 많은 $x_{11} = 2x_{12}$ 중에 하나는 $\underline{\mathbf{x}}_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ 이고,

정규화된 첫 번째 고유벡터는 $\underline{\mathbf{e}}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 0.8944 \\ 0.4472 \end{pmatrix}$이다.

3. 고유값과 고유벡터 구하는 방법

Example 1 (계속) :

- 행렬 $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ 의 고유값과 고유 벡터는? (행렬 $\boldsymbol{A}$ 는 대칭행렬)

Sol) $\lambda_1 = 2, \quad \lambda_2 = -3$

(2) $\lambda_2 = -3$ 에 대응하는 고유벡터 $\underline{\mathbf{x}}_2 = \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}$ 는

$$\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = -3\begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} \rightarrow \begin{pmatrix} x_{21} + 2x_{22} \\ 2x_{21} - 2x_{22} \end{pmatrix} = \begin{pmatrix} -3x_{21} \\ -3x_{22} \end{pmatrix} \rightarrow \begin{matrix} 4x_{21} + 2x_{22} = 0 \\ 2x_{21} + x_{22} = 0 \end{matrix}$$

$\rightarrow 2x_{21} = -x_{22}$ ($x_{22} = 0$ 이 아닌 상수)

무수히 많은 $2x_{21} = -x_{22}$ 중에 하나는 $\underline{\mathbf{x}}_2 = \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ 이고,

정규화된 두 번째 고유벡터는 $\underline{\mathbf{e}}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 0.4472 \\ -0.8944 \end{pmatrix}$ 이다.

두 개의 정규화된 고유벡터의 내적 $\underline{\mathbf{e}}_1^T\underline{\mathbf{e}}_2 = 0$ 이므로 두 정규화된 고유벡터는 서로 직교이며 독립이다 (행렬 A 가 정방행렬이면서 대칭).

3. 고유값과 고유벡터 구하는 방법

Example2 : (행렬 $\boldsymbol{A}$ 는 대칭행렬이 아님)

행렬 $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$ 이라 하면 고유값은 $\begin{vmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = (\lambda^2 - 5\lambda + 6) - 2 = 0 \quad \Rightarrow \lambda = 4,\ 1$

$\lambda = 4$ 에 대응하는 고유 벡터 $\underline{u}_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}$ 은

$$\begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = 4\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} \Rightarrow \begin{pmatrix} 2u_{11} + 2u_{12} \\ u_{11} + 3u_{12} \end{pmatrix} = \begin{pmatrix} 4u_{11} \\ 4u_{12} \end{pmatrix}$$

$$\Rightarrow \begin{matrix} -2u_{11} + 2u_{12} = 0 \\ -3u_{11} + 3u_{12} = 0 \end{matrix} \Rightarrow u_{11} = u_{12}$$

(u_{11} =0 이 아닌 상수)

$\lambda = 1$ 에 대응하는 고유벡터 $\underline{u}_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$ 라 하면

$$\begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = 1\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} \Rightarrow \begin{pmatrix} 2u_{21} + 2u_{22} \\ u_{21} + 3u_{22} \end{pmatrix} = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

$$\Rightarrow \begin{matrix} u_{21} + 2u_{22} = 0 \\ u_{21} + 2u_{22} = 0 \end{matrix} \Rightarrow u_{21} = -2u_{22}$$

3. 고유값과 고유벡터 구하는 방법

Example 3 :

◎ 예제5.16 on p.166

$$A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\det(A - \lambda I) = \det\begin{pmatrix} (2-\lambda) & -1 & 0 \\ 0 & (1-\lambda) & 0 \\ 0 & 0 & (1-\lambda) \end{pmatrix} = (2-\lambda)(1-\lambda)^2 = 0$$

$\Rightarrow$ 고유값 : $\lambda = 1, 2$ (1은 중근)

- $\lambda = 2$ 에 대하여 : $A\boldsymbol{x} - 2I\boldsymbol{x} = (A - 2I)\boldsymbol{x} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$

$\Rightarrow$ $x_2 = 0,\ x_3 = 0$ $\Rightarrow$ 고유벡터 : $\boldsymbol{x} = (a, 0, 0)'$, eg. $\boldsymbol{x} = (1, 0, 0)'$

- $\lambda = 1$ 에 대하여 : $A\boldsymbol{x} - I\boldsymbol{x} = (A - I)\boldsymbol{x} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$

$\Rightarrow$ $x_1 = x_2$ $\Rightarrow$ 고유벡터 : $\boldsymbol{x} = (a, a, b)'$, eg. $\boldsymbol{x} = (1, 1, 0)'$, $\boldsymbol{x} = (0, 0, 1)'$

4. 고유값과 고유벡터의 성질

정방 행렬 $\mathbf{A}$ 의 고유값이 λ, 고유벡터가 $\mathbf{x}$ 라고 할 때, 임의의 상수 c에 대하여,

$$\mathbf{A}(c\mathbf{x}) = \lambda(c\mathbf{x})$$

가 성립하므로, 고유값 λ 에 대한 고유벡터는 $\mathbf{x}$ 뿐 아니라 $c\mathbf{x}$ 도 된다.
따라서, 고유값에 대한 고유벡터는 무한개 존재한다. 모든 상수배가 고유벡터이다.

그 중에서 정규화된 고유벡터는 $\mathbf{e_i} = \mathbf{x_i} / \sqrt{\mathbf{x_i'x_i}}$ 이다.

만약 $\mathbf{A}$ 가 대칭행렬인 경우 서로 다른 고유벡터는 서로 직교한다.

$$\text{즉, } \mathbf{e_i'e_j} = \mathbf{0} \quad (\mathbf{i} \neq \mathbf{j})$$

또한, 임의의 상수가 c 일때 $c\mathbf{A}$ 의 고유값은 $c\lambda$ 이다.

4. 고유값과 고유벡터의 성질

- 정방행렬의 고유값과 고유벡터는 행렬의 여러 가지 성질과 밀접한 연관이 있다. 그 중에서 행렬식(determinant)과 궤적(trace)에 대한 성질이 있다.

- 행렬 $\mathbf{A}$의 고유값이 $\{\lambda_i\}_{i=1}^{n}$라고 하면 다음과 같은 성질이 성립한다.

$$i)\quad \operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$$

$$ii)\quad \det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$$

4. 고유값과 고유벡터의 성질

- Proof)
  다음 행렬의 고유값을 생각해보자.

