빅데이터분석를 위한 수학

이용희

2023-12-10

목차

서론

이 온라인 연습장은 **빅데이터분석를 위한 수학**의 강의 보충 노트와 연습문제를 모아 놓은 사이트입니다.

- 이 연습장은 강의에 사용된 슬라이드와 부교재 Deisenroth, Faisal, 와/과 Ong (2020) 를 참고하였다.
  - 강의에 사용된 슬라이드는 서울시립대학교 온라인 강의실에서 다운로드 받을 수 있다.
  - 강의에 사용된 부교재는 교과서 웹사이트에서 다운로드 받을 수 있다.
- 이 교재의 각 장에서는 강의에 사용된 슬라이드에서 배운 내용을 보충 설명하고 반드시 학습해야 할 주요 주제를 설명한다.

i 노트

- 이 연습장의 각 장(chapter)의 내용은 강의에 사용된 슬라이드 번호의 내용과 일치합니다.
- 이 연습장에서 벡터와 행렬은 각각 $\boldsymbol{x}$, $\boldsymbol{A}$ 와 같이 볼드체로 표기하며 하나의 숫자를 나타내는 변수는 보통의 서체 x 로 표기한다.
- 정의, 정리, 예제 등이 끝나는 표시는 ■ 로 나타낸다.

1 강의 일정과 내용

1.1 강의 진도

- 1주차

슬라이드	주제	페이지 번호	부교재 내용
2	행렬의 도입	2-9, 13-17	
3	행렬의 연산	1-3, 6-7, 9-12	

- 2주차

슬라이드	주제	페이지 번호	부교재 내용
4	역행렬과 연립방정식의 해	1-5	
5	가우스소거법과 연립방정식의 해	1-21	27-32 페이지
6	행연산 행렬과 역행렬		33-34 페이지

- 3주차

슬라이드	주제	페이지 번호	부교재 내용
7	벡터공간	6-10	37-40 페이지
8	벡터공간의 기저와 차원	1-15	40-47 페이지
9	행렬의 계수	4, 7	47-48 페이지

- 4주차
  - 추석연휴
- 5주차

슬라이드	주제	페이지 번호	부교재 내용
10	선형사상	1,2,4,5,6,7	48-49 페이지
11	선형사상		50- 53 페이지
12	기저변환과 변환행렬		
13	선형변환의 핵과 상	1-2	58-60 페이지

슬라이드	주제	페이지 번호	부교재 내용
14	아핀공간		

- 6주차

슬라이드	주제	페이지 번호	부교재 내용
15	유클리드공간 위에서의 내적	2,3,5,6,7	71-78 페이지
16	벡터공간 위에서의 내적	1,5,7-12	71-78 페이지
17	직교기저	1-6	78-80 페이지
18	직선으로의 정사영	1-5	81-84 페이지

2 행렬의 도입

2.1 일차연립방정식

다음과 같이 n 개의 변수 $x_1, x_2, \dots, x_n$ 에 대한 m 개의 일차 방정식이 있다면 이를 일차연립방정식(a system of linear equations) 이라고 한다.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned} \tag{2.1}$$

위의 일차연립방정식(식 2.1) 에 사용된 변수 $x_1, x_2, \dots, x_n$ 와 계수 a_{ij}, y_i 으로 좀 더 보기 좋고 효율적으로 표현하기 위하여 행렬 $\boldsymbol{A}$ 와 벡터 $\boldsymbol{x}, \boldsymbol{y}$ 를 다음과 같이 표기하여

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

식 2.1 의 일차연립방정식을 다음과 같이 표현할 수 있다.

$$\boldsymbol{Ax} = \boldsymbol{y}, \text{ 즉 } \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \tag{2.2}$$

식 2.2 은 y_i 의 값을 계산하는 방법이 벡터 $\boldsymbol{x}$ 의 변수 $x_1, x_2, \dots, x_n$ 와 행렬 $\boldsymbol{A}$ 의 i 번째 행에 있는 계수들 $a_{i1}, a_{i2}, \dots a_{in}$ 을 다음과 같은 식으로 계산한다는 의미이다. 즉 일차연립방정식(식 2.1) 을 행렬 $\boldsymbol{A}$ 와 벡터 $\boldsymbol{x}, \boldsymbol{y}$ 로 표현한 것이다.

$$\sum_{i=j}^{n} a_{ij}x_j = y_i, \quad i = 1, 2, \dots, m$$

이제 위에서 일차연립방정식을 표현할 때 사용한 벡터와 행렬의 정의와 기본 연산에 대하여 알아보자.

2.2 행렬과 벡터

2.2.1 행렬

m 개의 행과 n 개의 열을 가진, 즉 $m \times n$ 행렬은 보통 알파벳 대문자(upper case letter)로 표현하며 다음과 같은 형태로 나타낸다.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij}) \ (i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$$

행렬 $\boldsymbol{A}$ 가 m 개의 행과 n 개의 열을 가진 행렬이라면 다음과 같이 표시한다.

$$\boldsymbol{A} \in \mathbb{R}^{m \times n}$$

2.2.2 벡터

벡터(vector)는 일반적인 행렬의 하나의 행 또는 하나의 열을 나타내는 이름으로 사용된다.

- 행렬의 각 행은 $1 \times n$ 행렬 혹은 행벡터 (row vector)라고 한다.
- 행렬의 각 열은 $m \times 1$ 행렬 혹은 열벡터 (column vector)라고 한다.

벡터는 다음과 같이 숫자를 모아 놓은 형태에 따라서 행벡터($\boldsymbol{r}$)와 열벡터($\boldsymbol{c}$)로 구분할 수 있다.

$$\boldsymbol{r} = \begin{bmatrix} 1 \ 2 \ 3 \ 4 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

또한 벡터는 위치를 나타내는 개체 (geometric vector)로 사용할 수 있다. 위치의 개념을 더 확장하면 벡터는 n 개의 숫자(element)를 순서 있게 모아 놓은 모든 집합, 즉 유클리디안 공간(Euclidean space; $\mathbb{R}^n$) 을 구성하는 개체로 사용할 수 있다.

2.3 중요한 내용과 정의

- 두 행렬이 같다는 정의

$$\boldsymbol{A} = \boldsymbol{B} \quad \Leftrightarrow \quad a_{ij} = b_{ij} \quad \forall i, j$$

- 정방행렬(square matrix)

- 대각행렬(diagonal matrix)
- 상삼각 행렬(upper triangular matrix)과 하삼각행렬(lower triangular matrix)
- 영행렬(zero matrix)
- 단위행렬(identity matrix)
- 대칭행렬(symmetric matrix)
- 스칼라(scalar)

3 행렬의 연산

3.1 행렬의 덧셈과 스칼라곱

3.1.1 덧셈

두 행렬 $\boldsymbol{A}$ 와 $\boldsymbol{B}$ 를 더하는 규칙은 다음과 같다.

- 두 행렬 $\boldsymbol{A}$ 와 $\boldsymbol{B}$ 는 행과 열의 갯수가 같아야 한다.
- $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{C}$ 라고 하면, 덧셈의 결과로 만들어진 행렬 $\boldsymbol{C}$는 두 행렬과 같은 수의 행과 열을 가지면 각 원소는 다음과 같다.

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{C} \quad \rightarrow \quad c_{ij}=a_{ij}+b_{ij}$$

3.1.2 스칼라곱

임의의 실수 λ (스칼라)가 주어졌을 때, λ 와 행렬 $\boldsymbol{A}$의 스칼라곱(scalar product) 는 행렬의 모든 원소에 λ 를 곱해준 행렬로 정의된다.

예를 들어 $\lambda=2$, $\boldsymbol{A}\in\mathbb{R}^{2\times3}$ 인 경우

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

3.2 행렬의 곱셈

먼저 두 행렬 $\boldsymbol{A}$ 와 $\boldsymbol{B}$ 의 곱셈

$$\boldsymbol{A}\times\boldsymbol{B}\equiv\boldsymbol{AB}$$

을 정의하려면 다음과 같은 조건이 만족되어야 한다.

- 행렬 $\boldsymbol{A}$ 의 열의 갯수와 행렬 $\boldsymbol{B}$ 의 행의 갯수가 같아야 한다

따라서 두 행렬의 곱셈은 순서를 바꾸면 정의 자체가 안될 수 있다.

정의 3.1 (곱셈의 정의). 이제 두 행렬 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 와 $\boldsymbol{B} \in \mathbb{R}^{n\times k}$ 의 곱셈은 다음과 같이 정의된다.

$$\boldsymbol{AB} = \boldsymbol{C}$$

행렬 $\boldsymbol{C}$ 는 m 개의 행과 k개의 열로 구성된 행렬이며($\boldsymbol{C} \in \mathbb{R}^{m\times k}$) 각 원소 c_{ij}는 다음과 같이 정의된다.

$$c_{ij} = \sum_{l=1}^{n} a_{il}b_{lk}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, k$$

먼저 간단한 예제로 다음과 같은 두 개의 행렬의 곱을 생각해 보자.

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} (1)(0) + (2)(-1) & (1)(1) + (2)(2) \\ (3)(0) + (4)(-1) & (3)(1) + (4)(2) \end{bmatrix} = \begin{bmatrix} -2 & 5 \\ -4 & 11 \end{bmatrix}$$

곱하는 순서를 바꾸어 계산해 보자.

$$\boldsymbol{BA} = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (0)(1) + (1)(3) & (0)(2) + (1)(4) \\ (-1)(1) + (2)(3) & (-1)(2) + (2)(4) \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$$

위 두 결과를 보면 행렬의 곱셈에서는 교환법칙이 성립하지 않음을 알 수 있다.

이제 차원이 다른 두 행렬의 곱셈을 살펴보자.

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

두 행렬의 곱셈은 정의 3.1 에 위하여 다음과 같이 계산할 수 있다.

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 2 & 5 \end{bmatrix}$$

두 행렬의 곱하는 순서를 바꾸면 차원이 전혀 다른 행렬이 얻어진다.

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

3.3 중요한 내용과 정의

- 행렬의 전치(transpose operation): $\boldsymbol{A}^T$
- 행렬의 더하기와 스칼라곱의 성질
- 행렬의 곱셈은 교환법칙이 성립하지 않는다.

$$\boldsymbol{AB} \neq \boldsymbol{BA} \tag{3.1}$$

> 주의
>
> **교환법칙이 성립하지 않는다**는 의미는 식 3.1 이 언제나 성립한다는 의미는 아니다. 아래와 같이 특별한 경우 교환법칙이 성립하는 경우도 있다.
>
> $$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

- 행렬의 곱셈은 결합법칙과 배분법칙은 성립한다.

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$

$$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$$

4 행렬과 연립방정식의 해

4.1 역행렬의 정의

정방행렬 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 의 역행렬(inverse metrix)이 존재하면 $\boldsymbol{A}^{-1}$ 로 표시하며 다음을 만족하는 행렬이다.

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

- 역행렬은 유일하다.
- 예를 들어 2차원 정방행렬의 역행렬은 다음과 같이 계산할 수 있다.

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, ad-bc \neq 0 \rightarrow \boldsymbol{A}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{4.1}$$

위의 2차원 정방행렬의 역행렬에서 만약 $ad - bc = 0$ 이면 역행렬이 존재하지 않는다. 일반적으로 모든 정방행렬의 역행렬이 존재하는 것은 아니다.

