공학 수학
강의 노트

이재호

jaeho.lee@snu.ac.kr

마지막 수정: 2018년 11월 7일

7 Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

2018.9.3. Field는 수학과 물리학에서 지칭하는 대상이 다르다:

- **체體**: 실수체, 복소수체
- **장場**: 전자기장, 벡터장

정의 7.1. 체는 $+, -, \times, \nabla\cdot$에 대해 닫혀 있는 수 집합을 말한다.

예시.

- $\mathbb{Q}$는 조밀 (dense) 하다.
- $\mathbb{R}$은 꽉차 (complete) 있다.
- $\mathbb{C}$는 대수적으로 닫혀 (closed) 있다.
- $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$
- $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$

정의 7.2. 벡터공간은 (roughly) 덧셈 (+) 과 (어떤 체에 속하는) 상수배가 정의된 집합이다.

예시.

- $\mathbb{R}^2 = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$은 덧셈에 대해 닫혀 있고, $k \in \mathbb{R}$의 곱에 대해 닫혀 있으므로 실수체에 대한 벡터공간이다.

- $\mathbb{R}^n$ 과 $\mathbb{C}^n$ 은 실수체에 대한 벡터공간이면서 유리수체에 대한 벡터공간이다. 이를 각각 $\mathbb{Q}$–벡터공간, $\mathbb{R}$–벡터공간이라고 한다.

- $\mathcal{C}_I^0 = \{f : I \to \mathbb{R} \mid f$는 연속 함수, $I \subset \mathbb{R}\}$ 에서는, $x \mapsto \sin x$가 하나의 벡터이다.

- $\mathcal{C}_I^n = \{f : I \to \mathbb{R} \mid \exists f^{(n)},\ f^{(n)}$는 연속 함수, $I \subset \mathbb{R}\}$ 이며, $\mathcal{C}_I^0 > \mathcal{C}_I^1 > \cdots > \mathcal{C}_I^\infty$ 이다.

정의 7.2. 체 F에 대한 벡터공간 V의 원소 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$에 대해서,

1. $c_1, \ldots, c_k \in F$일 때
$$c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$$
는 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$의 일차결합, 혹은 선형결합 (linear combination) 이라고 한다.

2. $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$의 선형 생성
$$\mathrm{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$$
은 모든 일차결합의 집합이다. 예를 들어, 어떤 평면 상의 평행하지 않은 두 벡터는 그 평면을 선형 생성한다.

3. $W \subset V$이면서 W가 벡터공간이면, W는 V의 부분공간이라고 하며 $W < V$로 표기한다. 예를 들어, 실수 평면으로 나타내어지는 $\mathbb{R}^2$ 벡터공간의 부분집합인 원점을 지나는 직선은 벡터공간이므로 부분공간이다. 반면, 원점을 지나지 않는 직선은 부분집합이지만 벡터공간은 아니므로 부분공간이 아니다.

4. $W = \mathrm{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} < V$일 때, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$는 W의 생성자 (generator) 라고 한다.

5. $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$가 일차종속 (linearly dependent) 라는 것은 어느 하나가 다른 벡터들의 일차결합이라는 것이다. 일차종속이 아니면 일차독립 (linearly independent) 이라고 한다. 예를 들어, $\mathbb{R}^3$에서 $(1,2,3), (4,5,6), (7,8,9)$는 일차종속인 반면, $(1,2,3), (4,5,6)$은 일차독립이다.

2018.9.5. **정리 7.3.** *체 F에 대한 벡터공간 V의 원소들 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$가 일차독립이라는 것은, $c_1, \ldots, c_k \in F$일 때*
$$c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k = \mathbf{0} \ \Rightarrow\ c_1 = \cdots = c_k = 0$$
이라는 것과 동치이다.

증명. (⇒) (재배열 가능하여) $c_k \neq 0$이라고 가정하자. 그렇다면,

$$\boldsymbol{v}_k = -\left(\frac{c_1}{c_k}\boldsymbol{v}_1 + \cdots + \frac{c_{k-1}}{c_k}\boldsymbol{v}_{k-1}\right)$$

이므로 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$가 일차독립이라는 가정에 모순된다.

(⇐) $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$가 일차종속이라고 가정하자. 즉, (재배열 가능하여) 어떤 $a_1, \ldots, a_{k-1} \in F$가 존재해서,

$$\boldsymbol{v}_k = a_1\boldsymbol{v}_1 + \cdots + a_{k-1}\boldsymbol{v}_{k-1}$$

이다. 그렇다면

$$a_1\boldsymbol{v}_1 + \cdots + a_{k-1}\boldsymbol{v}_{k-1} + (-1)\boldsymbol{v}_k = \mathbf{0}$$

이므로 모순이다. □

예시.

- $a, b, c \in \mathbb{R}$에 대해

  $$a(1,2,3) + b(4,5,6) + c(7,8,10) = (0,0,0)$$

  을 만족하는 a, b, c는 0밖에 없으므로 $(1,2,3), (4,5,6), (7,8,10)$은 일차독립이다.

- $\mathbf{0}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$는 일차종속이다.

- $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$가 일차독립이면 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 또한 일차독립이다. ($\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$가 종속이면 $\boldsymbol{v}_3 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + 0\boldsymbol{v}_4$이기 때문이다.)

정의 7.4. 벡터공간 V의 부분공간 W가 일차독립인 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$에 의해 생성될 때, $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$를 W의 기저(basis)라고 한다.

정리 7.5 (차원(dimension) 정리). *W가 유한생성된(finitely generated) (부분)공간이면 W의 기저의 원소의 개수는 동일하다. 이때, 기저의 원소 개수를 W의 차원(dimension)이라고 하며,* $\dim W$*로 표기한다.*

예시.

- $\mathcal{P}_n = \{n$차 이하 실계수 다항식$\}$일 때, $f \in \mathcal{P}_n$는 항상 $\mathcal{B} = \{1, x, \ldots, x^n\}$의 원소의 상수배의 합으로 나타낼 수 있다. 또한,

  $$c_0 \cdot 1 + c_1 x + \cdots + c_n x^n = 0$$

  이면 $c_0 = \ldots c_n = 0$이므로, $\mathcal{B}$는 기저이다.

- $\mathcal{P} = \{$모든 다항식$\}$의 기저는 $\{1, x, x^2, \dots\}$이며 $\dim \mathcal{P} = \infty$이다.

참고.

1. a_1, a_2, a_3는 서로 다른 실수이다. 이때, $(1, a_1, a_1^2), (1, a_2, a_2^2), (1, a_3, a_3^2)$이 일차독립임이다.

2. $a_1, \dots, a_n$는 서로 다른 실수이다. 이때, $(1, a_1, a_1^2), \dots, (1, a_n, a_n^2)$이 일차독립이다.

   증명.
$$\begin{pmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \\ \vdots & \ddots & \vdots \\ a_1^{n-1} & \dots & a_n^{n-1} \end{pmatrix}$$
   의 열벡터들이 독립을 보이면 되는데, 이는 뒤에서 보겠지만 행벡터들이 독립임을 보이는 것과 동치이다. 따라서, 다음의 일차결합을 생각할 수 있다:
$$c_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} a_1 \\ \vdots \\ a_n^{n-1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} a_1^{n-1} \\ \vdots \\ a_n^{n-1} \end{pmatrix} = \mathbf{0}$$
   그런데 $n-1$차 식
$$c_1 + c_2 x + \cdots + c_n x^{n-1} = 0$$
   의 해가 $x = a_1, \dots, a_n$으로 n개라는 것은 위 식이 항등식이라는 것이다. 따라서
$$c_1 = \cdots = c_n = 0$$
   이다. □

정리 7.6. *체 F에 대한 벡터공간 V의 일차독립인 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$에 대해서, $\boldsymbol{w} \in V \setminus \langle \boldsymbol{v}_1, \dots, \boldsymbol{v}_k \rangle$이면 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{w}$는 일차독립이다.*

증명. $\boldsymbol{w} \notin \langle \boldsymbol{v}_1, \dots, \boldsymbol{v}_k \rangle$이므로
$$\boldsymbol{w} = c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k$$
인 $c_i \in F$ $(1 \le i \le k)$들이 존재하지 않는다.

만약 $1 \le j \le k$인 어떤 $\boldsymbol{v}_j$가

$$\boldsymbol{v}_j = c_1\boldsymbol{v}_1 + \cdots + c_{i-1}\boldsymbol{v}_{i-1} + c_{i+1}\boldsymbol{v}_{i+1} + \cdots + c_k\boldsymbol{v}_k + d\boldsymbol{w}$$

와 같은 일차결합으로 나타내진다면, $d = 0$일 경우 $\boldsymbol{v}_j$가 나머지들의 일차결합이 되어 일차독립성에 모순이다. $d \neq 0$일 경우에는

$$\boldsymbol{w} = \frac{1}{d}(\boldsymbol{v}_j - c_1\boldsymbol{v}_1 - \cdots - c_{i-1}\boldsymbol{v}_{i-1} - c_{i+1}\boldsymbol{v}_{i+1} - \cdots - c_k\boldsymbol{v}_k)$$

이 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$에 속하게 되므로 모순이다. 따라서 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}$는 일차독립이다. □

정리 7.6에 따르면, 일차독립인 벡터들의 선형 생성에 포함되지 않는 벡터 또한 이들과 일차독립이다. 이에 따라 어떤 유한생성된 벡터공간 V의 일차독립인 벡터들 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$가 주어졌을 때, 선형 생성에 속하지 않는 벡터 $\boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_{\dim V}$를 순차적으로 추가해서 기저를 구성할 수 있다. 이를 기저 확장 (basis extension)이라고 부른다.