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

고유값을 구해보면

$$\left|\left(\mathbf{A} - \lambda\mathbf{I}\right)\right| = \left|\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}\right|$$

$$= \lambda^2 - \left(a+d\right)\lambda + \left(ad - bc\right)$$

가 되어 특성방정식이

$$p\left(\lambda\right) = \lambda^2 - \left(a+d\right)\lambda + \left(ad - bc\right)$$

$$= \left(\lambda - \lambda_1\right)\left(\lambda - \lambda_2\right)$$

라고 표현 가능하다. 이는 $p\left(\lambda\right)$ 전개를 통해

$$\lambda_1 + \lambda_2 = a + d = tr\left(A\right)$$

$$\lambda_1 \cdot \lambda_2 = ad - bc = \left|A\right|$$

임을 확인할 수 있다.

4. 고유값과 고유벡터의 성질

◎ <u>예제 5.19 on 170</u>

$A = \begin{pmatrix} 5 & -3 & 3 \\ 4 & -2 & 3 \\ 4 & -4 & 5 \end{pmatrix}$의 고유값은 1, 2, 5이므로

- $tr(A) = 5 - 2 + 5 = 8 = 1 + 2 + 5 = \lambda_1 + \lambda_2 + \lambda_3$
- $\det(A) = 5(-10 + 12) + 3(20 - 12) + 3(-16 + 8) = 10 = (1)(2)(5) = \lambda_1 \lambda_2 \lambda_3$

◎ <u>#5.32 on p.198</u> $\det(A) = 6,$ λ: 1,2,3

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ -1 & -1 & 2 \end{pmatrix} \Rightarrow \quad \det(A) = 6, \quad \lambda\text{: } 1,2,3$$

4. 고유값과 고유벡터의 성질

- 만일 행렬 $\mathbf{A}$ 의 고유값 중 하나 이상의 0 이 존재하면 행렬 $\mathbf{A}$ 는 비정칙행렬(singular matrix)로서 역행렬이 존재하지 않는다.

proof)

고유값 중에 하나 이상이 0 이 존재한다고 가정했을 때, 행렬식은

$$|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i = 0$$

이 되므로 $\mathbf{A}$ 는 비정칙행렬 이다.

- 행렬 $\mathbf{A}$ 의 고유값이 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 이고 고유벡터가 $\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_n}$ 일 때, 만일 고유값들이 서로 다른 값이면 $(\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_n)$, 고유벡터 $\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_n}$ 은 선형독립(linearly independent)이다.

n 차 대칭 정방행렬 $\mathbf{A}$ 의 어떤 두 개의 고유값 λ_1, λ_2 이고, 이에 대응하는 고유벡터가 $\mathbf{x_1}$, $\mathbf{x_2}$ 라 하면, $\lambda_1 \neq \lambda_2$ 일 때 $\mathbf{x_1}$ 과 $\mathbf{x_2}$ 는 직교한다.

즉, $\mathbf{x_1}^T \cdot \mathbf{x_2} = \mathbf{0}$ 이다.

5. 대칭-양정치행렬의 고유값

- 정방행렬 $\mathbf{A}_{n\times n}$ 에 대해서 $\mathbf{x}'\mathbf{A}\mathbf{x}$ 를 행렬 $\mathbf{A}$ 의 이차형식 (quadratic form) 이라 부른다. 이차형식은 하나의 값(스칼라)으로 나타난다. 일반적으로

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\ \sum_{j=i+1}^{n} a_{ij}x_i\, x_j$$ 로 표시된다.

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$$

$$\boldsymbol{x}'A\boldsymbol{x} = (x_1, x_2)\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 - 2x_2,\ -2x_1 + 5x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 - 4x_1x_2 + 5x_2^2$$

- 행렬에는 여러 개의 숫자가 존재하기 때문에 스칼라와 달리 어떤 행렬이 0보다 크다(>0) 혹은 적다(<0)라는 표현은 적합하지 않다. 그러나 행렬의 이차형식은 스칼라이므로 부호를 가질 수 있기 때문에 부호의 개념이 존재한다. 어떤 행렬의 이차형식이 항상 양일 때, 그 행렬을 "양정치 행렬(positive definite matrix)"라고 한다.

5. 대칭-양정치행렬의 고유값

- 행렬 $\mathbf{A}=\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ 이 양정치행렬 ?

  Sol) 영이 아닌 임의의 벡터 $\mathbf{x}$ 에 대하여,

  $$\mathbf{x}'\mathbf{A}\mathbf{x}=\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=x_1^2+x_2^2>0$$

  가 항상 성립하므로 ($\mathbf{x}\neq 0$), 행렬 $\mathbf{A}$ 는 양정치행렬이다

- 행렬 $\mathbf{A}_{n\times n}$ 에 대하여, 모든 고유값(eigenvalue)이 $\lambda_i>0, \quad i=1,2,\ldots n$ 이면 행렬 $\mathbf{A}$ 는 양정치 행렬이다.

- 요약: 양정치행렬 $\mathbf{A}$ 는:

  (1) 0 이 아닌 임의의 벡터 $\mathbf{x}$ 에 대하여 $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ 이 항상 성립한다.

  (2) 고유값이 모두 양수 (>0) 이다.

8. 대칭-양정치행렬의 고유값

- 스칼라에 부호를 부여하듯이 행렬에도 부호를 부여할 수 있을까?

- 양의 정부호 행렬(positive definite matrix) $\mathbf{A}$ if

양의 정부호 행렬 : $\mathbf{0}$이 아닌 모든 벡터 $\mathbf{x}$에 대해, $\mathbf{x}^T\mathbf{A}\mathbf{x}>0$

- 예를 들어, $\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}$는 $(x_1 \quad x_2)\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 + 2x_2^2$ 이므로

$\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}$는 양의 정부호 행렬

양의 정부호 행렬의 고유값은 모두 양수이다.
양의 준정부호 행렬의 고유값은 음수가 아니다. 즉 양수이거나 0이다.

양의 준정부호positive semi-definite 행렬: $\mathbf{0}$이 아닌 모든 벡터 $\mathbf{x}$에 대해, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$

음의 정부호negative definite 행렬: $\mathbf{0}$이 아닌 모든 벡터 $\mathbf{x}$에 대해, $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$

음의 준정부호negative semi-definite 행렬: $\mathbf{0}$이 아닌 모든 벡터 $\mathbf{x}$에 대해, $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$

정부호(definite) 행렬이 아닌 부정부호(indefinite) 행렬: 벡터에 따라서 음수가 되기도 하고, 양수가 되기도 한다.
부정부호 행렬의 고유값은 음수도 있고 양수도 있다.