4.2 중요한 내용과 정의

- 역행렬의 성질

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

$$(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$$

- 연립방정식의 해

  n개의 n 변수 일차연립방정식 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$가 주어졌다고 하자. 여기서 $\boldsymbol{A}$는 $n \times n$ 정방행렬이다. 만약 $\boldsymbol{A}^{-1}$가 존재하면

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}$$

5 가우스소거법과 연립방정식의 해

! 중요

- 슬라이드의 1-7페이지를 반드시 먼저 학습하세요
- 부교재 Deisenroth, Faisal, 와/과 Ong (2020) 의 다음 절을 반드시 학습하세요
  - 2.3.1 Particular and General Solution
  - 2.3.2 Elementary Transformations
  - 2.3.3 The Minus-1 Trick

5.1 행렬의 기본연산

5.1.1 일차연립방정식의 기본연산

기본연산은 일차연립방정식의 해집합을 변화시키지 않으면서 방정식을 변화시켜 해집합을 구하는 다음의 세 가지 연산을 말한다.

1. 한개의 방정식의 양변에 0이 아닌 상수를 곱하기
2. 한개의 방정식의 양변에 0이 아닌 상수를 곱한 것을 다른방정식의 양변에 각각 더하기

3. 방정식의 위치를 바꾸기

5.1.2 행렬의 기본 행연산

기본 행연산(elementary row operation)은 기본연산을 행렬의 행에 시행하는 것을 말한다.

1. 한 행에 0이 아닌 상수를 곱하기
2. 한 행에 0이 아닌 상수를 곱한 결과를 다른 행 에더하기
3. 두 행의 위치를 교환하기

5.2 행렬과 벡터의 곱

$m \times n$ 인 행렬 $\boldsymbol{A}$ 와 n-차원 벡터 $\boldsymbol{x}$를 곱하는 과정을 다음과 같이 두 개의 서로 다른 형태로 나타낼 수 있다.

1. **행렬 계산법의 이용**

먼저 행렬과 벡터의 곱셈은 행렬 계산법의 이용하여 나타낼 수 있다.

$$\boldsymbol{Ax} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{l=1}^{n} a_{1l}x_l \\ \sum_{l=1}^{n} a_{2l}x_l \\ \vdots \\ \sum_{l=1}^{n} a_{ml}x_l \end{bmatrix}$$

2. **열벡터의 선형조합**

이제 행렬과 벡터의 곱셈을 행렬 $\boldsymbol{A}$을 구성하는 열벡터들의 선형조합(linear combination)으로 나타낼 수 있다.

$$\boldsymbol{Ax} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$$

위의 식에서 벡터 $\boldsymbol{a}_j$ 는 행렬 $\boldsymbol{A}$ 의 j 번째 열벡터이다.

예제 5.1. 먼저 간단한 예제로 다음과 같은 행렬과 벡터의 곱셈을 생각해 보자.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

행렬과 벡터의 곱셈은 앞에서 배운 행렬의 곱셈 방법으로 다음과 같이 나타낼 수 있다.

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} (1)(1) + (2)(-1) \\ (1)(1) + (-1)(-1) \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

이제 행렬과 벡터의 곱셈을 행렬 $\boldsymbol{A}$ 의 열들의 선형 조합으로 표시할 수 있다는 것도 알아두자.

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = (1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

행렬과 벡터의 곱셈을을 **앞 행렬의 열**과 **뒷 벡터의 원소**의 선형조합으로 나타낼 수 있다는 사실은 다양한 주제에서 유용하게 사용된다.

■

5.3 방정식의 근이 무한개인 경우

교재 슬라이드 4번의 5-7 페이지에는 변수가 3개이고 방정식의 개수가 2개인 경우에

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 6 \\ x_2 - x_3 = 2 \end{cases} \quad \leftrightarrow \quad \boldsymbol{Ax} = \boldsymbol{y} \quad \leftrightarrow \quad \begin{bmatrix} 1 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \tag{5.1}$$

아래와 같이 첨가행렬(augmented matrix)을 만들고 기본 행연산을 적용하여 일반해를 구하는 예제가 있다.

$$\left[\begin{array}{ccc|c} 1 & -2 & 2 & 6 \\ 0 & 1 & -1 & 2 \end{array}\right] \tag{5.2}$$

위의 식 5.2 에서 첨가행렬의 왼쪽 부분이 행사다리꼴 행렬(row echelon form)임을 유의하자. 식 5.2 의 두 번째 행에 2를 곱해서 첫번 째 행에 더하면 다음과 같이 기약행사다리꼴 행렬(reduced row echelon form)이 된다.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 0 & 1 & -1 & 2 \end{array}\right] \tag{5.3}$$

이제 방정식을 푸는 3가지 방법에 대하여 알아보자.

(1) 매개변수와 가우스소거법를 이용하는 법

슬라이드에 나오는 방법처럼 식 5.3 에 매개변수($x_3 = t$)를 사용하기 위하여 첨가행렬의 마지막 행에 $(0, 0, 1, t)$를 추가한 후 첨가행렬의 왼쪽을 항등행렬로 바꾸는 행연산을 적용하면(가우스 소거법) 다음과 같은 결과를 얻고

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & t+2 \\ 0 & 0 & 1 & t \end{array}\right]$$

해집합은 다음과 같이 주어진다.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ t+2 \\ t \end{bmatrix} = t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 10 \\ 2 \\ 0 \end{bmatrix}, \quad t \in \mathbb{R} \tag{5.4}$$

(2) 특수해와 선형조합을 이용

다시 식 5.3 에서 제시된 방정식을 열벡터의 선형조합의 형태로 아래와 같이 써보자

$$\boldsymbol{Ax} = \boldsymbol{y} \quad \leftrightarrow \quad x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \end{bmatrix} \tag{5.5}$$

이제 위의 식에서 $x_3 = 0$ 으로 놓으면 다음과 같이 x_1 과 x_2 만 포함된 간단한 방정식이 나타나며 이를 만족하는 특수해(particular solution) $\boldsymbol{x}^*$ 를 다음과 같이 구할 수 있다.

$$x_3 = 0,\ \ x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x}^* = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ 0 \end{bmatrix} \tag{5.6}$$

위의 식 5.6 에서 구한 특수해 $\boldsymbol{x}^*$ 는 식 5.4 에서 주어진 해집합에서 나타난 마지막 벡터이다.

이제 식 5.1 에 주어진 방정식은 특별한 해 $\boldsymbol{x}^*$ 만 만족하는 것이 아니므로 일반해(general solution)을 구해야 한다. 일반해를 구하는 방법은 다음과 같이 행렬 $\boldsymbol{A}$ 의 열들의 선형 조합이 영벡터가 되는 $\boldsymbol{x}^{**}$ 를 찾는 것이다.

$$\boldsymbol{A}\boldsymbol{x}^{**} = \boldsymbol{0} \quad \leftrightarrow \quad x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5.7}$$

식 5.7 를 만족하는 해는 유일하지 않다. 하지만 행렬 $\boldsymbol{A}$ 의 두 번째 열과 세 번째 열의 부호가 반대인 점을 이용하면 다음과 같은 식을 얻을 수 있다.

$$(0) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} + (1) \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

따라서 식 5.7 를 만족하는 해 $\boldsymbol{x}^{**}$ 는 쉽게 찾을 수 있다.

$$\boldsymbol{x}^{**} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

이제 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{y}$ 를 만족하는 특수해 $\boldsymbol{x}^*$ 와 $\boldsymbol{A}\boldsymbol{x}^{**} = \boldsymbol{0}$ 를 만족하는 $\boldsymbol{x}^{**}$ 를 이용하여 일반해를 구해보자. 임의의 실수 t 에 대하여 다음과 같은 식이 만족한다.

$$\boldsymbol{A}(\boldsymbol{x}^* + t\boldsymbol{x}^{**}) = \boldsymbol{A}\boldsymbol{x}^* + t\boldsymbol{A}\boldsymbol{x}^{**} = \boldsymbol{y} + t\boldsymbol{0} = \boldsymbol{y} \quad t \in \mathbb{R}$$

위의 식에 의하여 이제 일반해를 구하면 다음과 같이 나타나며 이는 매개변수를 이용하여 얻은 해(식 5.4)와 동일하게 나타난다.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \boldsymbol{x}^* + t\boldsymbol{x}^{**} = \begin{bmatrix} 10 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

(3) (-1)-추가법

이제 마지막 방법은 부교재 2.3.3 절(32 페이지)에 나온 (-1)-추가법(Minus-1 Trick)에 대하여 배워보자.

식 5.3 는 기약행사다리꼴 행렬로서 첫 번째 행과 두 번째 행이 피봇을 포함한 행이다. 이제 기약행사다리꼴 행렬에 대각원소 위치에 피봇이 없는 행에 대각원소가 -1 인 행을 추가하여 정방행렬로 만들어 보자. 식 5.3 에서 3행에 피봇이 없으므로 대각원소가 -1 이고 나머지가 0인 행을 3행에 다음과 같이 추가한다.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -1 & 0 \end{array}\right] \tag{5.8}$$

이제 식 5.8 에서 가장 오른 쪽에 있는 열이 특수해가 되며 대각원소가 -1 은 열벡터가 $\boldsymbol{Ax} = \mathbf{0}$ 을 만족하는 해가 된다.

따라서 방정식의 일반해를 다음과 같이 쓸수 있다.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ 0 \end{bmatrix} + t^* \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad t = -t^*, \;\; t \in \mathbb{R}$$

5.4 영공간과 일반해

식 5.7 에 나타난 것과 같이 주어진 행렬 $\boldsymbol{A}$ 의 열들의 선형조합을 영벡터로 만드는 해의 집합을 영공간(Null space) 라고 한다.

정의 5.1. 일차연립방정식(또는 행렬 방정식) $\boldsymbol{Ax} = \mathbf{0}$ 의 해집합을 $\boldsymbol{A}$의 영공간이라고 하고 $N(\boldsymbol{A})$라고 표시한다. 즉,

$$N(\boldsymbol{A}) = \{\boldsymbol{x} : \boldsymbol{Ax} = \mathbf{0}\}$$

위에서 방정식의 해를 찾을 때 특수한 해를 먼저 찾고 일반해를 만드는 작업은 영공간을 찾는 작업과 동일하다.

5.5 중요한 내용과 정의

- 행사다리꼴 행렬(row echelon form)의 정의
- 피벗(pivot 혹은 leading entry)의 정의
- 기약행사다리꼴(reduced row echelon form)의 정의
- 가우스 소거법의 절차
  - 기본변수(basic variable)와 자유변수(free variable)
  - 매개변수의 이용
- 연립일차방정식이 유일한 해를 가질 조건
- 정방행렬의 역행령이 존재하면 영공간은 영벡터와 같다.

6 행연산 행렬과 역행렬

> **!** 중요
>
> 강의자료 슬라이드 6번의 기본행렬은 강의 범위에 포함되지 않습니다.
> 단, 첨가행렬과 행연산을 이용하여 역행렬을 구하는 방법은 반드시 알아야 합니다.
> 이 연습장에 포함된 예제와 부교재 33-34 페이지 Calculating the Inverse 의 Example 2.9 를 공부하세요.

6.1 역행렬의 공식

먼저 2×2 행렬의 역행렬을 구하는 공식을 이용해 보자. 다음과 같이 2×2 행렬 $\boldsymbol{A}$ 가 주어졌을 때

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

식 4.1 를 사용하면 다음과 같이 2×2 행렬의 역행렬을 구할 수 있다.

$$\boldsymbol{A}^{-1} = \frac{1}{(1)(4) - (2)(3)} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$$

6.2 행연산과 역행렬

아래에 주어진 두 예제에서 행연산을 이용하여 역행렬을 구해보자.