정리 7.7. *체 F에 대한 벡터공간 V의 부분공간 W_1, W_2가 주어졌을 때, $W_1 \cap W_2$ 또한 V의 부분공간이다.*

증명. $\boldsymbol{v}_1, \boldsymbol{v}_2 \in W_1$이면서 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in W_2$이면, 각각 벡터공간이므로 임의의 $c_1, c_2 \in F$에 대해서 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 \in W_1$이면서 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 \in W_2$이다. 즉,

$$\boldsymbol{v}_1, \boldsymbol{v}_2 \in W_1 \cap W_2 \;\Rightarrow\; \forall c_1, c_2 \in F \quad c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 \in W_1 \cap W_2$$

이다. 따라서, $W_1 \cap W_2 < V$가 성립한다. □

정의 7.8. 벡터공간 V의 부분집합 W_1, W_2에 대해서 $W_1 + W_2$를 다음과 같이 정의한다:

$$W_1 + W_2 := \{\boldsymbol{v}_1 + \boldsymbol{v}_2 \mid \boldsymbol{v}_1 \in W_1, \boldsymbol{v}_2 \in W_2\}.$$

예시. 좌표 평면으로 나타내어지는 벡터공간 $\mathbb{R}^2$의 부분집합 (이면서 부분공간인)인 원점을 지나는 직선을 W_1이라고 하자. 원점이 아닌 한 점만을 원소로 하는 집합 $W_2 \subset \mathbb{R}^2$가 주어졌을 때, $W_1 + W_2$는 W_1의 직선을 W_2의 (유일한) 원소의 점으로 평행 이동한 직선을 나타낸다.

정리 7.9. *체 F에 대한 벡터공간 V의 부분공간 W_1과 W_2에 대해서, $W_1 + W_2$도 V의 부분공간이다.*

증명. 임의의 $\boldsymbol{v}_1 \in W_1 + W_2$에 대해 어떤 $\boldsymbol{w}_1 \in W_1$과 $\boldsymbol{w}_2 \in W_2$가 존재하며, 마찬가지로 임의의 $\boldsymbol{v}_2 \in W_1 + W_2$에 대해 어떤 $\boldsymbol{u}_1 \in W_1$과 $\boldsymbol{u}_2 \in W_2$가 존재한다.

W_1과 W_2는 부분공간이므로 임의의 $c_1, c_2 \in F$에 대해

$$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 \in W_1$$
$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 \in W_2$$

이고,

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = (c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2) + (c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2)$$

이므로,

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 \in W_1 + W_2$$

이다. 따라서 $W_1 + W_2 < V$이다. □

2018.9.10. **정리 7.10** (그라스만 (Graßmann) 공식). *체 F에 대한 유한생성된 벡터공간 U와 W에 대해, 다음이 성립한다:*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

증명. l, m, n을

$$l = \dim(U \cap W)$$
$$n = \dim U - l$$
$$m = \dim W - l$$

이라 하고, $U \cap W$의 기저를 $\mathcal{B}_\cap = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l\}$라고 하자. $U \setminus W$에서 $\mathcal{B}_\cap$을 확장해 U의 기저 $\mathcal{B}_U = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$와 $W \setminus U$에서 확장해 W의 기저 $\mathcal{B}_W = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$을 구성할 수 있다.

이제 $\mathcal{B}_+ = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$가 $U + W$의 기저임을 보이자. 먼저 $\mathcal{B}_+$가 일차독립임을 보이자.

다음의 일차결합

$$c_1\boldsymbol{v}_1 + \cdots + c_l\boldsymbol{v}_l + d_1\boldsymbol{u}_1 + \cdots + d_n\boldsymbol{u}_n + e_1\boldsymbol{w}_1 + \cdots + e_m\boldsymbol{w}_m = \boldsymbol{0}$$

을 생각하자. 이때 $\boldsymbol{z}$를

$$\boldsymbol{z} = c_1\boldsymbol{v}_1 + \cdots + c_l\boldsymbol{v}_l + d_1\boldsymbol{u}_1 + \cdots + d_n\boldsymbol{u}_n = -e_1\boldsymbol{w}_1 - \cdots - e_m\boldsymbol{w}_m$$

로 정의하면, $\boldsymbol{z} \in U \cap W = \operatorname{Span}\mathcal{B}_\cap$이다. 따라서

$$\boldsymbol{z} = c_1'\boldsymbol{v}_1 + \cdots + c_l'\boldsymbol{v}_l$$

로 쓸 수 있고, 기저에 의한 일차결합은 유일하게 표현되므로 모든 계수 c_i, d_j, e_k, c_i' 은 0이다. 그러므로 $\mathcal{B}_+$는 일차독립이다.

이제 $\operatorname{Span}\mathcal{B}_+$가 $U+W$임을 보이자. $\mathcal{B}_+ \subset U+W$이므로 $\operatorname{Span}\mathcal{B}_+ \subseteq U+W$ 이고, $\operatorname{Span}\mathcal{B}_+ \supseteq U+W$임을 보이면 된다.

$U+W$의 원소 $\boldsymbol{x}$를 고르면, $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w}$가 되는

$$\boldsymbol{u} = c_1\boldsymbol{v}_1 + \cdots + c_l\boldsymbol{v}_l + d_1\boldsymbol{u}_1 + \cdots + d_n\boldsymbol{u}_n$$

과

$$\boldsymbol{v} = c_1'\boldsymbol{v}_1 + \cdots + c_l'\boldsymbol{v}_l + e_1\boldsymbol{w}_1 + \cdots + e_n\boldsymbol{w}_k$$

이 존재한다. 따라서

$$\boldsymbol{x} = (c_1 + c_1')\boldsymbol{v}_1 + \cdots + (c_l + c_l')\boldsymbol{v}_l + d_1\boldsymbol{u}_1 + \cdots + d_n\boldsymbol{u}_n + e_1\boldsymbol{w}_1 + \cdots + e_k\boldsymbol{w}_k$$

는 $\mathcal{B}_+$의 일차결합이므로 $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w} \in \operatorname{Span}\mathcal{B}_+$이다. 따라서 $\operatorname{Span}\mathcal{B}_+ \supseteq U+W$ 이고, $\operatorname{Span}\mathcal{B}_+ = U+W$이다.

그러므로 $\mathcal{B}_+$는 $U+W$의 기저가 되어,

$$\dim(U+W) = l + n + m = \dim U + \dim W - \dim(U \cap W)$$

이다. □

정의 7.11. 체 F에 대한 벡터공간 V, W에서, 임의의 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$와 $c_1, c_2 \in F$에 대해

$$L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1L(\boldsymbol{v}_1) + c_2L(\boldsymbol{v}_2)$$

을 만족하는 함수 $L : V \to W$을 선형변환 (linear transformation) 혹은 선형사상 (linear map) 이라고 한다.

예시.

1. $L : \mathbb{R}^2 \to \mathbb{R}^3$에 대해,

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+2y \\ 3x+4y \\ 5x+6y \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

   은 선형변환이다.

2. $L : \mathbb{R}^2 \to \mathbb{R}^3$ 에 대해,

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+2y \\ 3x+4y \\ 5x+6y+1 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

   은 선형변환이 아니다.

3. 어떤 사상 $L : \mathbb{R}^n \to \mathbb{R}^m$ 이 존재한다는 것은 어떤 $m \times n$ 행렬 A가 존재해서 $L(\boldsymbol{x}) = A\boldsymbol{x}$라는 것이다.

4. $f \mapsto f'$ 의 미분 연산자 $L : \mathcal{C}_I^\infty \to \mathcal{C}_I^\infty$ 는 $L(c_1 f_1 + c_2 f_2) = c_1 f_1' + c_2 f_2' = c_1 L(f_1) + c_2 L(f_2)$ 이므로 선형변환이다.

5. $f(x) \mapsto \int_0^x f(t)\,\mathrm{d}t$ 의 적분 연산자 $L : \mathcal{C}_I^\infty \to \mathcal{C}_I^\infty$ 또한 선형변환이다.

정의 7.12. 두 벡터공간 V와 W 간의 선형변환 $L : V \to W$에 대해 다음이 정의된다:

1. L의 핵(kernel) 혹은 영 공간(null space)[1]

$$\ker L = \mathrm{N}(L) := \{\boldsymbol{v} \in V \mid L(\boldsymbol{v}) = \boldsymbol{0}\}.$$

2. L의 치역(range) 혹은 상(image)[2]

$$\mathrm{R}(L) = \mathrm{im}(L) := \{L(\boldsymbol{v}) \mid \boldsymbol{v} \in V\}.$$

3. L이 일대일 함수(one-to-one function) 혹은 단사 함수(injective function, injection)라는 것은 다음을 의미한다:

$$\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V \quad \big(L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2) \ \Rightarrow\ \boldsymbol{v}_1 = \boldsymbol{v}_2\big)$$

4. L이 위로의(onto) 함수 혹은 단사 함수(surjective function, surjection)라는 것은 다음을 의미한다:

$$\forall \boldsymbol{w} \in W \quad \exists \boldsymbol{v} \in V \quad L(\boldsymbol{v}) = \boldsymbol{w}$$

예시.