출처: 기계학습 2장, pp.89-90, 오일석 지음, 한빛아카데미 (2017)

정방행렬 A_{n*n} 에 대하여

Inverse existing	Inverse not existing
A^{-1} 이 존재한다.	A^{-1} 이 존재하지 않는다.
full-rank 이다. rank(A)=n	full-rank 아니다. rank(A)<n
A는 정칙(non-singular)이다.	A는 비정칙(singular)이다.
|A| 는 0이 아니다.	|A| 는 0이다.
Ax=0의 해는 오로지 x=0만이다.	Ax=0의 해는 x=0 이 아닌 x도 존재한다.
A는 n개의 선형독립인 열(행)벡터를 가진다.	A는 n개 보다 적은 개수의 선형 독립인 열(행)벡터를 가진다.

▷ **정칙행렬(nonsingular matrix) (완전계수 행렬)**

A : $n \times n$ 행렬이고, $rank(A) = n \Rightarrow A$: 정칙행렬

정리

정리 2-1 다음 성질은 서로 필요충분조건이다.

- $\mathbf{A}$는 역행렬을 가진다. 즉, 특이행렬이 아니다.
- $\mathbf{A}$는 최대계수를 가진다.
- $\mathbf{A}$의 모든 행이 선형독립이다.
- $\mathbf{A}$의 모든 열이 선형독립이다.
- $\mathbf{A}$의 행렬식은 0이 아니다.
- $\mathbf{A}^{\mathrm{T}}\mathbf{A}$는 양의 정부호positive definite 대칭 행렬이다.
- $\mathbf{A}$의 고윳값은 모두 0이 아니다.

역행렬이 없는 행렬: 특이행렬(singular matrix)

출처: 기계학습, 오일석 지음, 한빛아카데미 (2017)

스펙트럼 분해

(spectral decomposition)
(eigenvalue decomposition)

특이값 분해

(singular value decomposition, SVD)

- 분해(decomposition)란?
  - 정수 3717은 특성이 보이지 않지만, 3*3*7*59로 소인수 분해를 하면 특성이 보이듯이, 행렬도 분해하면 여러모로 유용함

행렬의 분해

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T =$$

$+$

$+ \quad \cdots$

4.1 스펙트럼 분해(Spectral Decomposition)

- 여기서는 대칭행렬은 '스펙트럼 분해' , 대칭이 아닌 일반행렬의 경우에는 '특이값 분해'를 통해 여러 개의 '행렬의 합'으로 표현 가능

4.1 스펙트럼 분해

- 한 개의 **대칭행렬**이 여러 개의 합으로 분해

Theorem 4.1 (스펙트럼 분해)

행렬 $A_{n\times n}$이 대칭행렬이고, 고유값이 $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ 일 때, 직교 정규벡터로 이루어진

행렬 $P = (\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n)$ 와 고유값으로 이루어진 대각행렬 $D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$

에 대하여 행렬 A 는 다음과 같이 스펙트럼 분해된다.

$$A = PDP^T = \lambda_1 \underline{v}_1 \underline{v}_1^T + \lambda_2 \underline{v}_2 \underline{v}_2^T + \cdots + \lambda_n \underline{v}_n \underline{v}_n^T = \sum_{i=1}^{n} \lambda_i \underline{v}_i \underline{v}_i^T$$

4.1 스펙트럼 분해

Example : 행렬 $\mathbf{A}$의 고유값이 5, -5, 2 이며 고유벡터가 아래와 같을 때,

$\mathbf{A}$ 를 스펙트럼 분해하라.

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 3 & -4 \end{pmatrix}, \quad \underline{u}_1 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}, \quad \underline{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

(SOL) 위의 고유벡터는 서로 직교하되 길이가 1이 아니므로 정규화 시키면,

$$\underline{v}_1 = \begin{pmatrix} 0 \\ \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}, \quad \underline{v}_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix}, \quad \underline{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$ 가 된다.

4.1 스펙트럼 분해

따라서 $\mathbf{A}$는 다음과 같이 스펙트럼 분해된다.

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \underline{v}_i \underline{v}_i^T = 5\begin{pmatrix} 0 \\ \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}\left(0 \ \frac{3}{\sqrt{10}} \ \frac{1}{\sqrt{10}}\right) + (-5)\begin{pmatrix} 0 \\ \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix}\left(0 \ \frac{1}{\sqrt{10}} \ -\frac{3}{\sqrt{10}}\right) + 2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}(1 \ 0 \ 0)$$

즉,

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 3 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4.5 & 1.5 \\ 0 & 1.5 & 0.5 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.5 & 1.5 \\ 0 & 1.5 & -4.5 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.1 스펙트럼 분해

Example 2 :

◎ 예제5.22 on p.178

$$A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}, \quad |A - \lambda I| = 0 \quad \Rightarrow \quad \lambda_1 = 18, \quad \lambda_2 = 9, \quad \lambda_3 = 9$$

$$\Rightarrow \quad \boldsymbol{v}_1 = \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} \frac{1}{\sqrt{18}} \\ -\frac{1}{\sqrt{18}} \\ -\frac{4}{\sqrt{18}} \end{pmatrix}$$

$$\Rightarrow \; D = diag(18, 9, 9), \qquad V = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$$

정방행렬의 대각화: 고유값 분해(eigenvalue decomposition)

[고유값 분해(eigenvalue decomposition)]

$$A = PDP^{-1} = PDP^{T}$$

$$= \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}^{-1}$$

A : n * n 정방행렬(n by n square matrix)

λ_i : 고유값(eigenvalue)

v_i : 고유값 λ_i 에 대응하는 고유벡터(eigenvector)

P : 고유벡터 v_i 로 이루어진 행렬

D : 고유값 λ_i 로 이루어진 대각행렬(*diagonal matrix*)

> 행렬 A가 역행렬을 가질 조건: 0이 아닌 고유값들을 갖는다.
>
> 행렬 A가 대각화될 조건: n개의 선형 독립인 고유벡터들을 갖는다.