예제 6.1 (2×2 행렬의 역행렬). 정방행렬의 역행렬을 구하는 다른 방법 중의 하나는 항등행렬 $\boldsymbol{I}$ 와 같이 첨가행렬을 만들고 행연산을 적용하는 것이다. 이제 행연산을 이용하여 $\boldsymbol{A}$ 의 역행렬을 구하는 방법을 연습해 보자

이제 $\boldsymbol{A}$ 과 이차원 항등행렬 $\boldsymbol{I}$ 을 붙여서 만든 첨가행렬은 다음과 같다.

$$[\,\boldsymbol{A} \mid \boldsymbol{I}\,] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$

이제 위의 첨가행렬에서 행연산을 이용하여 행렬 $\boldsymbol{A}$ 부분을 항등행렬로 만들어 보자.

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \quad (-3)(\text{1st row}) + (\text{2nd row})$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right] \quad (1)(\text{2nd row}) + (\text{1st row})$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right] \quad (-1/2)(\text{2nd row})$$

$$\rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

이렇게 첨가행렬에서 행렬 $\boldsymbol{A}$ 부분을 행연산을 이용하여 항등행렬로 만들어 주면 오른쪽의 항등행렬이 $\boldsymbol{A}^{-1}$ 로 나타난다.

■

예제 6.2 (4 × 4 행렬의 역행렬).

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

첨가행렬에서 행연산을 이용하여 행렬 $\boldsymbol{A}$ 부분을 항등행렬로 만들어 보자.

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ (-1)(\text{1st row}) + (\text{3nd row}) \\ (-1)(\text{1st row}) + (\text{4th row}) \end{array}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right] \text{swap with 2nd and 4th row}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ (-1)(\text{2nd row}) + (\text{3rd row}) \\ (-1)(\text{2nd row}) + (\text{4th row}) \end{array}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & -1 \end{array}\right] \begin{array}{l} (-1)(\text{4th row}) + (\text{1st row}) \\ \\ \\ \\ \end{array}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & -1 \end{array}\right] \begin{array}{l} \\ \\ \\ \text{swap with 3rd and 4th row} \end{array}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} \\ \\ \\ (-1)(\text{3rd row}) + (\text{4th row}) \end{array}$$

$$\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & -2 \end{array}\right]$$

위의 마지막 결과로 다음과 같은 역행렬이 얻어진다.

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

■

7 벡터공간

> **노트**
>
> 강의자료 슬라이드 2-5번(군, 필드의 정의)은 강의 범위에 포함되지 않습니다.
> 부교재 37-40 페이지를 공부하세요.

7.1 벡터공간의 정의와 의미

벡터공간(vector space) 은 어떤 집합 S 에 다음과 같은 두 개의 연산이 정의된 공간을 말한다.

1. 두 개의 원소에 대한 더하기(addition, $+$) 연산의 정의되어 있다.

$$+ \quad : S + S \to S \tag{7.1}$$

2. 하나의 실수와 한 개의 원소에 대한 스칼라곱(scalar product, $\cdot$) 연산이 정의되어 있다.

$$\cdot \quad : \mathbb{R} \cdot S \to S \tag{7.2}$$

위에서 **더하기 연산이 정의되어 있다**는 의미는 다음에 주어진 규칙이 성립한다는 의미이다.

- 집합 S 가 연산에 대하여 닫혀있다 (closure).

$$s_1 + b \in S \quad \forall s_1, b \in S$$

- 결합법칙이 성립한다 (Associativity).

$$(s_1 + s_2) + s_3 = s_1 + (s_2 + s_3) \quad \forall s_1, s_2, s_3 \in S$$

- 항등원이 존재한다 (Neutral element).

$$s + e = e + s = s \quad \exists e \ \ \forall s \in S$$

- 역원이 존재한다 (Inverse element).

$$s + i = i + s = 0 \quad \exists i \quad \forall s \in S$$

일반적으로 항등원(e) 는 0 으로 표시하며 역원(i) 는 $-s$ 로 표시한다.

- 교환법칙이 성립한다 (Commutativity).

$$s_1 + s_2 = s_2 + s_1 \quad \forall s_1, s_2 \in S$$

또한 위에서 **스칼라곱 연산이 정의되어 있다**는 의미는 다음에 주어진 규칙이 성립한다는 의미이다.

- 스칼라곱 연산의 분배법칙이 성립한다 (Distributivity).

$$r_1(s_1 + s_2) = r_1 s_1 + r_2 s_2, \quad (r_1 + r_2)s = r_1 s + r_2 s \quad \forall s_1, s_2 \in S, \quad \forall r_1, r_2 in \mathbb{R}$$

- 스칼라곱 연산의 결합법칙이 성립한다

$$r_1(r_2 s) = (r_1 r_2)s \quad \forall s \in S, \quad \forall r_1, r_2 in \mathbb{R}$$

- 스칼라곱 연산의 항등원이 존재한다 (Neutral element).

$$1 \cdot s = s \quad \forall s \in S$$

일반적으로 벡터공간은 $(S, +, f)$ 라고 표시한다. 이러한 표시에서 함수 f 는 스칼라곱 연산에 대한 정의를 나타내는 것이며 식 7.2 에 나타나는 대응을 의미한다.

이 강좌에서는 스칼라로 실수만 사용하고 있으므로 벡터공간을 실벡터(real vector space) 라고 부른다.

$$f : \mathbb{R} \cdot S \to S, \quad \text{즉} \quad f(rs) = r \cdot s = rs$$

주의

벡터 공간에서 주의할 점은 **두 벡터의 곱하기** 가 정의되어 있다는 것이 아니라 하나의 스칼라와 하나의 벡터에 대한 스칼라 곱하기가 정의되어 있다는 것이다.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = ? \quad but \quad 3 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

두 벡터의 곱하기 는 나중에 내적(inner product) 란 이름으로 따로 정의한다.

7.2 중요한 내용과 정의

- 벡터공간의 예제 (슬라이드 참조)
- 부분공간(subspace)의 정의와 예제(부교재 39 페이지 Example 2.12 참조)

8 벡터공간의 기저와 차원

노트

강의자료 슬라이드 2-5 페이지(군, 필드의 정의)은 강의 범위에 포함되지 않습니다.
이 연습장에 포함된 예제와 부교재 40-47 페이지를 공부하세요.

8.1 벡터의 일차독립

벡터공간에 속한 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 의 일차결합(또는 선형결합, linear combination)이란 각 벡터에 스칼라를 곱하여 더한 것들이다. 즉 다음과 같은 형태의 식을 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 의 일차결합(linear combination)이라고 한다:

$$r_1\boldsymbol{v}_1 + r_2\boldsymbol{v}_2 + \cdots + r_n\boldsymbol{v}_n, \quad r_1, r_2, \dots, r_n \in \mathbb{R} \tag{8.1}$$

정의 8.1 (벡터의 일차독립과 일차종속). 벡터공간에 속한 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 가 있다고 하자. 만약 다음 식이 만약 모두 0인 n개의 스칼라 $x_1, x_2, \dots, x_n$에 대해서만 성립하면 n개 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 들은 일차독립(linearly independent)라고 한다.

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n = \mathbf{0} \quad \Longleftrightarrow x_1 = x_2 = \cdots = x_n = 0 \tag{8.2}$$

또한 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 가 일차독립이 아니면 일차종속(linear dependent)라고 한다. 벡터 $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ \dots\ ,\boldsymbol{v}_n$ 가 일차종속이면 **모두 0이 아닌** $x_1, x_2, \dots, x_n$ 이 존재하여 다음이 성립한다는 것이다.

$$\exists\ x_1, x_2, \dots, x_n \in \mathbb{R} \text{ s.t. } (x_1, x_2, \dots, x_n) \neq \mathbf{0}, \quad \boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n = \mathbf{0} \tag{8.3}$$

■

예를 들어 다음과 같이 주어진 3개의 3-차원 벡터들은 선형종속이다.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \tag{8.4}$$

왜냐하면 다음과 같이 모두 0이 아닌 스칼라에 의해서 다음 식이 성립하기 때문이다. 즉 벡터 $\boldsymbol{v}_3$는 $\boldsymbol{v}_2$ 에 2를 곱하여 $\boldsymbol{v}_1$ 에 더한 값과 같다.

$$\boldsymbol{v}_3 = \boldsymbol{v}_1 + 2\boldsymbol{v}_2 \quad \Longleftrightarrow \quad \boldsymbol{v}_1 + 2\boldsymbol{v}_2 - \boldsymbol{v}_3 = 0$$

주어진 벡터들이 서로 일차독립임을 확인할 수 있는 일반적인 방법은 다음과 같이 가우스소거법을 이용하는 것이다.

1. 주어진 벡터들을 열로 구성하는 행렬을 만들고 가우스소거법(또는 행사다리꼴)을 적용한다.
2. 이때 피봇을 포함하는 열의 개수가 선형독립인 벡터의 개수이다.

다음과 같이 식 8.4 의 3개의 벡터를 각 열로 합친 3×3-차원 행렬에 행연산을 적용하여 피봇이 1인 행사다리꼴을 만들어보자.

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 0 & 2 \\ 3 & 1 & 5 \end{bmatrix} \begin{array}{l} \\ (-2)(\text{1st row}) + (\text{2nd row}) \\ (-3)(\text{1st row}) + (\text{3rd row}) \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & -2 & -4 \\ 0 & -2 & -4 \end{bmatrix} \begin{array}{l} \\ \\ (-1)(\text{2nd row}) + (\text{3rd row}) \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \\ (-1/2)(\text{2nd row}) \\ \\ \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} (-1)(\text{2nd row}) + (\text{1st row}) \\ \\ \\ \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

위에서 마지막 행렬의 비봇(빨간 숫자 1)을 포함한 열은 첫 번째 열과 두 번째 열이고 세 번째 열은 첫 번째 열과 두 번째 열의 선형조합으로 나타낼 수 있음을 보여주고 있다. 피봇을 포함하지 않는 세번 째 열의 숫자가 각각 1 과 2 라는 것은 세 번째 벡터 $\boldsymbol{v}_3 = (1)\boldsymbol{v}_1 + (2)\boldsymbol{v}_2$ 로 나타날 수 있다는 것을 보여준다.

이제 다음과 같이 주어진 3개의 3-차원 벡터들은 일차독립이다. 즉 3개 벡터의 선형 조합이 0이 될 수 있도록 만드는 스칼라는 모두 0인 경우 밖에 없다.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} \tag{8.5}$$

이제 식 8.5 의 3개의 벡터를 각 열로 합친 3×3-차원 행렬에 행연산을 적용하여 피봇이 1인 행사다리꼴을 만들어 보자.

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 0 & 2 \\ 3 & 1 & 4 \end{bmatrix} \begin{matrix} \\ (-2)(\text{1st row}) + (\text{2nd row}) \\ (-3)(\text{1st row}) + (\text{3rd row}) \end{matrix}$$

$$\to \begin{bmatrix} 1 & 1 & 3 \\ 0 & -2 & -4 \\ 0 & -2 & -5 \end{bmatrix} \begin{matrix} \\ \\ (-1)(\text{2nd row}) + (\text{3rd row}) \end{matrix}$$

$$\to \begin{bmatrix} 1 & 1 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & -1 \end{bmatrix} \begin{matrix} \\ (-1/2)(\text{2nd row}) \\ (-1)(\text{3rd row}) \end{matrix}$$

$$\to \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} (-1)(\text{2nd row}) + (\text{1st row}) \\ (-2)(\text{3rd row}) + (\text{2nd row}) \\ \end{matrix}$$

$$\to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} (-1)(\text{3rd row}) + (\text{1st row}) \\ \\ \end{matrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

식 8.5 의 3개의 벡터로 구성된 행렬에 가우스 소거법을 적용하면 위와 같이 모든 열이 피봇을 포함한 열로 나타난다. 따라서 3개의 벡터는 서로 일차독립이다.