---

[1] 영 공간이라는 용어는 보통 행렬에 국한되어 사용되고, 핵은 추상적인 선형변환에 대해 많이 사용된다.

[2] 치역은 정의역의 상이다. 상은 정의역의 부분집합에 대해서도 정의될 수 있다.

1. $L\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$ 일 때 $\ker L\ \big(=\mathrm{N}(L)\big)$ 을 구하시오.

풀이.

$$\begin{aligned}
\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} &\to \begin{pmatrix}1&2&3\\0&-3&-6\\0&-6&-12\end{pmatrix}\\
\xrightarrow{\text{Echelon}} \begin{pmatrix}1&2&3\\0&1&2\\0&0&0\end{pmatrix} &\xrightarrow{\text{Reduce}} \begin{pmatrix}1&0&-1\\0&1&2\\0&0&0\end{pmatrix}
\end{aligned}$$

즉

$$\begin{pmatrix}1&0&-1\\0&1&2\\0&0&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \Rightarrow \begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}1\\-2\\1\end{pmatrix} z$$

이 되어, $\mathrm{N}(L) = \langle(1,-2,1)\rangle$ 이다. ■

2. $\mathrm{R}(L)\ \big(=\mathrm{im}(L)\big)$ 을 구하시오.

풀이. $L\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$ 이므로,

$$\begin{aligned}
L(\boldsymbol{e}_1) &= (1,4,7)\\
L(\boldsymbol{e}_2) &= (2,5,8)\\
L(\boldsymbol{e}_3) &= (3,6,9)
\end{aligned}$$

가 된다. 따라서

$$\begin{aligned}
\mathrm{R}(L) &= \{L(x,y,z) \mid (x,y,z)\in V\}\\
&= \{L(x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3) \mid (x,y,z)\in V\}\\
&= \{xL(\boldsymbol{e}_1) + yL(\boldsymbol{e}_2) + zL(\boldsymbol{e}_3) \mid (x,y,z)\in V\}\\
&= \mathrm{Span}\{L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), L(\boldsymbol{e}_3)\}\\
&= \langle(1,4,7),(2,5,8),(3,6,9)\rangle\\
&= \langle(1,4,7),(2,5,8)\rangle
\end{aligned}$$

이다. ■

정리 7.13. *벡터공간 V 와 W 간의 선형사상 $L : V \to W$ 에 대해* $\mathrm{N}(L) < V$ *이고* $\mathrm{R}(L) < W$ *이다.*

증명. $\mathrm{N}(L)$은

$$\mathrm{N}(L) = \{\boldsymbol{v} \in V \mid L(\boldsymbol{v}) = \boldsymbol{0}\}$$

이므로, $\boldsymbol{v}_1$과 $\boldsymbol{v}_2$를 $\mathrm{N}(L)$에서 고르면 둘의 일차결합 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2$의 함숫값은

$$L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1 L(\boldsymbol{v}_1) + c_2 L(\boldsymbol{v}_2) = \boldsymbol{0}$$

이다. 따라서 $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2$도 $\mathrm{N}(L)$의 원소이다. $\mathrm{N}(L)$은 일차결합에 대해 닫혀있고 V의 부분집합이므로 V의 부분공간이다.

$\mathrm{R}(L)$은

$$\mathrm{R}(L) = \{\boldsymbol{w} \mid \boldsymbol{u} \in V \wedge \boldsymbol{w} = L(\boldsymbol{v})\}$$

와 같이 정의되고, $\mathrm{R}(L)$의 원소 $\boldsymbol{w}_1, \boldsymbol{w}_2$에 대해서

$$\begin{aligned} \boldsymbol{w}_1 &= L(\boldsymbol{v}_1) \\ \boldsymbol{w}_2 &= L(\boldsymbol{v}_2) \end{aligned}$$

인 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$가 존재한다. 따라서

$$L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1 L(\boldsymbol{v}_1) + c_2 L(\boldsymbol{v}_2) = c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2$$

는 $\mathrm{R}(L)$에 속한다. 그러므로 $\mathrm{R}(L)$도 일차결합에 대해 닫혀있고 W의 부분집합이므로 W의 부분공간이다. □

정리 7.14 (계수(rank)–퇴화 차수(nullity) 정리)**.** *벡터공간 V 와 W 간의 선형사상 $L : V \to W$ 에 대해 다음이 성립한다:*

$$\dim \mathrm{N}(L) + \dim \mathrm{R}(L) = \dim V.$$[3]

증명. 정리 7.13에 의해 $\mathrm{N}(L)$은 V의 부분공간이므로, $\mathrm{N}(L)$의 기저를 확장해 V의 기저를 구성할 수 있다. 따라서 $\mathrm{N}(L)$의 기저 $\mathcal{B}_N$를 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$으로 놓는다면, V의 기저 $\mathcal{B}_V$를 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l\}$로 확장할 수 있다. 이제 $\mathcal{B}_R = \{L(\boldsymbol{u}_1), \ldots, L(\boldsymbol{u}_l)\}$이 $\mathrm{R}(L)$의 기저임을 보이자. 먼저 $\mathcal{B}_R$이 일차독립임을 보이자.

[3]계수와 퇴화 차수에 대해서는 다음 강좌에 정의한다.

$\mathcal{B}_R$ 의 일차결합

$$c_1L(\boldsymbol{u}_1)+\cdots+c_lL(\boldsymbol{u}_l)=\mathbf{0}$$

이면

$$L(c_1\boldsymbol{u}_1+\cdots+c_l\boldsymbol{u}_l)=\mathbf{0}$$

이므로

$$\boldsymbol{z}=c_1\boldsymbol{u}_1+\cdots+c_l\boldsymbol{u}_l$$

로 두면 $\boldsymbol{z}\in \mathrm{N}(L)$ 이다. 그런데 $\boldsymbol{z}\in V$ 이기도 하므로

$$\boldsymbol{z}=c_1\boldsymbol{u}_1+\cdots+c_l\boldsymbol{u}_l+0\boldsymbol{v}_1+\cdots+0\boldsymbol{v}_n=d_1\boldsymbol{v}_1+\dots d_n\boldsymbol{v}_n$$

이다. 따라서 기저의 일차결합 표현의 유일성에 의해 계수 c_i, d_j 는 모두 0이다. 그러므로 $\mathcal{B}_R$ 은 일차독립이다.

이제 $\operatorname{Span}\mathcal{B}_R=\mathrm{R}(L)$ 임을 보이다. $\mathcal{B}_R\subset\mathrm{R}(L)$ 이고 $\mathrm{R}(L)$ 은 정리 7.13에 따라 벡터공간이므로 $\operatorname{Span}\mathcal{B}_R\subseteq\mathrm{R}(L)$ 이다. 따라서 $\operatorname{Span}\mathcal{B}_R\supseteq\mathrm{R}(L)$ 임을 보이면 된다.

$\boldsymbol{w}\in\mathrm{R}(L)$ 를 고르면 $\boldsymbol{w}=L(\boldsymbol{v})$ 인 $\boldsymbol{v}\in V$ 가 존재한다. $\boldsymbol{v}\in\operatorname{Span}\mathcal{B}_V$ 이므로

$$\boldsymbol{v}=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n+d_1\boldsymbol{u}_1+\cdots+d_l\boldsymbol{u}_l$$

로 나타낼 수 있다. 그러므로

$$\boldsymbol{w}=c_1L(\boldsymbol{v}_1)+\cdots+c_nL(\boldsymbol{v}_n)+d_1L(\boldsymbol{u}_1)+\cdots+d_lL(\boldsymbol{u}_l)$$

이고, $L(\boldsymbol{v}_i)=\mathbf{0}$ 이므로

$$\boldsymbol{w}=d_1L(\boldsymbol{u}_1)+\cdots+d_lL(\boldsymbol{u}_l)$$

이다. 즉, $\boldsymbol{w}\in\operatorname{Span}\mathcal{B}_R$ 이고, $\mathrm{R}(L)\subseteq\operatorname{Span}\mathcal{B}_R$ 이다. 결국 $\operatorname{Span}\mathcal{B}_R=\mathrm{R}(L)$ 이므로, $\mathcal{B}_R$ 은 $\mathrm{R}(L)$ 의 기저이다.

따라서 $\dim\mathrm{R}(L)=l=\dim V-\dim\mathrm{N}(L)$ 이다. □

예시. 이전 예시에서 $\mathrm{N}(L)=\langle(1,-2,1)\rangle$ 이고 $\mathrm{R}(L)=\langle(1,4,7),(2,5,8)\rangle$ 이므로, $\dim\mathrm{N}(L)+\dim\mathrm{R}(L)=1+2=3$ 이 성립한다.

2018.9.12. **정리 7.15.** *벡터 공간 V 와 W 간의 선형변환 $L:V\to W$ 에 대해서 $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ 이 V 의 기저라면,*

$$\mathrm{R}(L)=\operatorname{Span}\{L(\boldsymbol{v}_1),\dots,L(\boldsymbol{v}_n)\}$$

이다.

증명. $L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)$은 $\mathrm{R}(L)$의 원소이므로 $\langle L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)\rangle \subseteq \mathrm{R}(L)$이다.