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

행렬 분해: 고유값 분해 (eigenvalue decomposition)

- 고유값 분해eigenvalue decomposition

$$A = PDP^{-1}$$

  - P 는 $\mathbf{A}$의 고유 벡터를 열에 배치한 행렬이고 D는 고유값을 대각선에 배치한 대각행렬
  - 예를 들어, $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$
  - 고유값 분해는 정사각행렬에만 적용 가능한데, 기계 학습에서는 정사각행렬이 아닌 경우의 분해도 필요하므로 고유값 분해는 한계를 가짐

=> 특이값 분해

출처: 기계학습 2장, pp.89-90, 오일석 지음, 한빛아카데미 (2017)

행렬 분해: 특이값 분해 (singular value decomposition, SVD)

- $n*m$ 행렬 **A**의 특이값 분해SVD(singular value decomposition)

$$A = UDV^T$$

  - 왼쪽 특이행렬 **U**는 $\mathbf{AA}^\mathrm{T}$의 고유 벡터를 열에 배치한 $n*n$ 행렬
  - 오른쪽 특이행렬 **V**는 $\mathbf{A}^\mathrm{T}\mathbf{A}$의 고유 벡터를 열에 배치한 $m*m$ 행렬
  - D는 $\mathbf{AA}^\mathrm{T}$의 고유값의 제곱근을 대각선에 배치한 $n*m$ 대각행렬

예를 들어, **A**를 4*3 행렬이라고 했을 때 다음과 같이 특잇값 분해가 된다.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -0.1914 & -0.2412 & 0.1195 & -0.9439 \\ -0.5144 & 0.6990 & -0.4781 & -0.1348 \\ -0.6946 & -0.6226 & -0.2390 & 0.2697 \\ -0.4651 & 0.2560 & 0.8367 & 0.1348 \end{pmatrix}$$

$$\begin{pmatrix} 3.7837 & 0 & 0 \\ 0 & 2.7719 & 0 \\ 0 & 0 & 1.4142 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -0.7242 & -0.4555 & -0.5177 \\ -0.6685 & 0.2797 & 0.6891 \\ 0.1690 & -0.8452 & 0.5071 \end{pmatrix}$$

출처: 기계학습 2장, pp.89-90, 오일석 지음, 한빛아카데미 (2017)

4.2 특이값 분해
(Singular Value Decomposition , SVD)

- **대칭행렬이 아닌 일반적인 $\boldsymbol{n \times m}$ 행렬에서** 적용 가능한 분해

  - 행렬근사(matrix approximation)나 최소제곱(Least squares)기법에서 적용되는 매우 유용한 분해법이다.

Theorem 4.2 (SVD)

행렬 $A_{n\times m}$에 대하여 $(n \geq m = Rank(A))$, 다음과 같은 분해를 특이값 분해라고 한다.

$$A_{n\times m} = U_{n\times m} D_{m\times m} V_{m\times m}^T$$

여기에서 (1) $U^T U = V^T V = I_m$

(2) $AA^T \underline{u}_i = \lambda_i \underline{u}_i \quad , i = 1,2,\ldots,m$

(3) $A^T A \underline{v}_i = \lambda_i \underline{v}_i \quad , i = 1,2,\ldots,m$

(4) D는 아래와 같이 정의 된다. $(\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \cdots \geq \sqrt{\lambda_m} > 0)$

$$D = \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_m} \end{pmatrix}$$

참고: 행렬 AA^T (혹은 $A^T A$) 은 대칭행렬

4.2 특이값 분해
(Singular Value Decomposition， SVD)

Proof : 이를 확인하기 위해 $A = UDV^T$ 라고 하면,

$$AA^T = UDV^TVDU^T = UD^2U^T = (\underline{u}_1, \underline{u}_2, \dots, \underline{u}_m)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix}(\underline{u}_1, \underline{u}_2, \dots, \underline{u}_m)^T$$

이것은 행렬 AA^T 의 스펙트럼 분해와 일치한다. 따라서 λ_i 는 AA^T의 고유값이고 $\underline{u}_i$는 직교하는 고유벡터이다. 마찬가지로

$$A^TA = VDU^TUDV^T = VD^2V^T = (\underline{v}_1, \underline{v}_2, \dots, \underline{v}_m)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix}(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_m)^T$$

가 되어 λ_i 는 A^TA의 고유값이고 $\underline{v}_i$는 직교하는 고유벡터이다.

참고: 행렬 AA^T (혹은 A^TA) 은 대칭행렬

4.2 특이값 분해
(Singular Value Decomposition， SVD)

Remark : 만일 λ_i 가 행렬 $\mathbf{A}^T\mathbf{A}$ (혹은 $\mathbf{AA}^T$)의 고유값이면, $\sqrt{\lambda_i}$ 는 행렬 $\mathbf{A}$ 의 특이값(singular value)이다.

Remark : 행렬 U 의 열벡터 $\underline{u}_i$ 를 ‘왼쪽 특이 벡터(left singular vector)’ , 행렬 V 의 열벡터 $\underline{v}_i$ 를 ‘오른쪽 특이벡터(right singular vector)’ 라고 한다.

Remark : Theorem 4.2 에서 $Rank(A) = n \le m$ 인 경우에는 다음과 같이 된다.

$$A_{n\times m} = U_{n\times n} D_{n\times n} (V_{m\times n})^T$$

4.2 특이값 분해
(Singular Value Decomposition， SVD)

Example1 행렬 $\mathbf{A}$ 를 특이값 분해(SVD)를 실시하여라.

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

(SOL) 행렬 $\mathbf{AA}^T$의 고유값과 고유벡터를 구하면,

$$\mathbf{AA}^T = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

그러므로 $\lambda_1 = 5$,$\lambda_2 = 1$ 이 고유값이 되며 고유벡터를 구하면,

$\underline{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 가 되어 $U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 이다.

한편 $\mathbf{A}^T\mathbf{A}$의 고유값과 고유벡터를 구하면,

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$ 에서

4.2 특이값 분해

고유값은 5, 1, 0 이 된다. 이제 0 이 아닌 고유값은 AA^T 때와 마찬가지로 5, 1 이 되고 직교 정규 벡터는

$$\underline{v}_1 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ -\frac{1}{\sqrt{5}} \end{pmatrix}, \quad \underline{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$ 가 되어 $$V = \begin{pmatrix} 2/\sqrt{5} & 0 \\ 0 & 1 \\ -1/\sqrt{5} & 0 \end{pmatrix}$$ 이다.

따라서 특이값 분해(SVD)는 다음과 같이 된다.

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 0 & -1/\sqrt{5} \\ 0 & 1 & 0 \end{pmatrix} = UDV^T$$

Proposition 4.1 (SVD의 전개)

행렬 $A_{n\times m}$에 대하여 $(Rank(A) = n \le m)$, 특이값 분해를 하면 행렬 A 는 다음과 같이 n개의 $n \times m$ 행렬의 합으로 표현 가능하다.

$$A_{n\times m} = U_{n\times n} D_{n\times n} (V_{m\times n})^T = \sum_{i=1}^{n} \sqrt{\lambda_i}\, \underline{u}_i \underline{v}_i^T$$

4.2 특이값 분해

Example1 : 앞의 Proposition 4.1을 이용하여

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = U_{2\times 2} D_{2\times 2} (V_{3\times 2})^T = \sum_{i=1}^{2} \sqrt{\lambda_i}\, \underline{u}_i \underline{v}_i^T$$

$$= \sqrt{5} \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} \left(\frac{2}{\sqrt{5}} \quad 0 \quad -\frac{1}{\sqrt{5}}\right) + 1 \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \quad 1 \quad 0)$$

$$= \begin{pmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$ 가 된다.