이제 다음과 같이 주어진 4개의 3-차원 벡터들은 일차종속이다.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} \quad \boldsymbol{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{8.6}$$

식 8.6 에 나타난 4개의 벡터가 일차종속임을 어떻게 알 수 있을까? 앞에서와 마찬가지로 식 8.6 에 있는 4개의 벡터들이 열로 구성된 3×4-행렬에 가우스소거법을 적용해보자.

$$
\begin{aligned}
&\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{v}_4 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 3 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 1 & 4 & 1 \end{bmatrix} \begin{matrix} \\ (-2)(\text{1st row}) + (\text{2nd row}) \\ (-3)(\text{1st row}) + (\text{3rd row}) \end{matrix} \\
&\to \begin{bmatrix} 1 & 1 & 3 & 0 \\ 0 & -2 & -4 & 0 \\ 0 & -2 & -5 & 1 \end{bmatrix} \begin{matrix} \\ \\ (-1)(\text{2nd row}) + (\text{3rd row}) \end{matrix} \\
&\to \begin{bmatrix} 1 & 1 & 3 & 0 \\ 0 & -2 & -4 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{matrix} (1/2)(\text{2nd row}) + (\text{1st row}) \\ \\ \\ \end{matrix} \\
&\to \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -2 & -4 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{matrix} \\ (-1/2)(\text{2nd row}) \\ \\ \end{matrix} \\
&\to \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}
\end{aligned}
$$

위와 같이 가우스소거법을 적용하여 얻은 행렬에서 피봇을 포함한 열은 1,2,4, 번째 열이고 포함하지 않은 열은 3번째 열임을 알 수 있다. 여기서 주어진 벡터들로 행렬을 구성할 때 기약행사다리꼴의 형태가 벡터들을 배열하는 순서에 따라 달라지는 것을 알 수 있다. 즉, 3×4-행렬을 구성할 때 순서를 바꾸어 다음과 같이 $[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4, \boldsymbol{v}_3]$ 로 배열하면 다음과 같은 기약행사다리꼴의 형태가 얻어진다.

$$
\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_4 & \boldsymbol{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 0 & 0 & 2 \\ 3 & 1 & 1 & 4 \end{bmatrix} \to_{\text{가우스소거법}} \to \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{bmatrix}
$$

위와 같이 피봇이 1인 기약행사다리꼴에서 식 8.6 의 벡터 $\boldsymbol{v}_3$가 나머지 3개의 벡터의 선형조합으로 표현될 수 있다는 의미이다. 따라서 식 8.6 의 벡터는 일차종속이며 기약행사다리꼴의 마지막 열에 나타나 숫자 $(1, 2, -1)$은 $\boldsymbol{v}_3$가 다음과 같이 다른 벡터의 일차결합으로 나타난는 것을 보여준다.

$$
\boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = (1)\boldsymbol{v}_1 + (2)\boldsymbol{v}_2 + (-1)\boldsymbol{v}_4 = (1) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + (2) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
$$

식 8.6 와 같이 3차원 벡터가 4개인 경우 벡터의 값에 관계없이 일차종속으로 나타난다. 이러한 사실은 $\mathbb{R}^n$ 의 $n+1$ 개의 벡터는 항상 일차종속이라는 정리(슬라이드 6페이지의 정리 참조)의 결과이다. 즉, $\mathbb{R}^n$ 에서 n개보다 더 많은 벡터들은 항상 일차종속이다.

8.2 생성집합과 기저

정의 8.2 (생성집합과 기저). 벡터공간 V 의 벡터 $\boldsymbol{v}_1, \boldsymbol{v}_n, \dots, \boldsymbol{v}_m$ 의 일차결합을 모두 모은 집합

$$W = span\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m\} = \{r_1\boldsymbol{v}_1 + r_2\boldsymbol{v}_2 + \dots + r_m\boldsymbol{v}_m : r_1, r_2, \dots, r_m \in \mathbb{R}\}$$

을 벡터 $\boldsymbol{v}_1, \boldsymbol{v}_n, \dots, \boldsymbol{v}_m$ 의 생성(span)이라고 하며 W 의 생성집합(generating set, spanning set) 이라고 한다.

또한 어떤 벡터공간(혹은 부분공간)의 **생성집합에 속한 벡터들이 일차독립일 때** 이 생성집합을 기저 (basis)라고 한다

■

8.3 중요한 내용과 정의

- $\mathbb{R}^n$ 의 모든 기저는 n개의 원소를 갖는다.
- 임의의 벡터공간 V에 대해서 V의 부분집합 $B = \{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ 가 V의 한 기저라고 하면 다음을 보일 수 있다.
  - V 의 모든 벡터들은 $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ 의 일차결합으로 나타낼 수 있으며 유일하다.
  - V 의 부분집합이 n 개보다 많은 벡터를 포함하면 이 부분집합의 벡터들은 일차종속이다.
  - V 의또다른기저 $C = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_m\}$ 가있다면$m = n$ 이다.
- 벡터공간 V의 차원(dimension) 은 기저의 개수로 정의되며 $dim(V)$로 표시한다.

9 행렬의 계수

9.1 계수의 정의

정의 9.1 (계수의 정의). 행렬의 계수(rank)란 **행렬의 일차 독립인 행들의 최대 수** 또는 **일차독립인 열들의 최대 수**로 정의된다

$$rank(\boldsymbol{A}) = rk(\boldsymbol{A}) = dim(Col(\boldsymbol{A})) = dim(Row(\boldsymbol{A}))$$

■

꼭 기억해야 할 것은 행렬의 계수는 열들을 이용하여 구한 계수와 행들을 이용하여 구한 계수가 같다는 것이다. 즉, 행렬의 계수는 열의 계수와 행의 계수 중 하나만 구해도 된다는 것이다.

9.2 중요공식

- $rank(\boldsymbol{A}) = rank(\boldsymbol{A}^T)$
- 행렬 $\boldsymbol{A} \in \mathbb{R}^n$ 가 정방행렬이고 계수가 n 이면 역행렬이 존재한다. 약행렬이 존재하면 정칙행렬(non-sigular matrix)이라고 한다.
- 더 나아가 다음에 나오는 문장은 모두 동치(equaivalent)이다.

  모든 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 에 대하여

  $\boldsymbol{A}$ 가 정칙행렬이다.

  $\Leftrightarrow \boldsymbol{A}$ 의 열들이 일차독립이다.

  $\Leftrightarrow \boldsymbol{A}$ 의 행들이 일차독립이다.

  $\Leftrightarrow \boldsymbol{A}$ 의 계수가 n 이다.

  $\Leftrightarrow \boldsymbol{A}$ 의 기약행사다리꼴행렬이 항등행렬이다.

  $\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{0}$ 해는 영벡터가 유일하다.
- 최대계수행렬
  - $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 에 대하여 $m \leq n$ 이고 $rank(\boldsymbol{A}) = m$ 이면 $\boldsymbol{A}$ 는 최대 행계수를 갖는다. 이 때 $\boldsymbol{A}$를 최대 행계수 행렬(full row rank matrix)이라고한다.

– $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 에 대하여 $m \leq n$ 이고 $m \geq n$이고 $rank(\boldsymbol{A}) = n$ 이면 $\boldsymbol{A}$는 최대 열계수를 갖는다 A는 최대 열계수 행렬이다라고 한다.이 때 $\boldsymbol{A}$를 최대 열계수행렬((full column rank matrix)이라고 한다.

– $\boldsymbol{A}$ 가 최대 행계수 행렬 또는 최대 열계수 행렬인 경우 $\boldsymbol{A}$ 는 최대 계수를 갖는다 또는 $\boldsymbol{A}$ 는 최대 계수 행렬(full rank matrix) 이라고 한다.

9.3 예제

9.3.1 부교재

- Example 2.18 (Rank)

9.3.2 연습문제 1

다음과 같은 3×4 행렬에 행연산을 적용하여 행사다리꼴 행렬로 만들어 보자.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 & 5 \\ 2 & 5 & 1 & 14 \\ 4 & 9 & 3 & 24 \end{bmatrix} \tag{9.1}$$

$$\begin{bmatrix} 1 & 2 & 1 & 5 \\ 2 & 5 & 1 & 14 \\ 4 & 9 & 3 & 24 \end{bmatrix} \quad (-2)(\text{1st row}) + (\text{2nd row})$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 1 & -1 & 4 \\ 4 & 9 & 3 & 24 \end{bmatrix} \quad (-4)(\text{1st row}) + (\text{3rd row})$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 1 & -1 & 4 \\ 0 & 1 & -1 & 4 \end{bmatrix} \quad (-1)(\text{2nd row}) + (\text{3rd row})$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

위의 결과에서 피봇 행이 2개가 나타나며 나머지 행은 모두 0으로 나타난다. 이러한 결과를 행렬의 계수가 2인 것을 의미하며($rank(\boldsymbol{A}) = 2$). 계수의 정의에 의하여 선형독립인 행의 갯수가 2이다.

따라서 마지막 행은 다른 2개의 행들의 선형조합이고 행연산의 결과로 모든 원소가 0이 되는 것을 알 수 있다.

만약 식 9.1 에 나타난 행렬의 열들을 고려하면 선형독립인 열들이 2개가 될까? 식 9.1 의 행렬 $\boldsymbol{A}$ 의 전치 행렬은 열들이 행으로 바뀐 행렬이므로 위와 유사하게 행 연산을 적용하면 선향독립인 열의 개수를 구할 수 있다.

$$\boldsymbol{A}^T = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 9 \\ 1 & 1 & 3 \\ 5 & 14 & 24 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 9 \\ 1 & 1 & 3 \\ 5 & 14 & 24 \end{bmatrix} \quad \text{(change 1st row and 3rd row)}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 2 & 5 & 9 \\ 1 & 2 & 4 \\ 5 & 14 & 24 \end{bmatrix} \quad \begin{array}{l} (-2)(\text{1st row}) + (\text{2nd row}) \\ (-1)(\text{1st row}) + (\text{3rd row}) \\ (-5)(\text{1st row}) + (\text{4th row}) \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & 3 & 3 \\ 0 & 1 & 1 \\ 0 & 9 & 9 \end{bmatrix} \quad \text{(change 2nd row and 3rd row)}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 3 & 3 \\ 0 & 9 & 9 \end{bmatrix} \quad \begin{array}{l} (-1)(\text{2nd row}) + (\text{3rd row}) \\ (-3)(\text{2nd row}) + (\text{4th row}) \end{array}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

위와 같이 4개의 열로 보아도 서로 독립인 열의 개수는 2개임을 알 수 있다.

따라서 행렬의 계수를 구하는 경우는 행을 이용한 연산과 열을 이용한 연산 중 하나만 선택하여 계산하면 된다.

10 선형사상

10.1 선형사상

V 와 W 가 벡터공간일 때 함수 $T : V \to W$ 가 다음 두 가지 조건을 만족하면 선형사상(linear mapping, linear transformation, homomorphism)이라고 한다.

$$
\begin{aligned}
&(1) \quad \forall v_1, v_2 \in V, \quad T(v_1 + v_2) = T(v_1) + T(v_2) \\
&(2) \quad \forall r \in \mathbb{R}, \quad T(rv_1) = rT(v_1)
\end{aligned}
$$

- 선형사상의 종류와 성질

다음에 나오는 특별한 성질을 가진 선형사상과 특별한 함수의 정의를 반드시 학습하고 암기하세요

- Injective(단사함수, one-to-one)
- Surjective(전사함수, onto)
- Bijective(전단사함수, one-to-one correspondence)
- 항등함수(id_V)
- 역함수

다음과 같은 용어는 참고로 알아두자.