$\mathrm{R}(L)$의 원소 $\boldsymbol{w}$를 고르자. 그러면 $L(\boldsymbol{v}) = \boldsymbol{w}$인 $\boldsymbol{v} \in V$가 존재한다. 따라서

$$\boldsymbol{w} = L(\boldsymbol{v}) = c_1 L(\boldsymbol{v}_1) + \cdots + c_n L(\boldsymbol{v}_n)$$

으로 표현할 수 있고, $\boldsymbol{w} \in \langle L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)\rangle$이다. 즉 $\mathrm{R}(L) \subseteq \langle L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)\rangle$이다. 결국 $\mathrm{R}(L) = \langle L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)\rangle$이다. □

참고. 위에서 $\{L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)\}$이 기저일 필요는 없다.

정리 7.16. *벡터공간 V와 W 간의 선형변환 $L : V \to W$가 일대일이라는 것은 $\mathrm{N}(L) = \{\mathbf{0}\}$라는 것과 동치이다.*

증명. ($\Rightarrow$) $L(\mathbf{0}) = \mathbf{0}$이고, L은 일대일이므로 $L(\boldsymbol{v}) = \mathbf{0}$를 만족하는 $\boldsymbol{v} = \mathbf{0}$가 유일하다.

($\Leftarrow$) $L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2)$인 임의의 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$를 고르자. 그렇다면

$$L(\boldsymbol{v}_1) - L(\boldsymbol{v}_2) = L(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \mathbf{0}$$

인데, $\mathrm{N}(L) = \{\mathbf{0}\}$이므로 $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \mathbf{0}$이다. 따라서 $\boldsymbol{v}_1 = \boldsymbol{v}_2$이며, L은 일대일 함수이다. □

정리 7.17. *벡터공간 V와 W 간의 선형변환 $L : V \to W$가 일대일이라는 것은 $\dim \mathrm{R}(L) = \dim V$라는 것과 동치이다.*

증명. L이 일대일이라는 것은 정리 7.16에 따라 $\mathrm{N}(L) = \{\mathbf{0}\}$, 즉 $\dim \mathrm{N}(L) = 0$이라는 것과 동치이다.[4] 그런데 정리 7.14에 따라서 $\dim \mathrm{N}(L) + \dim \mathrm{R}(L) = \dim V$이므로, $\dim \mathrm{R}(L) = \dim V$가 성립한다. □

정리 7.18. *벡터공간 V와 W 간의 선형변환 $L : V \to W$가 전단사라는 것은 L의 역함수 $L^{-1} : W \to V$가 존재한다는 것과 동치이다. 이때, L^{-1}도 선형이다.*

증명. 전단사 함수가 역함수를 가진다는 증명은 생략한다. L^{-1}이 선형임을 보인다.

W에서 두 원소 $\boldsymbol{w}_1$와 $\boldsymbol{w}_2$를 고르자. 이때,

$$L^{-1}(\boldsymbol{w}_1) = \boldsymbol{v}_1$$
$$L^{-1}(\boldsymbol{w}_2) = \boldsymbol{v}_2$$

[4] $\dim \mathrm{N}(L) = 0 \Leftrightarrow \mathrm{N}(L) = \{\mathbf{0}\}$

라고 하자. 그렇다면 $\boldsymbol{w}_1$과 $\boldsymbol{w}_2$의 선형결합은 다음과 같이 표현된다:

$$\begin{aligned} c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 &= c_1L(\boldsymbol{v}_1) + c_2L(\boldsymbol{v}_2) \\ &= L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) \end{aligned}$$

양변에 역함수를 취해주면

$$\begin{aligned} L^{-1}(c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2) &= c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 \\ &= c_1L^{-1}(\boldsymbol{w}_1) + c_2L^{-1}(\boldsymbol{w}_2) \end{aligned}$$

이 되어 L^{-1} 또한 선형변환임을 알 수 있다. □

정의 7.19. 벡터공간 V와 W간의 선형변환 $L : V \to W$에 대해서, L의 계수

$$\operatorname{rank} L := \dim \mathrm{R}(L)$$

로 정의한다.

정리 7.20. *벡터공간 U, V, W에 대해 선형변환 $L_1 : U \to V$와 $L_2 : V \to W$이 주어졌다고 하자. 만약 L_2가 일대일 함수라면, 다음이 성립한다:*

$$\operatorname{rank}(L_2 \circ L_1) = \operatorname{rank} L_1$$

증명. 정리 7.13에 의해 $\mathrm{R}(L_1)$은 벡터공간이므로, 선형변환 $\tilde{L}_2 : \mathrm{R}(L_1) \to W$를 정의하자:

$$\forall \boldsymbol{v} \in \mathrm{R}(L_1) < V \qquad \tilde{L}_2(\boldsymbol{v}) = L_2(\boldsymbol{v})$$

따라서 $\tilde{L}_2$ 또한 일대일 선형변환임을 알 수 있다. 정리 7.17에 따라

$$\dim \mathrm{R}(\tilde{L}_2) = \dim \mathrm{R}(L_1)$$

이다. 그런데

$$\begin{aligned} \mathrm{R}(\tilde{L}_2) &= \{\tilde{L}_2(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathrm{R}(L_1)\} \\ &= \{L_2(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathrm{R}(L_1)\} \end{aligned}$$

이고

$$\begin{aligned}\mathrm{R}(L_2 \circ L_1) &= \{(L_2 \circ L_1)(\boldsymbol{u}) \mid \boldsymbol{u} \in U\} \\ &= \{L_2(\boldsymbol{v}) \mid \boldsymbol{v} = L_1(\boldsymbol{u}),\ \boldsymbol{u} \in U\} \\ &= \{L_2(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathrm{R}(L_1)\}\end{aligned}$$

이므로 $\mathrm{R}(\tilde{L}_2) = \mathrm{R}(L_2 \circ L_1)$ 이다. 따라서 $\dim \mathrm{R}(\tilde{L}_2) = \dim \mathrm{R}(L_2 \circ L_1) = \dim \mathrm{R}(L_1)$ 이 성립한다. □

참고. 위 증명에서 $\mathrm{R}(L_2 \circ L_1)$ 의 기저와 $\mathrm{R}(L_1)$ 의 기저의 개수가 같음을 보이면 되므로 $\mathrm{R}(L_1)$ 의 기저 각각에 L_2 를 취한 벡터들이 다시 $\mathrm{R}(L_2 \circ L_1)$ 의 기저가 됨을 보여도 된다.

정의 7.21. $m \times n$ 행렬 (matrix) 이란, m 개의 행과 n 개의 열로 이루어진 원소들의 나열이다:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

두 행렬의 곱은 각각의 행렬에 대응되는 선형변환의 합성에 대응된다. 즉, 어떤 선형변환 $L_A : \mathbb{R}^l \to \mathbb{R}^m$ 이 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 로 대응시키고 $L_B : \mathbb{R}^m \to \mathbb{R}^n$ 이 $A\boldsymbol{x} \mapsto B(A\boldsymbol{x}) = BA\boldsymbol{x}$ 로 대응시킨다면, BA 는 $L_B \circ L_A$ 에 대응된다.

$n \times n$ 행렬 A, B 에 대해서 $AB = BA = I$ 이라면, $A = B^{-1}$ 이고 $B = A^{-1}$ 이다. 이때,

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

을 말한다.

정의 7.22. 선형변환 L_A 에 대응되는 행렬 A 에 대해서, 열공간 (column space) 은 A 의 열벡터들의 생성이다. 이때 열공간의 차원을 열계수 (column rank) 라고 부른다. 마찬가지로, 행공간 (row space) 은 A 의 행벡터들의 생성이다. 이때 행공간의 차원을 행계수 (row rank) 라고 부른다.

예시. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ 에서 열벡터 $(1, 4, 7)$ 과 $(3, 6, 9)$ 는 일차독립이고, $(2, 5, 8)$ 은 두 벡터의 합의 절반이므로 일차종속이다. 따라서 열계수는 2이다. 또한 행벡터

$(1,2,3)$과 $(7,8,9)$는 일차독립이고, $(4,5,6)$은 두 벡터의 합의 절반이므로 일차종속이다. 따라서 행계수도 2이다.

정리 7.23. *선형변환 L_A에 대응되는 행렬 A에 대해서, 열공간은 $\mathrm{R}(L_A)$와 동일하다. 따라서 열계수는 L_A의 계수 $\dim \mathrm{R}(L_A) = \operatorname{rank} L_A$와 같다.*

정리 7.24. *선형변환 L_A에 대응되는 행렬 A에 대해서, 열계수와 행계수는 동일하다. 즉, 열계수와 행계수 모두 $\operatorname{rank} L_A$로 나타내어지며 바로 행렬 A를 사용해 $\operatorname{rank} A$라고 쓸 수 있다.*

두 행을 교환하거나, 어떤 행의 상수배를 다른 행에 더하거나, 한 행에 0이 아닌 상수를 곱하는 연산을 행렬의 기본 행 연산이라고 한다. 마찬가지로 열에 대해서도 기본 열 연산을 정의할 수 있다.

정리 7.25. *두 행렬 A와 B의 곱을 C라고 하자. 행렬 A에 기본 행 연산을 취한 새로운 행렬을 $\tilde{A}$라고 하자. 행렬 C에 동일한 기본 행 연산을 취한 새로운 행렬을 $\tilde{C}$라고 하면, 여전히 $\tilde{A}B = \tilde{C}$가 성립한다.*

정의 7.26. 항등행렬 I에 기본 행/열 연산을 한 번한 것을 기본 행렬이라고 부른다.