Remark :

행렬 $A_{n\times m}$에 대하여 $(n \geq m = Rank(A))$, 특이값 분해를 하면 행렬 A 는 다음과 같이 m개의 $n \times m$ 행렬의 합으로 표현 가능하다.

$$A_{n\times m} = U_{n\times m} D_{m\times m} V_{m\times m}{}^T = \sum_{i=1}^{m} \sqrt{\lambda_i}\, \underline{u}_i \underline{v}_i^T$$

특이값 분해(singular value decomposition)

Example 2 :

[1] 먼저 A X t(A) 의 고유벡터인 U를 구해보자

특이값 분해 예제 (example of SVD)

$$A=\begin{pmatrix}3 & 6\\2 & 3\\0 & 0\\0 & 0\end{pmatrix}$$ 라고 했을 때,

$$AA^T=\begin{pmatrix}3 & 6\\2 & 3\\0 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}3 & 2 & 0 & 0\\6 & 3 & 0 & 0\end{pmatrix}=\begin{pmatrix}45 & 24 & 0 & 0\\24 & 13 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}=W$$ 라고 하면,

$Wx=\lambda x$ 에서 $(W-\lambda I)x=0$

$$(W-\lambda I)x=\begin{pmatrix}45-\lambda & 24 & 0 & 0\\24 & 13-\lambda & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}=0$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

Example 2 :

위의 특성방정식(characteristic equations)을 풀면 (by using R),

$AA^T=W$의 고유값(eigenvalue) λ_i : $\lambda_1 = 57.844,\ \lambda_2 = 0.155,\ \lambda_3 = 0,\ \lambda_4 = 0$

고유값 (eigenvalue) λ_i에 대응하는
고유벡터(eigenvectors) x_i :

$$x_1^T = \begin{pmatrix} 0.881 \\ 0.471 \\ 0 \\ 0 \end{pmatrix} \quad x_2^T = \begin{pmatrix} -0.471 \\ 0.881 \\ 0 \\ 0 \end{pmatrix} \quad x_3^T = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad x_4^T = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\therefore U = \text{eigenvectors of } AA^T = \begin{pmatrix} 0.881 & -0.471 & 0 & 0 \\ 0.471 & 0.881 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

Example 2 :

[2] 다음으로, `t(A) X A` 의 고유벡터(eigenvectors)를 구해보겠습니다. 위의 (1)번 풀이 과정과 동일합니다.

$$A^T A = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 6 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 6 \\ 2 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 13 & 24 \\ 24 & 45 \end{pmatrix} = W_2 \quad \text{라고 하면,}$$

$$W_2 x = \lambda x \quad \text{에서} \quad (W_2 - \lambda I)x = 0$$

$$(W_2 - \lambda I)x = \begin{pmatrix} 13 - \lambda & 24 \\ 24 & 45 - \lambda \end{pmatrix} = 0$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

Example 2 :

위의 특성방정식(characteristic equations)을 풀면 (by using R),

$A^TA=W_2$의 고유값(eigenvalue) λ_i : $\lambda_1 = 57.844$, $\lambda_2 = 0.155$

고유값 (eigenvalue) λ_i에 대응하는
고유벡터(eigenvectors) x_i : $x_1^T = \begin{pmatrix} 0.471 \\ 0.881 \end{pmatrix}$ $x_2^T = \begin{pmatrix} -0.881 \\ 0.471 \end{pmatrix}$

$$\therefore V^T = \text{eigenvectors of } A^TA = \begin{pmatrix} 0.471 & -0.881 \\ -0.881 & 0.471 \end{pmatrix}$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

Example 2 :

[3] 다음으로, `A X t(A)` , `t(A)*A` 의 고유값(eigenvalue)의 제곱근(square root)을 특이값(singular value) 대각원소로 가지고 나머지는 '0'인 대각행렬 Σ 를 구해보겠습니다.

$AA^T=W$의 고유값(eigenvalue) λ_i : $\lambda_1 = 57.844, \quad \lambda_2 = 0.155, \quad \lambda_3 = 0, \quad \lambda_4 = 0$

특이값(singular value) σ_i :
$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{57.844} = 7.605$$
$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{0.155} = 0.394$$

$$\therefore \Sigma = \begin{pmatrix} 7.605 & 0 \\ 0 & 0.394 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (\sigma_1 \geq \sigma_2 \geq 0)$$

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

Example 2 :

[4] 위에서 구한 U, V^T, Σ 를 종합하면 끝이네요.

$$A = \begin{pmatrix} 3 & 6 \\ 2 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$ 일 때,

A의 특이값 분해 (Singular Value Decomposition)은

$$A = U\Sigma V^T = \underbrace{\begin{pmatrix} 0.881 & -0.471 & 0 & 0 \\ 0.471 & 0.881 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 7.605 & 0 \\ 0 & 0.394 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} 0.471 & -0.881 \\ 0.881 & 0.471 \end{pmatrix}}_{V^T}$$

AA^T 의 고유벡터
(eigenvectors of AA^T)

$\sigma_i = \sqrt{\lambda_i}$
($\lambda_1 \geq \lambda_2 \ldots \geq 0$이며,
대각원소이외 모두0)

A^TA 의 고유벡터
(eigenvectors of A^TA)

출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

특이값 분해(singular value decomposition)