- Isomorphism (동형사상): $T : V \to W$ linear and bijective
- Automorphism (자기동형사상): $T : V \to V$ linear and bijective

정리 10.1 (동형사상). 한 벡터공간 V 가 있어서 $dim(V) = n$ 이면 동형사상 $f : V \to \mathbb{R}^n$ 이있다.

■

11 선형사상의 변환행렬

ℹ 노트

아래 열거된 사이트는 선형 변환을 시각적으로 설명하는 도구를 제공한다.

- 그림의 변환
- 2차원 좌표의 변환
- 3차원 좌표의 변환

11.1 좌표

벡터공간 V 의 차원이 $dim(V) = n$ 이고 주어진 기저를 $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ 이라고 하자. 참고로 기저 B 는 순서가 있는 집합(oerdered set)이다.

벡터공간 V 의 원소 $\boldsymbol{x} \in V$ 는 다음과 같이 기저 B 의 선형조합으로 나타난다,

$$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n \tag{11.1}$$

참고로 $\boldsymbol{x}$를 식 11.1 과 같이 나타낼 수 있는 $\alpha_1, \alpha_2, \dots, \alpha_n$ 의 조합은 유일하다.

정의 11.1 (좌표(coordinator)). 벡터공간 V 의 원소 $\boldsymbol{x} \in V$ 가 기저 B 에 대하여 식 11.1 으로 표현된다면 $\alpha_1, \alpha_2, \dots, \alpha_n$ 를 $\boldsymbol{x}$ 의 좌표(coordinator) 라고 부른다.

또한 벡터공간 V 의 모든 원소에 대하여 좌표를 대응해주는 좌표사상(coordinator map) 라고 부르고 다음과 같이 함수 $f : V \to \mathbb{R}^n$ 로 나타낼 수 있다.

$$f(\boldsymbol{x}) = \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

■

11.1.1 예제 (부교재 Example 2.20)

이차원 벡터공간 $V = \mathbb{R}^2$ 에서 다음과 같은 하나의 벡터를 생각하자.

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

이제 다음과 같은 표준 기저(standard basis) $B = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ 를 생각하자.

$$B = (\boldsymbol{e}_1, \boldsymbol{e}_2) = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \boldsymbol{I}$$

이제 $\boldsymbol{x}$ 를 기저 B 에 대하여 식 11.1 으로 표현하면 다음과 같다.

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2\boldsymbol{e}_1 + 3\boldsymbol{e}_2 = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \boldsymbol{B}\boldsymbol{\alpha}_B \tag{11.2}$$

따라서 기저 B 에 대하여 $\boldsymbol{x}$ 의 좌표 $\boldsymbol{\alpha}$는 다음과 같다.

$$\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

이제 기저를 다음과 같이 $\tilde{B}$ 로 바꾸어 보자.

$$\tilde{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2) = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

이제 $\boldsymbol{x}$ 를 기저 $\tilde{B}$ 에 대하여 식 11.1 으로 표현하면 다음과 같다.

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = (-1/2)\boldsymbol{b}_1 + (5/2)\boldsymbol{b}_2 = (-\frac{1}{2}) \begin{bmatrix} 1 \\ -1 \end{bmatrix} + (\frac{5}{2}) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1/2 \\ 5/2 \end{bmatrix} = \tilde{\boldsymbol{B}}\boldsymbol{\alpha}_{\tilde{B}} \tag{11.3}$$

따라서 기저 $\tilde{B}$ 에 대하여 $\boldsymbol{x}$ 의 좌표 $\boldsymbol{\alpha}$는 다음과 같다.

$$\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}_{\tilde{B}} = \begin{bmatrix} -\frac{1}{2} \\ 5 \end{bmatrix}$$

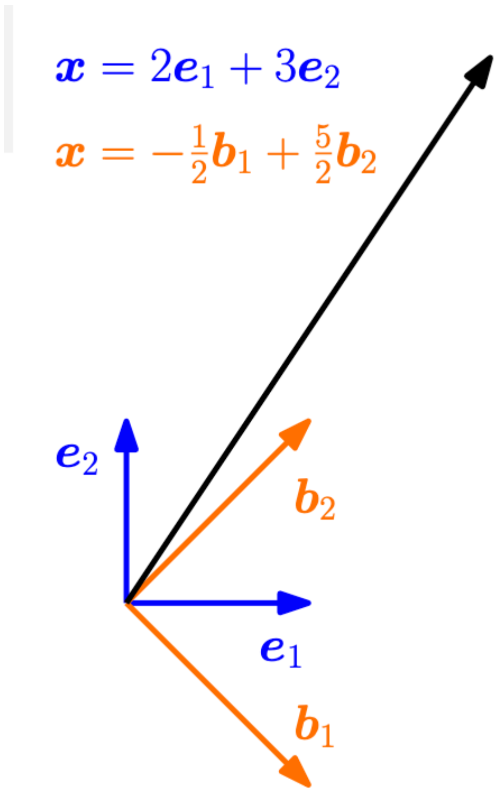


그림 11.1: 기저의 변화에 대한 좌표의 변화

이 예제에서 본 것처럼 기저가 변하면 동일한 원소에 대해서도 좌표가 달라진다.

만약 기저가 변했다면 좌표는 어떻게 변하는지를 알아야 한다. 이차원 공간의 임의의 벡터 $\boldsymbol{x}$ 에 대하여 두 기저 B 와 $\tilde{B}$ 에 대하여 식 11.2 과 식 11.3 의 관계를 이용하면 다음과 같은 식을 얻을 수 있다.

$$
\begin{aligned}
& \tilde{\boldsymbol{B}}\boldsymbol{\alpha}_{\tilde{B}} = \boldsymbol{B}\boldsymbol{\alpha}_B \\
\Rightarrow \quad & \boldsymbol{\alpha}_{\tilde{B}} = \tilde{\boldsymbol{B}}^{-1}\boldsymbol{B}\boldsymbol{\alpha}_B \\
\Rightarrow \quad & \boldsymbol{\alpha}_{\tilde{B}} = \tilde{\boldsymbol{B}}^{-1}\boldsymbol{I}\boldsymbol{\alpha}_B \\
\Rightarrow \quad & \boldsymbol{\alpha}_{\tilde{B}} = \tilde{\boldsymbol{B}}^{-1}\boldsymbol{\alpha}_B \\
\Rightarrow \quad & \boldsymbol{\alpha}_{\tilde{B}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \boldsymbol{\alpha}_B \\
\Rightarrow \quad & \boldsymbol{\alpha}_{\tilde{B}} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \boldsymbol{\alpha}_B
\end{aligned}
$$

따라서 기저가 변하면 좌표는 위와 같은 식으로 변한다.

11.2 변환행렬

11.2.1 정의

이제 두 개의 벡터공간 V 와 W 에 대하여 선형사상 Φ 가 정의되어 있고

$$\Phi : V \to W$$

벡터공간 V 와 W 에 대한 기저가 각각 $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ 와 $C = (\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m)$ 이라고 하자.

이제 벡터공간 V 의 기저에 대한 선향사상의 원소가 벡터공간 W 의 기저로 다음과 같이 표현된다고 하자.

$$\begin{aligned}
\Phi(\boldsymbol{b}_1) &= a_{11}\boldsymbol{c}_1 + a_{21}\boldsymbol{c}_2 + \cdots + a_{m1}\boldsymbol{c}_m \\
\Phi(\boldsymbol{b}_2) &= a_{12}\boldsymbol{c}_1 + a_{22}\boldsymbol{c}_2 + \cdots + a_{m2}\boldsymbol{c}_m \\
\cdots \\
\Phi(\boldsymbol{b}_n) &= a_{1n}\boldsymbol{c}_1 + a_{2n}\boldsymbol{c}_2 + \cdots + a_{mn}\boldsymbol{c}_m
\end{aligned} \tag{11.4}$$

식 11.4 에서 나타난 계수 a_{ij} 을 $m \times n$-행렬 $\boldsymbol{A}_\Phi$ 로 다음과 같이 표현할 수 있다.

$$\boldsymbol{A}_\Phi = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{11.5}$$

식 11.4 에 나타난 행렬 $\boldsymbol{A}_\Phi$ 를 **변환행렬(transformation matrix)**이라고 부르며 이 변환행렬은 각 벡터공간의 기저 B 와 C 에 따라 정의되는 것에 유의하자.

11.2.2 좌표와 변환행렬

만약 $\hat{\boldsymbol{x}}$ 가 벡터공간 V 에서 기저 B 에 대한 원소 $\boldsymbol{x}$ 의 좌표이고

$$\boldsymbol{x} = \boldsymbol{B}\hat{\boldsymbol{x}}$$

$\hat{\boldsymbol{y}}$ 가 벡터공간 W 에서 기저 C 에 대한 선형사상 $\boldsymbol{y} = \Phi(\boldsymbol{x})$ 의 좌표이면

$$\boldsymbol{y} = \boldsymbol{C}\hat{\boldsymbol{y}}$$

두 좌표 $\hat{\boldsymbol{x}}$ 와 $\hat{\boldsymbol{y}}$ 사이의 관계는 다음과 같다.

$$\hat{\boldsymbol{y}} = \boldsymbol{A}_\Phi \hat{\boldsymbol{x}} \tag{11.6}$$

참고로 $\mathbb{R}^n$ 에서 $\mathbb{R}^m$ 으로의 선형사상 $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ 를 생각하고 각 공간의 기저를 표준 기저(standard basis)로 고려하면 변환행렬 $\boldsymbol{A}_\Phi$ 는 선형사상 $\boldsymbol{y} = \Phi(\boldsymbol{x})$ 을 나타내는 $m \times n$ 행렬이다.

$$\boldsymbol{y} = \boldsymbol{A}_\Phi \boldsymbol{x} \tag{11.7}$$

11.2.3 예제

- 부교재 Example 2.21 (Transformation Matrix) 참조
- 부교재 Example 2.22 (Linear Transformations of Vectors) 참조

12 기저변환과 변환행렬

> **노트**
>
> 강의 슬라이드 12번 기저변환과 변환행렬의 주제는 추후에 강의합니다.

13 선형변환의 핵과 상

13.1 핵과 상의 정의

정의 13.1 (선형변환의 핵과 상). 벡터공간 V, W 사이의 선형사상 $T : V \to W$를 고려하자. 선형사상 T 의핵 (kernel) $Ker(T)$ 또는 영공간(null space) $N(T)$ 는 다음과 같이 정의된다:

$$ker(T) = N(T) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$$

또한 선형사상 T 의 상(range) 또는 치역(image) $Im(T)$ 는 다음과 같이 정의된다:

$$Im(T) = T(V) = \{T(\boldsymbol{v}) \mid \in V\}$$

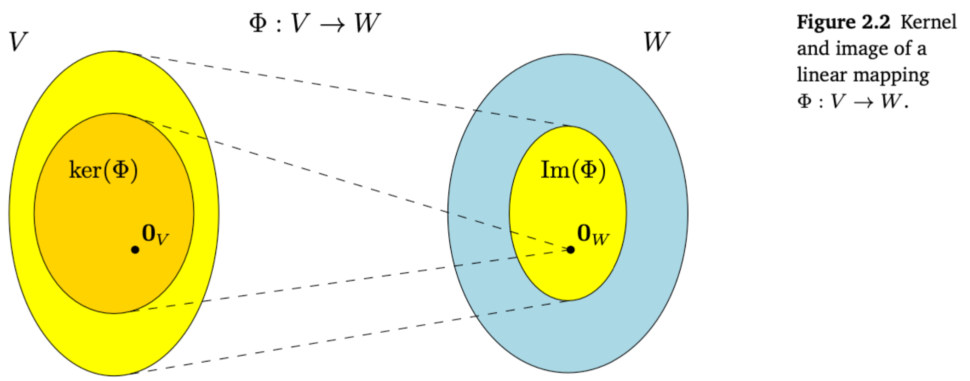


그림 13.1: 선형변환의 핵과 상

13.1.1 예제

- 부교재 Example 2.25 (Image and Kernel of a Linear Mapping)

14 아핀공간

> 노트
>
> 강의 슬라이드 14번 아핀공간의 주제는 추후에 강의합니다.