예시. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$은 모두 계수가 3인 기본 행렬들이다.

정리 7.27. *$n \times n$ 기본 행렬의 계수는 n이다.*

증명. $n \times n$ 항등행렬 I에 대해서, 각 열벡터를 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$이라고 하자. 이때 열벡터들은 서로 독립이므로 $\operatorname{rank} I = n$이다.

I에 어떤 기본 열 연산을 하여 기본 행렬 A를 만들었다고 하자. 기본 열 연산이 1. 열의 교환인 경우, 2. 열의 상수배를 하여 다른 열에 더한 경우, 3. 열에 0이 아닌 상수를 곱한 경우에 대해서 생각하자.

1. 열을 교환했을 경우
   i 번째 열과 j 번째 열을 교환했다고 하자. 그렇다면 A의 열벡터들은 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_j(i \text{ 번째}), \ldots, \boldsymbol{e}_i(j \text{ 번째}), \ldots, \boldsymbol{e}_n$이 된다. 그런데 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$은 순서와 상관없이 일차독립이므로 A의 열벡터들도 일차독립이다.

2. 열의 상수배를 하여 다른 열에 더한 경우
   i 번째 열에 k배를 하여 j 열에 더했다고 하자. ($k \neq 0$이다.) 그렇다면 A의 열벡터들은 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_j + k\boldsymbol{e}_i, \ldots, \boldsymbol{e}_n$이 된다. 다음의 선형결합을 생각하자:

$$c_1\boldsymbol{e}_1 + \cdots + c_j(\boldsymbol{e}_j + k\boldsymbol{e}_1) + \cdots + c_n\boldsymbol{e}_n = \boldsymbol{0}$$

만약 c_j가 0이 아니라면, 다음과 같이 식을 재작성할 수 있다:

$$\boldsymbol{e}_j = -\frac{1}{c_j}\left(c_1\boldsymbol{e}_1 + \cdots + c_{j-1}\boldsymbol{e}_{j-1} + c_{j+1}\boldsymbol{e}_{j+1} + \cdots + c_n\boldsymbol{e}_n\right) + \frac{k}{c_j}\boldsymbol{e}_1$$

이때 $\frac{k}{c_j} \neq 0$이므로 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$이 독립이라는 것에 모순이다. 따라서 $c_j = 0$이다. 결국,

$$c_1\boldsymbol{e}_1 + \cdots + c_{j-1}\boldsymbol{e}_{j-1} + c_{j+1}\boldsymbol{e}_{j+1} + \cdots + c_n\boldsymbol{e}_n = \mathbf{0}$$

가 되어, $c_1 = \cdots = c_{j-1} = c_{j+1} = \cdots = c_n = 0$이다. 그런데 $c_j = 0$이므로 $c_1 = \cdots = c_n = 0$이어서, A의 열벡터들도 일차독립이다.

3. 열에 0이 아닌 상수를 곱한 경우
   i 번째 열에 $k \neq 0$을 곱했다고 하자. 그렇다면 A의 열벡터들은 $\boldsymbol{e}_1, \ldots, k\boldsymbol{e}_i, \boldsymbol{e}_n$이 된다. 다음의 선형결합을 생각하자:

$$c_1\boldsymbol{e}_1 + \cdots + c_i(k\boldsymbol{e}_i) + \cdots + c_n\boldsymbol{e}_n = \mathbf{0}$$

   $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$의 독립성에 의해 $c_1 = \cdots = c_ik = \cdots = c_n = 0$이 된다. 그런데 $k \neq 0$이므로 $c_1 = \cdots = c_n = 0$이다. 따라서 A의 열벡터들도 일차독립이다.

따라서, A의 열공간의 차원은 n이고 정리 7.24 의해 $\operatorname{rank} A = n$이다. 마찬가지로 기본 행 연산의 경우에도 모든 과정을 행벡터에 대입하여 진행하면 동일하게 계수가 n임을 알 수 있다. □

정리 7.28. *어떤 유한생성된 벡터공간 V와 부분공간 W가 주어졌다고 하자. 그렇다면, 다음이 성립한다:*

$$\dim W = \dim V \Rightarrow W = V$$

증명. $\dim W = \dim V = n$일 때, 벡터공간 W의 기저 $\mathcal{B} = \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$를 구성하자. $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$은 일차독립이고 V의 원소이기도 하므로, 기저 확장을 통해 V의 기저를 구성할 수 있다. 그런데 이미 $|\mathcal{B}| = \dim V$이므로, $\mathcal{B}$는 V의 기저이다. 따라서 $\operatorname{Span}\mathcal{B} = W = V$이다. □

정리 7.29. *선형변환 $L : V \to V$에 대해서, 다음의 명제는 동치이다:*

1. *L는 일대일 함수이다.*

2. *L는 위로의 함수이다.*

3. $\operatorname{rank} L = \dim V$ 이다.

증명. 정리 7.17에서 L가 일대일 함수라는 것은 $\dim \mathrm{R}(L) = \dim V$, 즉 $\operatorname{rank} L = \dim V$라는 것과 동치이다. 따라서 1과 3은 동치이다.

이제 2와 3이 동치임을 보이자. 만약 L가 위로의 함수이면, 공역과 치역이 동일하다. 즉, $\mathrm{R}(L) = V$다. 따라서 $\operatorname{rank} L = \dim \mathrm{R}(L) = \dim V$이 성립한다.

역으로, $\operatorname{rank} L = \dim V$라고 하자. $\dim \mathrm{R}(L) = \dim V$이므로, 정리 7.28에 따라 $\mathrm{R}(L) = V$이다. 그러므로 L는 위로의 함수이다.

따라서, 1, 2, 3은 모두 동치이다. □

2018.9.17. 지금까지의 내용을 간략히 요약하고, 약간의 내용을 덧붙인다. 체란 간단히 가감승제에 대해서 닫힌 집합이고, 벡터공간은 덧셈과 스칼라 곱에 대해 닫힌 집합이다. 또한 일차결합, 일차독립, 일차종속과 기저 등에 대해서 다루었다.

정리에는 기저의 원소의 개수가 항상 일정하다는 정리 7.5), 벡터공간의 교집합도 벡터공간이라는 것을 보여주는 정리 7.7, 벡터공간의 합 또한 벡터공간이며 (정리 7.9) 그 차원은 각각의 차원의 합에서 교집합의 차원을 제한 것이라는 정리 7.10이 있었다. 어떤 선형사상의 영공간과 치역은 각각 정의역과 공역의 부분공간이며 (정리 7.13), 그 차원의 합은 정의역의 차원임을 알려주는 정리 7.14가 있었다. 또한 행렬의 행계수, 열계수, 그리고 대응되는 선형변환의 계수가 모두 같다는 정리 7.24가 있었다.

정리 7.30. *벡터공간 V와 W 간의 선형변환 $L : V \to W$이 일대일 함수라는 것은 임의의 일차독립인 벡터들 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in V$을 골랐을 때, $L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_k)$ 또한 일차독립이라는 것과 동치이다.*

증명. 먼저 L이 일대일 함수라고 가정하자. 임의의 일차독립인 벡터들 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$에 선형변환 L을 취한 벡터들의 선형결합

$$c_1 L(\boldsymbol{v}_1) + \cdots + c_n L(\boldsymbol{v}_n) = \mathbf{0}$$

를 생각하자. 그렇다면,

$$L(c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n) = \mathbf{0}$$

인데, 정리 7.16에 따라 $\mathrm{N}(L) = \{\mathbf{0}\}$이므로

$$c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n = \mathbf{0}$$

임을 알 수 있다. 그런데 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$은 일차독립이므로 $c_1 = \cdots = c_n = 0$이다. 따라서 $L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_k)$는 일차독립이다.

이제 임의의 일차독립인 벡터들 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$에 대해서 $L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_k)$가 일차독립이라고 가정하자. $L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2)$인 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$를 고르자. 그런데 $L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2)$이므로, 가정한 명제의 대우에 따라서 $\boldsymbol{v}_1$과 $\boldsymbol{v}_2$는 일차종속이다. 따라서 일반성을 잃지 않고

$$\boldsymbol{v}_1 = k\boldsymbol{v}_2$$

라고 놓을 수 있다. 그러면

$$L(\boldsymbol{v}_2) = L(\boldsymbol{v}_1) = L(k\boldsymbol{v}_2) = kL(\boldsymbol{v}_2)$$

이 되어

$$(k-1)L(\boldsymbol{v}_2) = \mathbf{0}$$

이다. 만약 $k = 1$이면 $\boldsymbol{v}_1 = \boldsymbol{v}_2$가 되어 L이 일대일 함수이다. 만약 $k \neq 1$일 경우, $L(\boldsymbol{v}_2) = \mathbf{0}$이며 $L(\boldsymbol{v}_1) = \mathbf{0}$가 된다. $\boldsymbol{v}_1 \neq \mathbf{0}$인 경우, $L(\boldsymbol{v}_1) = \mathbf{0}$가 되어 일차독립인 벡터들에 대해 함숫값이 일차독립이라는 가정에 모순이다.[5] 따라서 $\boldsymbol{v}_1 = \boldsymbol{v}_2 = \mathbf{0}$이어야 한다. 즉 $L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2)$인 벡터에 대해서 항상 $\boldsymbol{v}_1 = \boldsymbol{v}_2$이므로 L은 일대일 함수이다. □