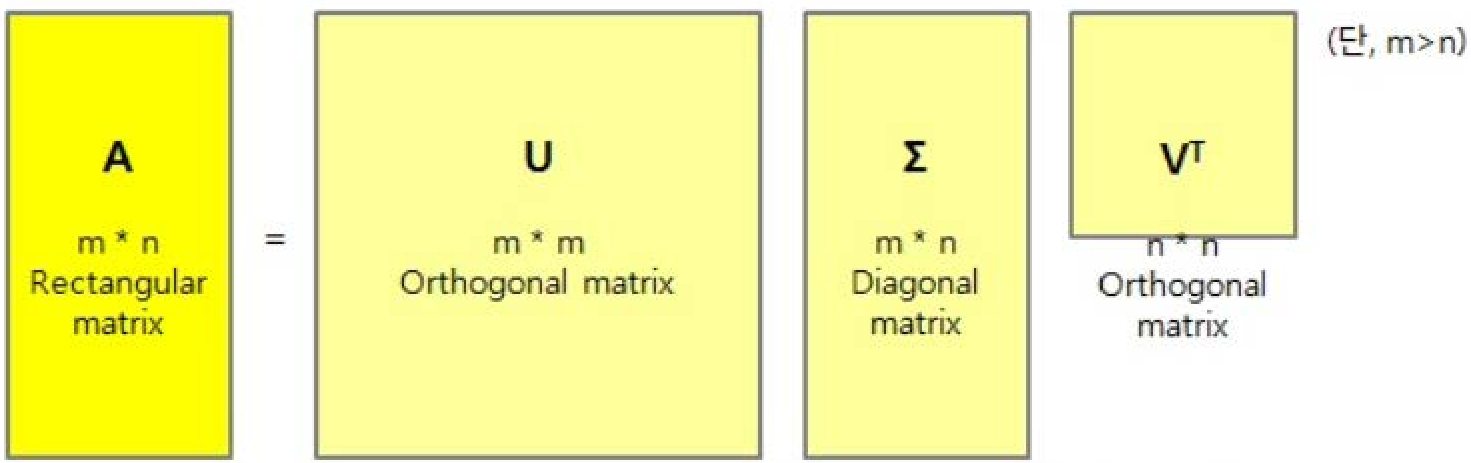
$$A = U\Sigma V^T$$

A : m * n 직사각행렬(m by n rectangular matrix)

U : A의 left singular vector로 이루어진 m * m 직교행렬(orthogonal matrix)

Σ : 주대각성분이 $\sqrt{\lambda_i}$로 이루어진 $m*n$직사각대각행렬(diagonal matrix)

V : A의 right singualr vector로 이루어진 n * n 직교행렬(orthogonal matrix)



출처: https://rstudio-pubs-static.s3.amazonaws.com/346037_78713fb40a144e749eaa6dab773a3571.html

4.3 특이값 분해의 활용

4.3 특이값 분해(SVD)의 활용

- SVD를 이용하여 여러 개의 행렬의 합으로 표현가능.

(1) 역행렬이 존재하지 않는 경우의 일반화역행렬(generalized inverse)계산

(2) 행렬의 계수(Rank) 계산(0 이 아닌 특이값의 개수 = 행렬의 계수)

(3) 행렬의 근사(matrix approximation)

(4) 영상 이미지 압축

4.3 특이값 분해의 활용

행렬의 근사

행렬 $A_{n\times m}$에 대하여 $(Rank(A) = n \leq m)$, 특이값 분해를 하면 행렬 A 는 다음과 같이 n개로 분해된 $n \times m$ 행렬의 합으로 표현 가능하다. ($\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \cdots \geq \sqrt{\lambda_m}$)

$$A_{n\times m} = U_{n\times n}D_{n\times n}(V_{m\times n})^T$$

$$= \sum_{i=1}^{n} \sqrt{\lambda_i}\underline{u}_i\underline{v}_i^T = \sqrt{\lambda_1}\underline{u}_1\underline{v}_1^T + \sqrt{\lambda_2}\underline{u}_2\underline{v}_2^T + \cdots + \sqrt{\lambda_k}\underline{u}_k\underline{v}_k^T + \sqrt{\lambda_{k+1}}\underline{u}_{k+1}\underline{v}_{k+1}^T + \cdots + \sqrt{\lambda_n}\underline{u}_n\underline{v}_n^T$$

여기서 어떤 $k(< n)$ 시점 이후의 특이값이 충분히 작다면,

즉, $\sqrt{\lambda_{k+1}}$, $\sqrt{\lambda_{k+2}}$, … , $\sqrt{\lambda_n}$ ≈ 0 ,

$$\mathbf{A} \approx \sqrt{\lambda_1}\underline{u}_1\underline{v}_1^T + \sqrt{\lambda_2}\underline{u}_2\underline{v}_2^T + \cdots \sqrt{\lambda_k}\underline{u}_k\underline{v}_k^T$$

행렬 $\mathbf{A}$ 는 $k(< n)$개의 행렬만으로도 충분히 근사 시킬 수 있을 것이다.

4.3 특이값 분해의 활용


Original (Rank 200)
Rank 1
Rank 2
Rank 3
Rank 5
Rank 10
Rank 15
Rank 25
Rank 50

4.3 특이값 분해의 활용

이와 같이 전체 m개로 이루어진 분해 중에서 일부분(k개)을 이용해 행렬을 근사 시키는 방법을 ‘k-계수근사(k-rank approximation)’ 이라고 하며, 아래는 행렬 $\mathbf{A}$의 2-계수 근사를 보여준다.

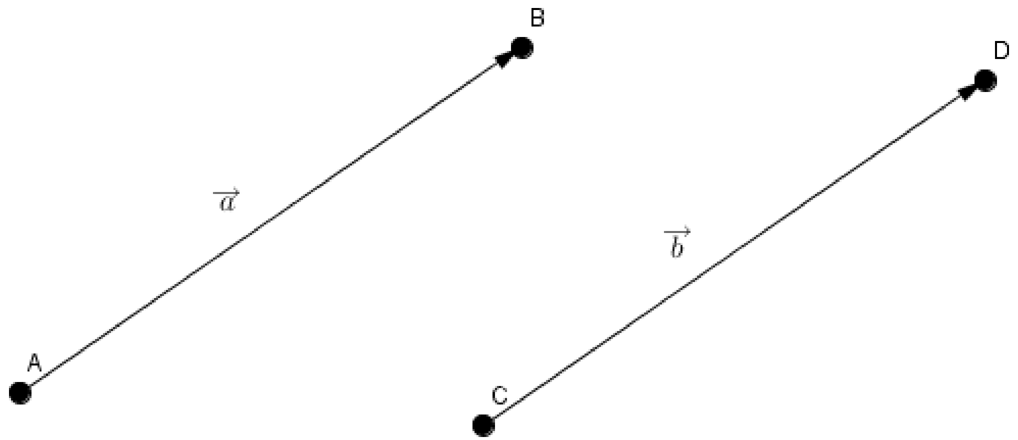
$$\mathbf{A} \approx \sqrt{\lambda_1}\underline{u}_1\underline{v}_1^T + \sqrt{\lambda_2}\underline{u}_2\underline{v}_2^T$$

이와 같이 행렬의 근사는 사이즈가 크고 복잡한 행렬을 간단한 몇 개의 벡터로 이루어진 행렬로 표현가능 하다는 의미.