15 유클리드공간 위에서의 내적

노트

- 15번 슬라이드에서 코쉬-슈바르츠 부등식은 강의 범위가 아닙니다.
- 15번 슬라이드에서 삼각부등식(triangle inequality) 은 내용만 이해하고 증명하지 않아도 됩니다.

15.1 중요한 내용과 정의

- 유클리드 공간에서 두 벡터의 내적(dot product, inner product)
- 내적의 성질

16 벡터공간 위에서의 내적

> **노트**
>
> - 16번 슬라이드에서 대각합을 이용한 행렬 내적의 예(슬라이드 2,3,4 페이지)는 강의 범위가 아닙니다.
> - 16번 슬라이드에서 유클리드공간에서 양정치행렬을 이용한 내적(슬라이드 6페이지)에 대한 정리(Theorem) 은 강의 범위가 아닙니다.

16.1 중요한 내용과 정의

- 벡터공간에서 내적의 정의
  - 슬라이드 15번에서 유클리드 공간에서 벡터의 내적 성질은 사실 벡터공간에서 내적의 정의입니다.
- 노름(Norm)의 정의와 성질
- 거리의 정의와 성질

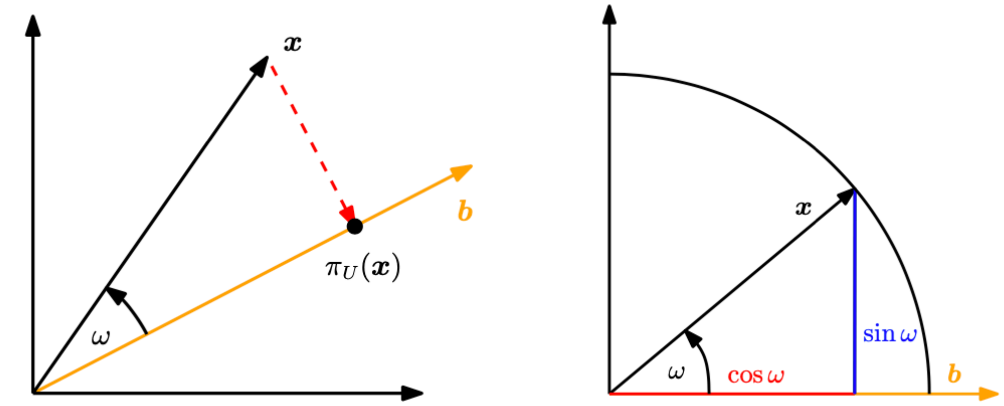
17 직교기저

17.1 중요한 내용과 정의

- 직교의 정의
- 직교 기저의 정의와 성질
- 직교여공간의 정의

18 직선으로의 정사영

18.1 직선으로의 사영



(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.

(b) Projection of a two-dimensional vector x with $\|x\| = 1$ onto a one-dimensional subspace spanned by b.

Figure 3.2 Examples of projections onto one-dimensional subspaces.

그림 18.1: 직선으로의 사영

1. 벡터 $\boldsymbol{b}$ 의 방향과 같은 벡터들 중에 벡터 $\boldsymbol{x}$ 와 가장 가까운 벡터를 $\pi_U(\boldsymbol{x})$ 라고 하자. 이 벡터는 $\boldsymbol{x}$ 에서 직선 $\boldsymbol{b}$ 에 내린 사영(projection)이며 다음을 만족해야 한다.

$$\langle \boldsymbol{x} - \pi_U(\boldsymbol{x}), \boldsymbol{b} \rangle = 0 \tag{18.1}$$

2. 사영 $\pi_U(\boldsymbol{x})$ 는 벡터 $\boldsymbol{b}$ 의 방향이므로 어떤 스칼라 λ 가 존재하여 다음을 만족해야 한다.

$$\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b}$$

3. 식 18.1 에서 제시된 직교하는 성질을 다시 쓰면 조건과 같다.

$$\langle \boldsymbol{x} - \lambda \boldsymbol{b}, \boldsymbol{b} \rangle = 0.$$

4. 내적의 성질을 이용하면 다음을 얻을 수 있다.

$$\langle \boldsymbol{x}, \boldsymbol{b} \rangle - \lambda \langle \boldsymbol{b}, \boldsymbol{b} \rangle = 0 \iff \lambda = \frac{\langle \boldsymbol{x}, \boldsymbol{b} \rangle}{\langle \boldsymbol{b}, \boldsymbol{b} \rangle} = \frac{\langle \boldsymbol{b}, \boldsymbol{x} \rangle}{\|\boldsymbol{b}\|^2}.$$

따라서 스칼라 λ 는 다음과 같다.

$$\lambda = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\boldsymbol{b}^\top \boldsymbol{b}} = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2}$$

5. 이제 벡터 $\boldsymbol{b}$ 의 방향으로의 벡터 $\boldsymbol{x}$ 의 사영 $\pi_U(\boldsymbol{x})$ 는 다음과 같다.

$$\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b} = \frac{\langle \boldsymbol{x}, \boldsymbol{b} \rangle}{\|\boldsymbol{b}\|^2} \boldsymbol{b} = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \boldsymbol{b} \tag{18.2}$$

6. 식 18.2 에서 $\boldsymbol{b}^\top \boldsymbol{x}$ 는 스칼라이므로 다음과 같이 쓸 수 있으며 노름(norm)의 정의와 행렬의 결합법칙을 이용하면 다음과 같다.

$$\begin{aligned}
\pi_U(\boldsymbol{x}) &= \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \boldsymbol{b} \\
&= \boldsymbol{b} \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \quad \text{(스칼라 성질을 이용)} \\
&= \frac{\boldsymbol{b}(\boldsymbol{b}^\top \boldsymbol{x})}{\boldsymbol{b}^\top \boldsymbol{b}} \\
&= \frac{(\boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{x}}{\boldsymbol{b}^\top \boldsymbol{b}} \quad \text{(결합법칙을 이용)} \\
&= \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\boldsymbol{b}^\top \boldsymbol{b}} \boldsymbol{x} \\
&= \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2} \boldsymbol{x} \\
&= \boldsymbol{P}_\pi \boldsymbol{x}
\end{aligned}$$

7. 벡터 $\boldsymbol{b}$ 의 방향으로 사영행렬 $\boldsymbol{P}_\pi$ 는 다음과 같다.

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2} \tag{18.3}$$

식 18.3 에서 $\boldsymbol{b}\boldsymbol{b}^\top$ 는 정방행렬이고 $\|\boldsymbol{b}\|^2$ 는 스칼라임에 유의하자.

! 사영행렬의 성질

이미 사영된 $\pi_U(x)$ 에 다시 사영행렬 P_π 을 곱해도 아무런 변화가 없다. 이는 벡터 $\boldsymbol{x}$ 를 이미 벡터 $\boldsymbol{b}$ 의 방향으로 사영했기 때문에 다시 사영해도 변화가 없다는 것을 의미한다.
즉, $P_\pi \pi_U(x) = \pi_U(x)$ 이다. 사영행렬 P_π 가 모든 벡터 x 에 대해 $P_\pi^2 x = P_\pi x$ 를 만족한다는 것을 의미한다.

$$P_\pi^2 = P_\pi$$

18.2 중요한 내용과 정의

- 슬라이드 18번의 5페이지에 나온 사영행렬에 대한 예제
- 부교재 Example 3.10

19 행렬식과 대각합

노트

이 노트의 내용은 부교재 100-105쪽의 내용을 요약한 것이다.

19.1 용어

- determinant : 행렬식
- trace : 대각합

주의

행렬식과 대각합은 정방행렬(square matrix)에 대해서만 정의된다.

19.2 행렬식

##행렬식의 정의

정방행렬(square matrix) $\boldsymbol{A}$ 의 행렬식(determinant)는 $\det(\boldsymbol{A})$ 로 표기한다.

$A \in \mathbb{R}^{2\times 2}$ 의 행렬식은 다음과 같이 계산한다.

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

19.2.1 역행렬과 Rank

- Theorm 4.1. 과 Theorem 4.3

행렬 $\boldsymbol{A}$ 가 \$n ×n \$ 정방행렬(square matrix) 인 경우 다음 3개의 문장이 동치(equivalent)임을 보여준다.

- $\det(A) \neq 0$
- $rank(A) = n$
- A은 역행렬이 존재한다

19.2.2 삼각행렬의 행렬식

행렬 $\boldsymbol{T}$ 가 상삼각행렬(upper triangular matrix) 또는 하삼각행렬(lower triangular matrix)이면 행렬식은 대각원소(diagonal element)의 곱과 같다.

$$\det(T) = \prod_{i=1}^{n} T_{ii}$$

19.2.3 Laplace Expansion

- 행렬식을 계산하는 방법 중 하나는 Laplace Expansion 이며 Theorem 4.2. 에서 설명한다.
- Example 4.3 꼭 읽어보기

19.2.4 행렬식의 성질

$$\det(AB) = \det(A)\det(B)$$
$$\det(A^T) = \det(A)$$

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

$$\det(\lambda A) = \lambda^n \det(A)$$

19.3 대각합

대각합의 정의는 부교재 식 4.18 에서 정의된다.

19.3.1 대각합의 성질

- $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$ for $A, B \in \mathbb{R}^{n\times n}$
- $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr}(A), \alpha \in \mathbb{R}$ for $A \in \mathbb{R}^{n\times n}$
- $\operatorname{tr}(I_n) = n$
- $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for $A \in \mathbb{R}^{n\times k}, B \in \mathbb{R}^{k\times n}$

대각합은 교환법칙이 성립히기 때문에 다음과 같은 성질이 성립한다.

$$\operatorname{tr}(AKL) = \operatorname{tr}(KLA)$$

벡터의 연산에서도 대각합의 교환법칙이 성립디어 다음과 같은 유용한 식이 성립한다.

$$\operatorname{tr}\left(xy^{\top}\right) = \operatorname{tr}\left(y^{\top}x\right) = y^{\top}x \in \mathbb{R}.$$

대각합의 교환법칙때문에 어떤 행렬의 앞에 특정 행렬을 곱하고, 뒤에 역행렬을 곱해도 대각합은 변하지 않는다.

$$\operatorname{tr}\left(S^{-1}AS\right) = \operatorname{tr}\left(ASS^{-1}\right) = \operatorname{tr}(A)$$

19.4 특성다항식

특성다항식(Characteristic polynomial)은 다음과 같이 정의된다 (부교재 definition 4.5)

실수 $\lambda \in \mathbb{R}$ 와 정방행렬(square matrix) $A \in \mathbb{R}^{n\times n}$ 에 대하여

$$\begin{aligned} p_A(\lambda) &:= \det(A - \lambda I) \\ &= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n, \end{aligned} \tag{19.1}$$

19.4.1 행렬식과 대각합과의 관계

$$c_0 = \det(A)$$

$$c_{n-1} = (-1)^n \operatorname{tr}(A)$$

20 고유값과 고유벡터의 정의

노트

- 이 노트의 내용은 부교재 105-110쪽의 내용을 요약한 것이다.
- 부교재 Example 4.5 반드시 공부하세요

20.1 용어

- eigenvalue : 고유값
- eigenvector : 고유벡터

주의

고유값과 고유벡터는 정방행렬(square matrix)에 대해서만 정의된다.