정리 7.31. *체 F에 대한 벡터공간 V가 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$에 의해 생성될 때, $\dim V \leq n$이다. 만약 $\dim V = n$이라면 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$은 기저를 이룬다.*[6]

증명. $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$의 부분집합 중 일차독립인 최대인 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$을 고르고, $\mathcal{B}$라고 쓰자. (즉 $r \leq n$이며, 원소의 순서는 재배열되었을 수 있다.) V의 임의의 원소 $\boldsymbol{v}$에 대해서, $V = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$이므로

$$\exists c_1, \ldots c_n \in F \quad \boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$$

이다. 그런데 $\boldsymbol{v}_{r+1}, \ldots$은 $\mathcal{B}$와 일차종속이므로 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$로 다시 쓸 수 있다. (일차독립이라면 해당 벡터를 $\mathcal{B}$에 추가하여도 여전히 일차독립이어야 하는데, 그러면 $\mathcal{B}$가 최대인 일차독립 부분집합이라는 가정에 모순된다.) 따라서,

$$\begin{aligned}
\boldsymbol{v} &= c_1\boldsymbol{v}_1 + \cdots + c_r\boldsymbol{v}_r + c_{r+1}\boldsymbol{v}_{r+1} + \ldots \\
&= c_1\boldsymbol{v}_1 + \cdots + c_r\boldsymbol{v}_r + c_{r+1}\sum_{i=1}^{r}\alpha_i\boldsymbol{v}_i + \ldots \\
&= \big(c_1 + \alpha_1(c_{r+1} + \ldots)\big)\boldsymbol{v}_1 + \cdots + \big(c_r + \alpha_r(c_{r+1} + \ldots)\big)\boldsymbol{v}_r
\end{aligned}$$

로 표현되어 $\operatorname{Span}\mathcal{B} = V$이다. 결국 $\mathcal{B}$는 V의 기저이며, $\dim V = r \leq n$이다.

---

[5] $\{\boldsymbol{v}_1\} \neq \{\mathbf{0}\}$는 일차독립인 집합이고, $\{\mathbf{0}\}$는 일차종속이기 때문이다.

[6] 강의 내용에 없었으나 본인의 필요에 따라 추가하였다.

이렇게 구성한 $\mathcal{B}$의 크기가 n이라고 하자. $\mathcal{B} \subseteq \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$인데 $|\mathcal{B}| = \big|\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}\big|$이므로, $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$이다. 따라서, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$은 기저를 이룬다. □

참고. 지금까지의 내용에 따라서, 벡터공간 V와 W간의 선형변환 L에 대해서 다음의 명제들은 서로 동치이다:

1. L은 일대일 함수이다.
2. $\mathrm{N}(L) = \{\mathbf{0}\}$이다. (정리 7.16)
3. $\operatorname{rank} L = \dim V$이다. (정리 7.17)
4. 임의의 선형독립인 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k \in V$에 대해서, $L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n)$도 선형독립이다. (정리 7.30)

만약 $V = W$이고, V가 유한생성된 벡터공간이라면, 다음의 명제들도 위 명제들과 동치이다:

5. L은 위로의 함수이다. (정리 7.29)
6. L은 일대일 대응이다. (1과 5로부터)
7. L은 가역 (invertible)이다. (정리 7.18)

나아가 $V = W = \mathbb{R}^n$일 경우, 선형변환 L에 대응되는 $n \times n$ 행렬 A가 있어서 다음의 명제들도 위 명제들과 동치이다:

8. $\forall \boldsymbol{v} \in \mathbb{R}^n \quad L^{-1}(\boldsymbol{v}) = A^{-1}\boldsymbol{v}$이다. (7로부터)
9. A의 열벡터들은 일차독립이다. (정리 7.24와 7.31)

2018.9.19. **정리 7.32.** *벡터 공간 U, V, W에 대해 선형변환 $L_1 : U \to V$와 $L_2 : V \to W$이 주어졌다고 하자. 만약 L_1가 위로의 함수라면, 다음이 성립한다:*

$$\operatorname{rank}(L_2 \circ L_1) = \operatorname{rank} L_2$$

증명. $\mathrm{R}(L_2 \circ L_1)$은 $L_2 \circ L_1$에 대한 U의 상이다. 즉,

$$\begin{aligned}\mathrm{R}(L_2 \circ L_1) &= \{(L_2 \circ L_1)(\boldsymbol{u}) \mid \boldsymbol{u} \in U\} \\ &= \{L_2(\boldsymbol{v}) \mid \boldsymbol{v} = L_1(\boldsymbol{u}),\ \boldsymbol{u} \in U\}\end{aligned}$$

이다. 그런데 L_1은 위로의 함수이므로

$$\forall \boldsymbol{v} \in V \quad \exists \boldsymbol{u} \in U \quad L_1(\boldsymbol{u}) = \boldsymbol{v}$$

가 성립한다. 따라서

$$\{L_2(\boldsymbol{v}) \mid \boldsymbol{v} = L_1(\boldsymbol{u}),\ \boldsymbol{u} \in U\} = \{L_2(\boldsymbol{v}) \mid \boldsymbol{v} \in V\} = \mathrm{R}(L_2)$$

이다. 따라서 $\mathrm{R}(L_2 \circ L_1) = \mathrm{R}(L_2)$이며, $\operatorname{rank}(L_2 \circ L_1) = \operatorname{rank} L_2$이 성립함을 알 수 있다. □

정리 7.33. *$m \times n$ 행렬 A의 계수는 기본 행/열 연산에 대해 불변이다.*

증명. 기본 행/열 연산에 대응되는 기본 행렬을 E라고 하고, A에 기본 행/열 연산을 취한 행렬을 $\tilde{A}$라고 하면,

$$EA = \tilde{A}$$

이다. 이때 E는 $m \times n$ 행렬을 $m \times n$ 행렬로 보내는 $m \times m$ 행렬이다. 각 행렬 $E, A, \tilde{A}$에 대응되는 행렬을 $L_E, L_A, L_{\tilde{A}}$라고 하자. 이때 $\tilde{A} = EA$이므로 $L_{\tilde{A}} = L_E \circ L_A$이다. 정리 7.23와 7.27에 따라 $\operatorname{rank} E = \operatorname{rank} L_E = m$이고, 따라서 정리 7.16에 따라 L_E는 일대일이다. 그러므로 정리 7.20에 의해

$$\operatorname{rank} L_{\tilde{A}} = \operatorname{rank}(L_E \circ L_A) = \operatorname{rank} L_A$$

임을 알 수 있다. 따라서 정리 7.23에 의해 $\operatorname{rank} A = \operatorname{rank} \tilde{A}$이다. □

정리 7.34. *$m \times n$ 행렬 A의 행계수는 기본 행/열 연산에 대해 불변이다.*

증명. 정리 7.24에 의해 자명하다.

혹은, 좀 더 직접적으로 정리 7.33에서 각 행렬에 전치를 취한 후, 기본 행 연산은 기본 열 연산으로, 기본 열 연산은 기본 행 연산으로 바꾼다면 행계수를 열계수로 바꾸어 생각할 수 있다. □

다음의 연립방정식

$$\begin{aligned} x + 2y + 3z &= 6 \\ 4x + 5y + 6z &= 15 \\ 7x + 8y + 9z &= 24 \end{aligned}$$

은 다음과 같이 행렬로 표현할 수 있다.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix}$$

정리 7.25에 의해 $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ 와 $\begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix}$ 에 동일한 기본 행 연산을 하여도 등식은 성립한다. 따라서

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ -9 \\ 0 \end{pmatrix}$$

이 된다. 위와 같은 형태를 행사다리꼴(row echelon form)이라고 한다. 여기서 z는 유일하게 정해지지 않으므로 $z = a$라고 놓으면, 후진 대입법을 통해서 y와 x를 구할 수 있다:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ 3-2a \\ a \end{pmatrix} = a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}.$$

혹은 여기서 그치지 않고 다음과 같이 기본 행 연산을 더 할 수도 있다:

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

위와 같은 행태를 기약행 사다리꼴(reduced row echelon form)이라고 한다. 여기서 바로 $z = a$로 둔다면, 쉽게 위와 동일한 해를 구할 수 있다.[7] 따라서 해집합은

$$\left\{ a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\}$$

임을 알 수 있다.

정리 7.35. *계수 행렬 A와 변수 열벡터 $\boldsymbol{x}$, 그리고 상수 열벡터 $\boldsymbol{b}$가 주어졌을 때, 방정식 $A\boldsymbol{x} = \boldsymbol{b}$의 해집합은*

$$\mathrm{N}(L_A) + \boldsymbol{x}_0$$

이때, L_A는 행렬 A에 대응되는 선형변환이고, $\boldsymbol{x}_0$는 특수해(particular solution)이다.

[7] 사실 동일한 표현이 나오는 것은 특수한 경우지만, 결과적으로는 동일한 해집합을 구하게 된다.