$\Rightarrow$ 데이터 마이닝 , microarray 발현분석 같은 유전자 분석에 많이 사용.

부 록

서로 같은 벡터라 함은 '크기'와 '방향'이 둘다 같은 벡터임을 의미합니다.



벡터 a(AB)와 벡터 b(CD)가 서로 같은 벡터는 당연히 관계를

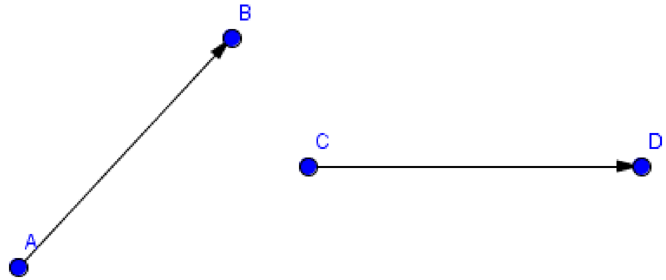
$\vec{a}=\vec{b}$, $\overrightarrow{AB}=\overrightarrow{CD}$

라고 나타내겠죠?

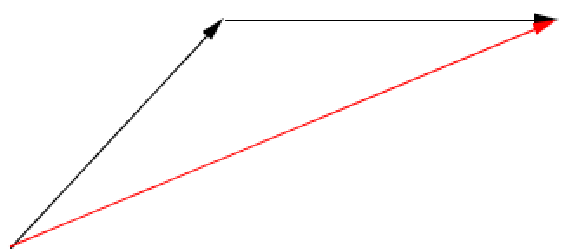
출처: http://blog.naver.com/PostView.nhn?blogId=ao9364&logNo=220740519031&categoryNo=9&parentCategoryNo=9&viewDate=&currentPage=1&postListTopCurrentPage=1&from=postView

<벡터의 덧셈>

아주 쉽습니다.



먼저 덧셈부터 해 볼께요. 벡터의 덧셈은 '시점과 종점' 만 잘 찾아내면 됩니다. 크게 두가지 방법이 있는데, 첫번째 방법은 삼각형 법 입니다.



한 벡터의 종점을 다른 벡터의 시점에 닿도록 평행이동 시킨 뒤, 닿지 않은 두 시점과 종점을 연결하면 됩니다.

결국 시점과 종점을 이으면 된다는 사실이죠!

출처:
http://blog.naver.com/PostView.nhn?blogId=ao9364&logNo=220740519031&categoryNo=9&parentCategoryNo=9&viewDate=&currentPage=1&postListTopCurrentPage=1&from=postView

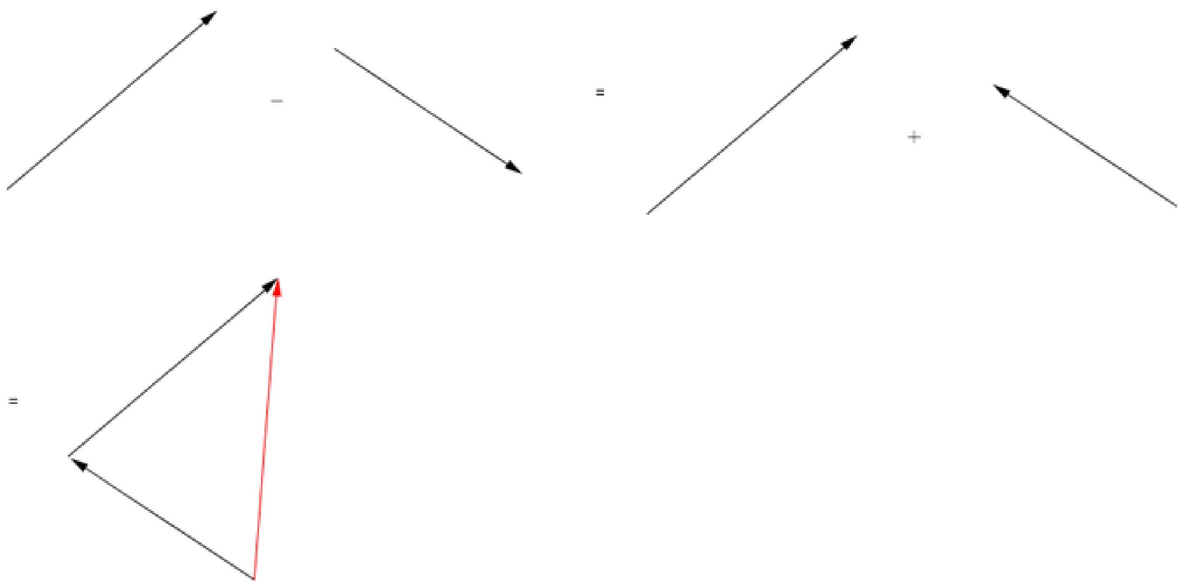
<벡터의 뺄셈>

가장 쉽게 생각하시는 방법은,

$\vec{a}-\vec{b}=\vec{a}+(-\vec{b})$

빼려는 벡터의 역벡터를 '더해준다' 라고 생각하시면 됩니다. 다른 방법으로는 뒤 벡터의 종점에서 앞 벡터의 종점으로 가는 벡터이다... 라는 설명도 있는데, 저는 개인적으로 헷갈릴때에는 역벡터를 더하는 방법이 제일 간단하더라고요.

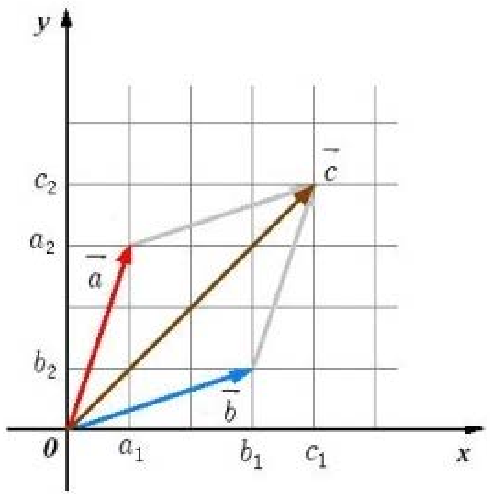
예를 들어보도록 하죠.