20.2 고유값과 고유벡터

20.2.1 정의

n-차원 정방행렬 $\boldsymbol{A}$ 이 있을 때, 다음 식을 만족하는 λ 와 벡터 $\boldsymbol{x}$가 존재하면 λ 를 행렬 $\boldsymbol{A}$ 의 고유값(eigenvalue), $\boldsymbol{x}$ 를 행렬 $\boldsymbol{A}$ 의 고유벡터(eigenvector)라고 한다 (부교재 definition 4.6)

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

- 고유벡터는 유일하지 않다. 즉, 벡터 $\boldsymbol{x}$ 가 고유벡터이면 $c\boldsymbol{x}$ 도 고유벡터이다.

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

20.2.2 계산

다음 3개의 문장은 동치이다

- λ 는 행렬 $\boldsymbol{A}$ 의 고유값이다.
- 방정식 $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ 은 영벡터이외의 해를 가진다(nontrivial solution)
- λ 는 행렬 $\boldsymbol{A} - \lambda \boldsymbol{I}$ 의 행렬식이 0이다.

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0 \tag{20.1}$$

- λ 는 행렬 $\boldsymbol{A} - \lambda \boldsymbol{I}$ 의 rank가 n 보다 작다.
- Theorem 4.8 에 의하면 위에서 행렬식이 0 인 방정식 식 20.1 을 푸는 것은 식 19.1 의 이 0 을 푸는 것과 동일하다는 것이다.

20.2.3 중복도와 고유공간

부교재의 Definition 4.9, 4.10, 4.11 에 대한 내용입니다.

- 대수적 중복도(algebraic multiplicity) 는 특성다항식 식 19.1 이 0인 방정식을 푸는 경우 다항식에서 고유값이 중근(multiple root)의 해로 나타나는 차수를 의미한다.
- 기하적 중복도(geometric multiplicity) 는 고유값에 대응하는 고유벡터들 중 선형독립인 고유벡터들의 최대 개수를 의미한다.
- 고유 공간(eigenspace)은 고유값에 대응하는 고유벡터들이 생성하는 벡터공간을 의미한다.

예제 20.1. 3차원 행렬 $\boldsymbol{A}$ 가 다음과 같을 때

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

행렬 $\boldsymbol{A}$의 특성다항식은 다음과 같다.

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$$

참고로 특성방정식을 푸는 경우, 방정식 $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$ 이나 $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ 중 어느 것을 사용해도 상관없다.

첫번째 고유값은 $\lambda_1 = 1$ 이다. 고유벡터를 구하기 위해서는 다음과 같은 방정식을 풀면 된다.

$$(\lambda_1 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$$

위의 방정식을 풀면

$$(\lambda_1 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & -1 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

아래와 같이 간단히 할 수 있으며

$$x_1 = -2x_3, \quad x_2 = x_3$$

다음과 같은 고유값과 고유벡터를 얻을 수 있다.

$$\lambda_1 = 1 \quad \rightarrow \quad x_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

첫번째 고유값은 $\lambda_1 = 1$ 이며 대수적 중복도는 1이고 기하적 중복도도 1이다. 이 경우 고유공간 E_1 은 한 개의 고유벡터 $\boldsymbol{x}_1$ 이 생성하는 부분공간을 의미한다.

$$E_1 = \text{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right\}$$

다음으로 두번째 고유값에 대한 방정식 $(\lambda_2 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 을 풀면 다음과 같다.

$$(\lambda_2 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = (2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

이 방정식은 아래와 같이 간단히 할 수 있으며

$$x_1 = -x_3$$

다음과 같은 두 개의 고유벡터를 얻을 수 있다.

$$\lambda_2 = 2 \quad \rightarrow \quad x_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

위에서 **두번째 고유값은** $\lambda_2 = 2$ **이며 대수적 중복도는 2이다.** 또한 **선형독립인 2개의 고유벡터를 구할 수 있으므로 기하적 중복도는 2이다.**

이 경우 E_2 는 두 개의 고유벡터 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 가 생성하는 부분공간을 의미한다.

$$E_2 = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

■

이제 대수적 중복도와 기하적 중복도가 다른 경우에 대한 예제를 들어보자.

예제 20.2. 3차원 행렬 $\boldsymbol{A}$ 가 다음과 같을 때

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

행렬 $\boldsymbol{A}$의 특성다항식은 다음과 같다.

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - 1 & 0 & -2 \\ 1 & \lambda - 1 & -3 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 1)^2 (\lambda - 2)$$

첫번째 고유값은 $\lambda_1 = 1$ 이다. 고유벡터를 구하기 위해서는 다음과 같은 방정식을 풀면 된다.

$$(\lambda_1 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$$

위의 방정식을 풀면

$$(\lambda_1 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 0 & -3 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

아래와 같이 간단히 할 수 있으며

$$x_1 = x_3 = 0$$

다음과 같은 하나의 고유벡터를 얻을 수 있다.

$$\lambda_1 = 1 \quad \rightarrow \quad x_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

첫번째 고유값은 $\lambda_1 = 1$ 이며 대수적 중복도는 2이지만 기하적 중복도는 1이다. 이 경우 고유공간 E_1 은 한 개의 고유벡터 $\boldsymbol{x}_1$ 이 생성하는 부분공간을 의미한다.

$$E_1 = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

다음으로 두번째 고유값에 대한 방정식 $(\lambda_2 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 을 풀면 다음과 같다.

$$(\lambda_2 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = (2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

이 방정식은 아래와 같이 간단히 할 수 있으며

$$x_1 = -2x_3, \quad x_2 = 5x_3$$

다음과 같은 한 개의 고유벡터를 얻을 수 있다.

$$\lambda_2 = 2 \quad \rightarrow \quad x_2 = \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$$

위에서 **두번째 고유값은 $\lambda_2 = 2$ 이며 대수적 중복도는 1이다.** 또한 **선형독립인 1개의 고유벡터를 구할 수 있으므로 기하적 중복도는 1이다.**

이 경우 E_2 는 한 개의 고유벡터 $\boldsymbol{x}_2$ 가 생성하는 부분공간을 의미한다.

$$E_2 = \text{span}\left\{ \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix} \right\}$$

■

---

21 고유값과 고유벡터의 성질

> **노트**
>
> 이 노트의 내용은 부교재 110-119 쪽의 내용을 요약한 것이다.
>
> - 부교재 4.3 Cholesky Decomposition 은 시험범위에서 제외합니다.
> - 부교재 Definition 4.13 의 defective matrix 는 시험범위에서 제외합니다.
> - 부교재 4.4절은 모두 시험범위에 포함됩니다. 특히 Example 4.11 중요하니 꼭 공부하세요

21.1 중요한 성질

- 행렬 $\boldsymbol{A}$ 와 그 전치(transpose) $\boldsymbol{A}^T$ 는 동일한 고유값을 가지나 고유벡터는 같지 않을 수 있다.
- 대칭이고 양정치 행렬(symmetrix positive definite)은 항상 양의 실수 고유값을 갖는다.
- 행렬 $\boldsymbol{A}$ 의 고유값이 모두 다르면 고유벡터들은 선형독립이다.
- 부교재 Theorem 4.15, Example 4.8
- 부교재 Theorem 4.16, Theorem 4.17

21.2 고유값 분해와 대각화

- 고유값분해: Eigendecomposition
- 대각화 : Diagonalization

22 특이값 분해

노트

- 시험 범위는 부교재 119-129 쪽입니다 (4.5절)
- 부교재 4.3 Example 4.13 중요하니 꼭 공부하세요
- 시험은 과제 3 번과 유사하게 출제됩니다.
- 4.6절은 시험범위에서 제외됩니다.

23 벡터 미분

ℹ 노트

- 이 장은 부교재 139-159쪽에 대한 정리입니다.
- 5.4 Gradients of Matrices 는 시험범위에 포함되지 않습니다.
- 5.6.2 Automatic Differentiation 은 Exaample 5.14 를 꼭 읽어보세요.

23.1 용어

- vector differential: 벡터 미분
- partial derivative: 편미분
- gradient: 그레디언트
- Jacobian: 야코비안, 자코비안

23.2 벡터 미분의 표기법

이제 다변량함수(multivariate function), $f : \mathbb{R}^n \to \mathbb{R}^m$ 에 대한 미분을 생각해보자.

먼저 간단한 예제를 고려해 보자. 두 열벡터 $\boldsymbol{x} = (x_1, x_2)^t \in \mathbb{R}_2$, $\boldsymbol{y} = (y_1, y_2, y_3)^t \in \mathbb{R}^3$ 를 고려하고 다음과 같은 함수로 두 벡터의 관계가 정의된다고 하자.

$$y_1 = x_1^2 + x_2, \quad y_2 = \exp(x_1) + 3x_2, \quad y_3 = \sin(x_1) + x_2^3 \tag{23.1}$$

위의 관계를 함수 관계 $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^3$ 로 나타내보면

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ f_3(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2 \\ \exp(x_1) + 3x_2 \\ \sin(x_1) + x_2^3 \end{bmatrix}$$

이러한 경우 다변량 함수 $\boldsymbol{f}$를 벡터 $\boldsymbol{x}$로 미분하려면, 즉 미분 표기법을 이용하려면 편미분을 한 결과를 행렬의 형태를 정해야한다.

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = (n \times m) - \text{matrix} \quad \text{or} \quad (m \times n) - \text{matrix?}$$

일단 각각의 편미분 $\frac{\partial f_i}{\partial x_j}$ 를 구해야 하며 이는 scalar 미분으로 쉽게 구해진다.

$$\begin{aligned}
\frac{\partial f_1}{\partial x_1} &= 2x_1, & \frac{\partial f_2}{\partial x_1} &= \exp(x_1), & \frac{\partial f_3}{\partial x_1} &= \cos(x_1) \\
\frac{\partial f_1}{\partial x_2} &= 1, & \frac{\partial f_2}{\partial x_1} &= 3, & \frac{\partial f_3}{\partial x_1} &= 3x_2^2
\end{aligned} \tag{23.2}$$

이제 이제 편미분값들을 행렬의 형태로 정리해보자. 편미분을 행렬에 배치할 때 다음과 같은 규칙을 사용할 것이다.

- 행렬의 행은 $\boldsymbol{f}$ 의 차원 m 과 같다.
- 행렬의 열은 $\boldsymbol{x}$ 의 차원 n 과 같다.

위와 같이 편미분을 배치하는 벡타 미분 표기법을 분자 표기법 (Numerator layout)이라고 한다.

분자 표기법 (Numerator layout)

$$\boldsymbol{J} = \nabla_x \boldsymbol{x} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \equiv \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}^t} \underset{def}{\equiv} \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 & 1 \\ \exp(x_1) & 3 \\ \cos(x_1) & 3x_2^2 \end{bmatrix}$$

$\boldsymbol{J}$는 야코비안 행렬(Jacobian matrix)이라고 부른다.