증명. $\boldsymbol{x}$를 임의의 해, $\boldsymbol{x}_0$를 특수해라고 하자. 즉,

$$A\boldsymbol{x} = \boldsymbol{b}$$
$$A\boldsymbol{x}_0 = \boldsymbol{b}$$

이다. 따라서

$$A(\boldsymbol{x} - \boldsymbol{x}_0) = 0$$

과 같이 정리할 수 있고, 정의에 따라 $\boldsymbol{x} - \boldsymbol{x}_0 \in \mathrm{N}(L_A)$임을 알 수 있다. 즉 $\boldsymbol{x} \in \mathrm{N}(L_A) + \boldsymbol{x}_0$이고, 해집합은 $\mathrm{N}(L_A) + \boldsymbol{x}_0$의 부분집합이다.

이제 $\mathrm{N}(L_A) + \boldsymbol{x}_0$의 한 원소 $\boldsymbol{x}$를 고르자. 즉, $\mathrm{N}(L_A)$에 속하는 한 원소 $\boldsymbol{x}_h$에 대해서

$$\boldsymbol{x} = \boldsymbol{x}_h + \boldsymbol{x}_0$$

를 만족한다. 양변에 행렬 A를 왼쪽에 곱해주면

$$A\boldsymbol{x} = A(\boldsymbol{x}_h + \boldsymbol{x}_0) = A\boldsymbol{x}_h + A\boldsymbol{x}_0$$

이다. 그런데 $\boldsymbol{x}_h \in \mathrm{N}(L_A)$이므로 $A\boldsymbol{x}_h = \mathbf{0}$이고, $\boldsymbol{x}_0$는 특수해이므로 $A\boldsymbol{x}_0 = \boldsymbol{b}$이다. 따라서

$$A\boldsymbol{x} = \mathbf{0} + \boldsymbol{b}$$

이어서, $\boldsymbol{x}$는 해집합에 속한다. 따라서 $\mathrm{N}(L_A) + \boldsymbol{x}_0$는 해집합의 부분집합이며, 위에서 해집합 또한 $\mathrm{N}(L_A) + \boldsymbol{x}_0$의 부분집합임을 보였으므로 둘은 같은 집합이다. □

예시. 방정식 $2x + y = 3$, 혹은

$$\begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3$$

의 해는 방정식 $2x + y = 0$의 해를 좌표 평면에 나타냈을 때 그려지는 직선을 평행 이동한 것이다. 특히 특수해 $(0, 3)$, $\left(\frac{3}{2}, 0\right)$, $(1, 1)$ 등에서 하나를 골라 원래 그래프가 해당 점을 지나도록 평행 이동한 것이다.

정리 7.35에서 상수 열벡터 $\boldsymbol{b} = \mathbf{0}$인 경우의 방정식을 동차 연립 일차 방정식 (homogeneous system of linear equations), $\boldsymbol{b} \neq \mathbf{0}$인 경우를 비동차 연립 일차 방정 (nonhomogeneous system of linear equations) 라고 한다. 다음의 동립 연

립 일차 방정식

$$\begin{pmatrix} 1 & 2 & 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

을 풀면,

$$(x_1, x_2, x_3, x_4, x_5, x_6) = a(2, 0, -2, -1, -1, 1) + b(-2, 1, 0, 0, 0, 0) \quad a, b \in \mathbb{R}$$

이 해가 된다. 이때 a와 b는 자유 변수이다. 따라서, 위 계수 행렬의 영공간은

$$\langle (2, 0, -2, -1, -1, 1), (-2, 1, 0, 0, 0, 0) \rangle$$

임을 알 수 있다. 여기서 주목할 사실은 변수 벡터의 원소 개수가 6, 즉 차원이 6인 벡터공간에 있다는 것이고, 해공간 혹은 영공간의 차원[8]이 2라는 것이다. 이 둘의 차이 $6 - 2 = 4$는 계수 행렬의 계수(rank)와 같은데, 이는 우연이 아니라 계수–퇴화 차수 정리(정리 7.14)에 의한 것이다.

정의 7.36. 어떤 $n \times n$ 행렬 A에 대해서 $n \times n$ 행렬 B가 존재해서

$$AB = I$$

일 때, A를 가역 행렬이라 하고 B를 A의 역행렬이라고 한다.

가역 행렬 A의 역행렬 B를 찾는 것은 식 $AB = I$에서 A를 I로 만드는 일련의 기본 행연산들을 A와 I에 동일하게 취하는 것이다. I에 해당 기본 행연산들을 취한 결과를 C라고 하면, 식 $AB = I$는 $IB = C$가 된다. 따라서 C는 A의 역행렬 B가 됨을 알 수 있다.

2018.10.1. 이렇게 역행렬을 구할 수 있으려면 기본 행연산들을 통해 A를 I로 만들 수 있어야 한다. 그런데 정리 7.34에 따르면 기본 행연산으로 행계수가 바뀌지 않으므로, A의 행계수가 A의 행/열의 개수와 같지 않으면 I로 변형할 수 없다. 따라서, A의 역행렬이 존재할 조건은 A의 계수가 행/열의 개수와 같은 것이다.

정의 7.37. 벡터공간 $V_1, \ldots, V_n, W$에 대해 어떤 함수 $f : V_1 \times \cdots \times V_n \to W$가 다중선형사상(multilinear map)이라는 것은 모든 $i \in \{1, \ldots, n\}$에 대한 $\boldsymbol{v}_i \in V_i$

[8]또는 자유변수의 개수, 즉 자유도

에 대해

$$f(\boldsymbol{v}_1, \ldots, c\boldsymbol{v}_i + \tilde{c}\tilde{\boldsymbol{v}}_i, \ldots, \boldsymbol{v}_n) = cf(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_n) + \tilde{c}f(\boldsymbol{v}_1, \ldots, \tilde{\boldsymbol{v}}_i, \ldots, \boldsymbol{v}_n)$$

을 만족한다는 것이다. 즉, 다중선형사상은 각각의 $\boldsymbol{v}_i$에 대해 f가 선형사상이다. 특히 $V_1 = \cdots = V_n$일 때, f를 n-형식(n-form)이라고 부른다.

정의 7.38. 벡터공간 V와 W에 대한 다중선형사상 $f : V^n \to W$가

$$\left(\exists i, j \in \{1, \ldots, n\} \quad \boldsymbol{v}_i = \boldsymbol{v}_j\right) \quad \Rightarrow \quad f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = 0$$

를 만족하면 f를 교대다중선형사상(alternating multilinear map)이라고 부른다.

정리 7.39. *벡터공간 V와 W에 대한 교대다중선형사상 $f : V^n \to W$은 두 벡터를 교환하면 부호가 바뀐다. 즉,*

$$f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_n) = -f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_n)$$

이다.

증명. 교대다중선형사상에서 두 벡터가 동일하면 그 함숫값은 0이므로,

$$f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i + \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i + \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) = 0$$

이다. 이는 다중선형성에 의해

$$\begin{aligned}
&f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i + \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i + \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) \\
&= f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_i, \ldots \boldsymbol{v}_n) + f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) \\
&\quad + f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) + f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i, \ldots \boldsymbol{v}_n) \\
&= f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) + f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i, \ldots \boldsymbol{v}_n) \\
&= 0
\end{aligned}$$

이다. 따라서,

$$f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_j, \ldots \boldsymbol{v}_n) = -f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i, \ldots \boldsymbol{v}_n)$$

를 만족한다. □

예시. 교대다중선형사상 $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$에 대해서, f가 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$의 열벡터를 인자로 받는다고 하자. 이를 계산하면

$$\begin{aligned} f\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= f(a\boldsymbol{e}_1 + c\boldsymbol{e}_2, b\boldsymbol{e}_1 + d\boldsymbol{e}_2) \\ &= abf(\boldsymbol{e}_1, \boldsymbol{e}_1) + adf(\boldsymbol{e}_1, \boldsymbol{e}_2) + cbf(\boldsymbol{e}_2, \boldsymbol{e}_1) + cdf(\boldsymbol{e}_2, \boldsymbol{e}_2) \\ &= (ad - bc)f(\boldsymbol{e}_1, \boldsymbol{e}_2) \\ &= (ad - bc)f\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

정의 7.40. 교대다중선형사상 $f : (\mathbb{R}^n)^n \to \mathbb{R}$에 대해서 $n \times n$ 단위행렬 I의 함숫값이 1, 즉 $f(I) = 1$이면 f를 행렬식(determinant)이라고 하고, det라고 표기한다.

행렬식을 계산할 때는 다중선형성(정의 7.37) 및 교대다중선형성(정의 7.38와 정리 7.39)에 따라 각 열벡터들의 원소들을 행이 겹치지 않도록 골라서 계산할 수 있다. 예를 들어

$$\begin{aligned} \det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} &= aej\det(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) + bfg\det(\boldsymbol{e}_3, \boldsymbol{e}_1, \boldsymbol{e}_2) + cdh\det(\boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_1) \\ &\quad + bdj\det(\boldsymbol{e}_2, \boldsymbol{e}_1, \boldsymbol{e}_3) + ceg\det(\boldsymbol{e}_3, \boldsymbol{e}_2, \boldsymbol{e}_1) + afh\det(\boldsymbol{e}_1, \boldsymbol{e}_3, \boldsymbol{e}_2) \\ &= \big((aej + bfg + cdh) - (bdj + ceg + afh)\big)\det I \end{aligned}$$

와 같이 계산하면 된다.