요렇게 방향만 바꿔서 더해주시면 된답니다~~

출처: http://blog.naver.com/PostView.nhn?blogId=ao9364&logNo=220740519031&categoryNo=9&parentCategoryNo=9&viewDate=&currentPage=1&postListTopCurrentPage=1&from=postView

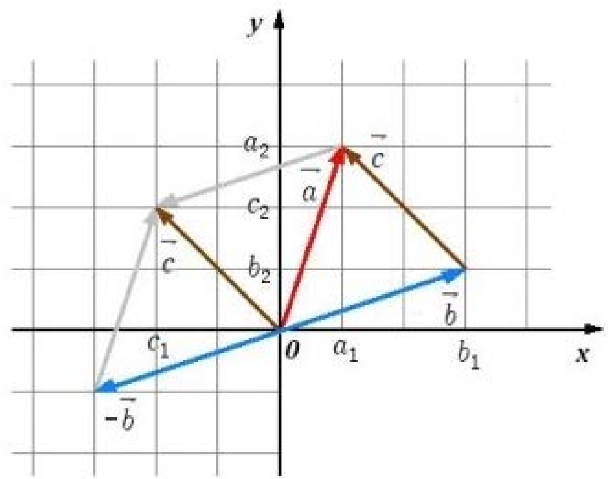
벡터의 덧셈



$\vec{a}=(a_1, a_2) \quad \vec{b}=(b_1, b_2)$

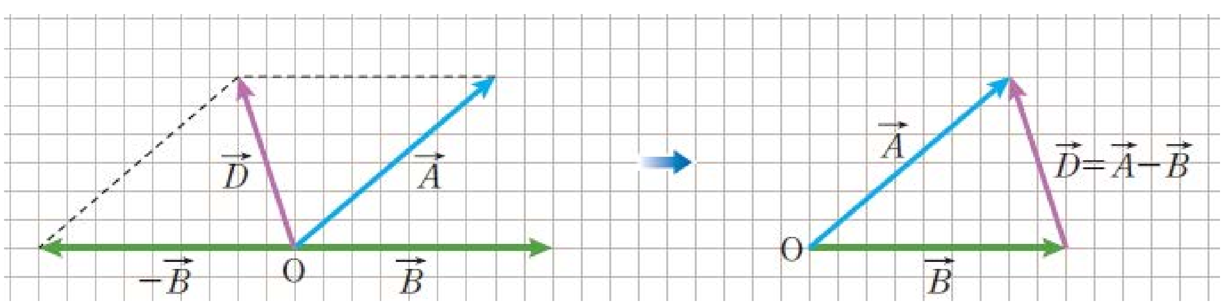
$\vec{a}+\vec{b}=\vec{c}=(a_1+b_1,\ a_2+b_2)=(c_1, c_2)$

벡터의 뺄셈



$\vec{a}=(a_1, a_2) \quad \vec{b}=(b_1, b_2)$

$\vec{a}-\vec{b}=\vec{c}=(a_1-b_1,\ a_2-b_2)=(c_1, c_2)$



사실 어떻게 정의하든 상관없지만 일반적인 내적의 정의는 다음과 같아요.

$a \cdot b = |a|\,|b| \cos\theta$

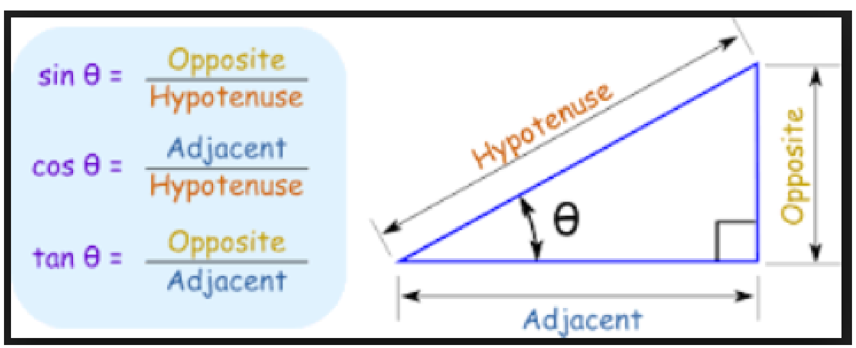
단 여기서 θ는 a라는 벡터와 b라는 벡터가 이루고 있는 0과 180도 사이의 각입니다.

$|a|$

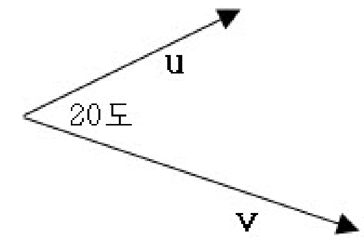
는 a 벡터의 크기를 의미하고

$|b|$

는 b 벡터의 크기를 의미합니다.



가령



가령 위와 같은 두개의 벡터가 있다고 해봐요.

u의 크기는 3, v의 크기는 5라고 가정해봅시당.

그럼

u와 v의 내적은 얼마일까요?

정의대로 하면 되지요.

$|u| = 3$

$|v| = 5$

$\theta = 20^\circ$

이 되니까 내적은

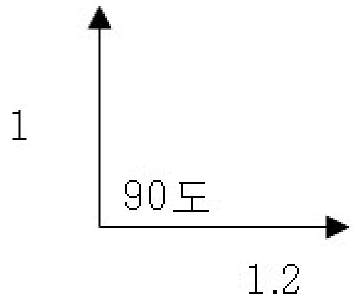
u와 v의 내적 =

$|u| \times |v| \times \cos\theta = 3 \times 5 \times \cos 20^\circ$

출처: http://tip.daum.net/question/54855462

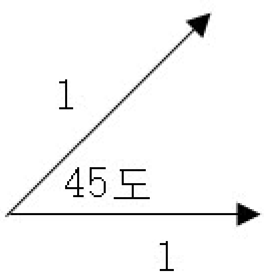
일단 내적은 그냥 정의로 받아들이시기 바랍니다.

그냥 연습삼아 문제를 몇개 풀어보죠.



이 문제의 내적은 얼마일까요?

답: 1*1.2*cos(90도)=0



이 문제의 내적은 얼마일까요?

답 :

$$1 \times 1 \times \cos 45^\circ = \frac{\sqrt{2}}{2}$$

원래 내적의 정의는

$$u \cdot v = |u||v|\cos\theta$$

이지만

만약 u와 v를 좌표계로

$$u = (u_1, u_2)$$

$$v = (v_1, v_2)$$

로 나타낼 수 있다면

그냥 간단하게

$$u \cdot v = u_1 v_1 + u_2 v_2$$

로 해도 결과가 똑같다는 말이죠.

이것은 다시 말해서

$$u \cdot v = |u||v|\cos\theta = u_1 v_1 + u_2 v_2$$

$$\Rightarrow \quad \cos\theta = \frac{u_1 v_1 + u_2 v_2}{|u||v|}$$

출처: http://tip.daum.net/question/54855462