이제 이러한 분자표기법의 특별한 결과를 알아보자

- $f : \mathbb{R}^n \to \mathbb{R}^1$ 인 경우

  $f : \mathbb{R}^n \to \mathbb{R}^1$ 인 경우 벡터미분 결과를 **그레디언트(gradient)**라고 부르며 다음과 같이 표기된다.

$$\nabla_x f = \frac{\partial f}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

- $f : \mathbb{R}^1 \to \mathbb{R}^m$ 인 경우

$$\frac{\partial \boldsymbol{f}}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \\ \vdots \\ \frac{\partial f_m}{\partial x} \end{bmatrix}$$

참고로 식 23.1 에서 정의한 함수 관계를 두 벡터 $\boldsymbol{x}$ 와 **y**$ 의 사상관계로 보면

$$\boldsymbol{f} : \boldsymbol{x} \mapsto \boldsymbol{y}$$

다음과 같이 그레디언트 벡터를 표기할 수 있다.

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} \end{bmatrix}$$

23.3 합성함수의 미분법

이제 합성함수의 미분법(chain rule)에 대하여 알아보자.

두 개의 함수

$$\boldsymbol{g} : \mathbb{R}^n \mapsto \mathbb{R}^m, \quad \boldsymbol{f} : \mathbb{R}^m \mapsto \mathbb{R}^p$$

가 있을 때, $\boldsymbol{f}$ 와 $\boldsymbol{g}$ 의 합성함수 $\boldsymbol{h}$ 는 다음과 같이 정의된다.

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x})) = \boldsymbol{f} \circ \boldsymbol{g}$$

즉,

$$\boldsymbol{h} : \mathbb{R}^n \mapsto \mathbb{R}^m \mapsto \mathbb{R}^p$$

이러한 합성함수의 미분은 다음과 같이 계산된다.

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f} \circ \boldsymbol{g}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{g}} \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} \tag{23.3}$$

식 23.3 에서 $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{g}}$ 는 $p \times m$ Jacovian 벡터이고

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{g}} \begin{bmatrix} \frac{\partial f_1}{\partial g_1} & \frac{\partial f_1}{\partial g_2} & \cdots & \frac{\partial f_1}{\partial g_m} \\ \frac{\partial f_2}{\partial g_1} & \frac{\partial f_2}{\partial g_2} & \cdots & \frac{\partial f_2}{\partial g_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial g_1} & \frac{\partial f_p}{\partial g_2} & \cdots & \frac{\partial f_p}{\partial g_m} \end{bmatrix} = (p \times m)$$

$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}$ 는 $n \times m$ Jacovian 벡터이다

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} = (n \times m)$$

합성함수의 미분 공식을 차원으로 나타내면 다음과 같다.

$$\underset{p\times n}{\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}} = \underset{p\times m}{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{g}}} \underset{m\times n}{\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}}$$

23.4 두 벡터 내적의 미분

23.4.1 상수벡터와 변수벡터의 내적

먼저 상수 벡터 $\boldsymbol{a}$와 변수 벡터 $\boldsymbol{x}$의 내적의 미분을 생각해 보자.

참고로 다음과 같이 두 벡터의 내적은 스칼라이다.

$$\boldsymbol{a}^T\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{a} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

따라서 그레이디언트를 구하는 방법과 같이 결과는 행벡터로 표기된다.

$$\frac{\partial \boldsymbol{a}^T\boldsymbol{x}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{x}^T\boldsymbol{a}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T = [a_1, a_2, \dots, a_n]$$

위의 식에서 상수벡터 $\boldsymbol{a}$는 가 전치로 앞에 나타나는 표현 $\boldsymbol{a}^T\boldsymbol{x}$ 를 사용하면 결과 벡터 $\boldsymbol{a}^T$가 행벡터로 그대로 나타나지므로 **내적의 미분 표기**로 사용할 것이다.

$$\frac{\partial \boldsymbol{a}^T\boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T \tag{23.4}$$

23.4.2 상수벡터와 함수벡터의 내적

더 나아가서 상수 벡터 $\boldsymbol{a}$와 함수 벡터 $\boldsymbol{f}$의 내적의 미분도 식 23.4 을 표시하는 동일한 논리로 다음과 같이 표기할 수 있다.

$$\frac{\partial \boldsymbol{a}^T\boldsymbol{f}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \tag{23.5}$$

참고로 식 23.5 에서 $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$는 행벡터가 아닌 행렬로 나타날 수 있다.

23.4.3 함수벡터와 함수벡터의 내적

이제 다음과 같이 같은 공간으로 사상되는 두 함수 $\boldsymbol{f}$ 와 $\boldsymbol{g}$ 의 내적을 생각해 보자.

$$\boldsymbol{f} : \mathbb{R}^n \mapsto \mathbb{R}^m, \quad \boldsymbol{g} : \mathbb{R}^n \mapsto \mathbb{R}^m$$

두 함수의 내적을 미분하는 경우 곱셉 법칙을 적용하여야 하는데 행렬의 곱셉에서는 교환법칙이 성립되지 않으므로 순서에 주의해야 한다.

내적 $\boldsymbol{f}^T\boldsymbol{g}$ 를 각각 따로 미분해야 하는데 각 벡터에 대해 따로 미분을 실행해 보자

- $\boldsymbol{f}$ 를 미분하는 경우 $\boldsymbol{g}$ 는 상수 벡터 $\boldsymbol{a}$ 로 취급한다. 그리고 식 23.5 를 적용한다.

$$\frac{\partial \boldsymbol{f}^T\boldsymbol{g}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f}^T\boldsymbol{a}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{a}^T\boldsymbol{f}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \boldsymbol{g}^T\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \tag{23.6}$$

- $\boldsymbol{g}$ 를 미분하는 경우 $\boldsymbol{f}$ 는 상수 벡터 $\boldsymbol{a}$ 로 취급한다. 그리고 식 23.5 를 적용한다.

$$\frac{\partial \boldsymbol{f}^T\boldsymbol{g}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{a}^T\boldsymbol{g}}{\partial \boldsymbol{x}} = \boldsymbol{a}^T\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} = \boldsymbol{f}^T\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \tag{23.7}$$

이제 위의 두 결과 식 23.6 과 식 23.7 를 합치면 다음과 같은 최종적인 결과를 얻을 수 있다.

$$\frac{\partial \boldsymbol{f}^T\boldsymbol{g}}{\partial \boldsymbol{x}} = \boldsymbol{g}^T\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} + \boldsymbol{f}^T\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} \tag{23.8}$$

23.5 벡터 미분의 응용

23.5.1 선형사상의 미분상

이제 앞에서 배운 벡터의 미분을 이용하여 유용한 응용 공식을 유도해보자.

먼저 선형변환 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 를 생각해 보자. 이때 (M $\times$N)-$\boldsymbol{A}$는 상수 행렬이다. 이때 $\boldsymbol{y}$를 $\boldsymbol{x}$로 미분하면 다음과 같다.

먼저 행렬 $\boldsymbol{A}$의 i 번째 행을 $\boldsymbol{a}_i^T$ 라고 하면

$$\boldsymbol{A} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \dots & A_{MN} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_M^T \end{bmatrix}$$

선형변환 $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ 로 정의하면 다음과 같이 나타낼 수 있다.

$$\boldsymbol{Ax} = \begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{x} \\ \boldsymbol{a}_2^T\boldsymbol{x} \\ \vdots \\ \boldsymbol{a}_M^T\boldsymbol{x} \end{bmatrix} = \begin{bmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_M(\boldsymbol{x}) \end{bmatrix} = \boldsymbol{f}(\boldsymbol{x})$$

따라서

$$\frac{\partial \boldsymbol{Ax}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \frac{\partial f_M}{\partial x_2} & \cdots & \frac{\partial f_M}{\partial x_N} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix} = \boldsymbol{A}$$

따라서 선형사상의 미분은 그 자신이다.

$$\frac{\partial \boldsymbol{Ax}}{\partial \boldsymbol{x}} = \boldsymbol{A} \tag{23.9}$$

23.5.2 이차형식의 미분

이차형식 $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Bx}$ 을 미분하는 경우 식 23.8 을 적용한다. 이 경우 $\boldsymbol{f} = \boldsymbol{x}^T$ 이고 $\boldsymbol{g} = \boldsymbol{Bx}$ 이라고 놓고 식 23.8 을 적용하면 다음과 같다.

$$\begin{aligned} \frac{\partial \boldsymbol{x}^T\boldsymbol{Bx}}{\partial \boldsymbol{x}} &= \frac{\partial \boldsymbol{x}^T(\boldsymbol{Bx})}{\partial \boldsymbol{x}} \\ &= (\boldsymbol{Bx})^T\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}} + \boldsymbol{x}^T\frac{\partial (\boldsymbol{Bx})}{\partial \boldsymbol{x}} \\ &= \boldsymbol{x}^T\boldsymbol{B}^T + \boldsymbol{x}^T\boldsymbol{B} \\ &= \boldsymbol{x}^T(\boldsymbol{B}^T + \boldsymbol{B}) \end{aligned}$$

만약 행렬 $\boldsymbol{B}$가 대칭행렬이면 $\boldsymbol{B}^T = \boldsymbol{B}$ 이므로 다음과 같이 간단하게 나타낼 수 있다.

$$\frac{\partial \boldsymbol{x}^T\boldsymbol{Bx}}{\partial \boldsymbol{x}} = \boldsymbol{x}^T(\boldsymbol{B}+\boldsymbol{B}) = 2\boldsymbol{x}^T\boldsymbol{B} \tag{23.10}$$

23.5.3 최소제곱법의 미분

부교재 Example 5.11 의 내용을 다음과 같이 정리할 수 있다.

$$\boldsymbol{y} = \boldsymbol{\Phi\theta},$$

여기서 $\theta \in \mathbb{R}^D$ 는 모수벡터(parameter vector), $\Phi \in \mathbb{R}^{N\times D}$ are 입력변수(input features), $y \in \mathbb{R}^N$ are the 관측값 벡터(observation vector).

다음으로 손실함수(loss function) L 과 오타벡터 $$ 를 다음과 같이 정의하자

$$
\begin{aligned}
L(e) &:= \|e\|^2 = \boldsymbol{e}^T\boldsymbol{e} \\
e(\theta) &:= y - \Phi\theta.
\end{aligned}
$$

이때 손실함수 L 을 최소화하는 θ 를 구하는 문제는 손실함수 L 을 θ 로 미분하여 0 이 되는 θ 를 구하는 문제로 바뀐다.

이러한 경우 손실함수 L 을 θ 로 미분하는 경우 다음과 같이 정리할 수 있다.

$$
\begin{aligned}
\frac{\partial L}{\partial \boldsymbol{\theta}} &= \frac{\partial L}{\partial \boldsymbol{e}}\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\theta}} \\
&= \frac{\partial \boldsymbol{e}^T\boldsymbol{e}}{\partial \boldsymbol{e}}\frac{\partial(\boldsymbol{y} - \boldsymbol{\Phi\theta})}{\partial \boldsymbol{\theta}} \\
&= 2\boldsymbol{e}^T\frac{\partial \boldsymbol{\Phi\theta}}{\partial \boldsymbol{\theta}} \\
&= 2\boldsymbol{e}^T\boldsymbol{\Phi} \\
&= 2(y - \Phi\theta)^T\boldsymbol{\Phi}
\end{aligned}
$$

참고로 위의 식에서 첫 번째 식은 합성함수의 공식(식 23.3), 세 번째 식은 이차형식의 미분(식 23.10) 과 선형사상의 미분(식 23.9) 을 적용하였다.

References

Deisenroth, Marc Peter, A Aldo Faisal, 와/과 Cheng Soon Ong. 2020. *Mathematics for machine learning*. Cambridge University Press.