도움정리 7.41. $\mathrm{Sym}\big(\{1, 2, \ldots, n\}\big)$ *의 항등 순열(permutation)* $(1\ 2\ \ldots\ n)$ *의 원소 두 개씩을 서로 교환(transposition) 하여 다시 항등 순열을 만들 때, 교환 횟수는 항상 짝수번이다.*

정리 7.42. $\mathrm{Sym}\big(\{1, 2, \ldots, n\}\big)$ *의 순열* σ_1 *에서* σ_2 *를 만드는 교환은 홀짝성을 보존한다.*

2018.10.5. **정리 7.43.** $\mathbb{R}^n$ *에 속하는 벡터* $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ *와* $n \times n$ *행렬* A, B *에 대해서, 다음이 성립한다:*

1. $\det(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_n) = \det\big(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i + k\boldsymbol{v}_j, \ldots, \boldsymbol{v}_n\big)$
2. $\det\big(A^t\big) = \det(A)$

3. $\det(AB) = \det(A)\det(B)$ [9]

정리 7.44. *$n \times n$ 행렬 A의 행렬식에 대해서,*

1. *두 행을 바꾸는 기본 행 연산은 행렬식의 부호를 바꾸며*
2. *어떤 행에 k를 곱하는 기본 행 연산은 행렬식의 값에 k배를 하며*
3. *어떤 행의 상수 배를 하여 다른 행에 더하는 연산에 대해 불변이다.*[10]

참고. 행렬식을 계산할 때 다중선형성과 교대다중선형성만을 사용해 계산하면 긴 계산을 거쳐야 하지만, 위의 정리 7.44을 사용하면 계산이 단순해질 수 있다. 이는 어떤 행렬

$$A = \begin{pmatrix} a & e & f & g \\ 0 & b & h & i \\ 0 & 0 & c & j \\ 0 & 0 & 0 & d \end{pmatrix}$$

에 대해서 $\det A = abcd$이기 때문이다. 따라서, 정리 7.44에 따라 기본 행 연산을 통해 이와 같이 대각선 밑의 모든 원소가 0 (혹은 위의 모든 원소) 인 꼴을 만들면 대각선의 곱을 통해 행렬식을 계산할 수 있다. 이러한 꼴을 상삼각행렬 (혹은 하삼각행렬) 이라고 부른다.

일반적으로 행렬식 계산은 이렇게 상삼각행렬과 하삼각행렬의 곱으로 분해하는 과정 (LU 분해) 을 통해 행렬을 곱하는데 필요한 시간복잡도와 동일하게 계산할 수 있다.

정리 7.45. *$n \times n$ 행렬 A에 대응되는 선형변환 L_A가 일대일 함수라는 것은* $\det A$ *가 0이 아니라는 것과 동치이다.*

증명. 만약 행렬 A에서 어떤 행의 k배를 하여 다른 행에 더하는 기본 행 연산을 반복하였을 때 상 (하) 삼각행렬을 만들었다고 하자. 상삼각행렬은 다시 어떤 행의 k배를 하여 다른 행에 더하는 기본 행 연산을 반복하여 (후진대입을 하는 것과 같이) 대각선 이외의 모든 원소를 0으로 만들 수 있다. 이렇게 만들어진 행렬을 $\tilde{A}$라고 하자. 이때 정리 7.44에 따라 $\det A = \det \tilde{A}$이다.

$\det \tilde{A} = 0$, 즉 $\tilde{A}$의 대각선의 한 원소가 0이라고 가정하자. 이는 $\tilde{A}$의 열벡터들이 서로 독립이 아니라는 것과 동치이다. ($\tilde{A}$의 대각선 이외의 모든 원소들은 0이기 때문이다.) 정리 7.33에 따라 기본 행 연산을 통해 열벡터들의 독립성이 바뀌지 않으므로, A의 열벡터들도 서로 독립이 되지 않는다. 따라서 A에 대응되는 L_A는 일대일 함수가 되지 않는다 (19쪽의 참고 9). □

[9] 원래 수업에서는 누락되었던 부분이지만, 2018년 10월 24일 수업에 추가로 언급하신 내용이다.
[10] 정리 7.43의 2에 따라, 정리 7.44의 기본 열 연산 버전을 만들 수 있다.

참고. 19쪽의 참고를 다시 옮긴다. 벡터공간 V와 W간의 선형변환 L에 대해서 다음의 명제들은 서로 동치이다:

1. L은 일대일 함수이다.
2. $\mathrm{N}(L) = \{\mathbf{0}\}$이다. (정리 7.16)
3. $\operatorname{rank} L = \dim V$이다. (정리 7.17)
4. 임의의 선형독립인 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in V$에 대해서, $L(\boldsymbol{v}_1), \ldots, L(\boldsymbol{v}_n)$도 선형독립이다. (정리 7.30)

만약 $V = W$이고, V가 유한생성된 벡터공간이라면, 다음의 명제들도 위 명제들과 동치이다:

5. L은 위로의 함수이다. (정리 7.29)
6. L은 일대일 대응이다. (1과 5로부터)
7. L은 가역(invertible)이다. (정리 7.18)

나아가 $V = W = \mathbb{R}^n$일 경우, 선형변환 L에 대응되는 $n \times n$ 행렬 A가 있어서 다음의 명제들도 위 명제들과 동치이다:

8. $\forall \boldsymbol{v} \in \mathbb{R}^n \quad L^{-1}(\boldsymbol{v}) = A^{-1}\boldsymbol{v}$이다. (7로부터)
9. A의 열벡터들은 일차독립이다. (정리 7.24와 7.31)
10. $A\boldsymbol{x} = \mathbf{0}$인 $\boldsymbol{x}$가 $\mathbf{0}$로 유일하게 존재한다. (정리 7.16)
11. $A\boldsymbol{x} = \boldsymbol{b}$인 $\boldsymbol{x}$가 유일하게 존재한다. (10과 정리 7.35)
12. $\det A$가 0이 아니다. (정리 7.45)

정리 7.46. *$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$에 대해서 다음이 성립한다:*

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$$

증명. 다음과 같이 놓자:

$$\begin{aligned} \boldsymbol{a} &= (a_1, a_2, a_3) \\ \boldsymbol{b} &= (b_1, b_2, b_3) \\ \boldsymbol{c} &= (c_1, c_2, c_3) \end{aligned}$$

이때,

$$\boldsymbol{a} \times \boldsymbol{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$$

이며, 따라서

$$\begin{aligned}(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c} &= (a_2b_3-a_3b_2)c_1+(a_3b_1-a_1b_3)c_2+(a_1b_2-a_2b_1)c_3\\ &= \det(\boldsymbol{c},\boldsymbol{a},\boldsymbol{b})\\ &= \det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\end{aligned}$$

이다. 마찬가지로 $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})$도 성분별로 계산을 해보면 $\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$임을 알 수 있다. □

정리 7.47. *$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in\mathbb{R}^3$가 이루는 평행체의 부피 $\mathrm{Vol}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$는 $|\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})|$이다.*

증명. $\boldsymbol{a}$와 $\boldsymbol{b}$가 밑면을 이룬다고 하면, 그 면적은

$$|\boldsymbol{a}\times\boldsymbol{b}|$$

이다. 이때 $\boldsymbol{a}\times\boldsymbol{b}$와 $\boldsymbol{c}$가 θ의 각도를 이룬다고 하면, $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$가 이루는 평행체의 부피는

$$\mathrm{Vol}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) = |\boldsymbol{a}\times\boldsymbol{b}|\big||\boldsymbol{c}|\cos\theta\big| = |(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}|$$

이다. 그런데 정리 7.46에 따라 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$이므로,

$$\mathrm{Vol}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) = |\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})|$$

이 성립한다. □

정의 7.48. 벡터 $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\in\mathbb{R}^n$이 양향 (positively oriented) 라는 것은

$$\det(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n) > 0$$

이라는 것이고, 음향 (negatively oriented) 이라는 것은

$$\det(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n) < 0$$

이라는 것이다.

이때 $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n$에 의해 생성되는 평행체의 부호수를 가지는 부피 (signed volume) $\widetilde{\mathrm{Vol}}(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$는

$$\widetilde{\mathrm{Vol}}(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n) = \det(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$$

으로 정의된다.

참고. 부호수를 가지는 부피 $\widetilde{\mathrm{Vol}}$는 행렬식으로 정의되었으므로, 정리 7.43와 7.44가 그대로 적용된다:

1. $\widetilde{\mathrm{Vol}}(\boldsymbol{a}, k\boldsymbol{b}, \boldsymbol{c}) = k\,\widetilde{\mathrm{Vol}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$

2. $\widetilde{\mathrm{Vol}}(\boldsymbol{a}, \boldsymbol{b} + \tilde{\boldsymbol{b}}, \boldsymbol{c}) = \widetilde{\mathrm{Vol}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) + \widetilde{\mathrm{Vol}}(\boldsymbol{a}, \tilde{\boldsymbol{b}}, \boldsymbol{c})$

3. $\widetilde{\mathrm{Vol}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = -\,\widetilde{\mathrm{Vol}}(\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{c})$

정의 7.49. $n \times n$ 행렬 A에 대해서 i행과 j열을 지운 행렬 M_{ij}를 소행렬식(minor)이라고 부르고, $C_{ij} = (-1)^{i+j} M_{ij}$를 여인수(cofactor)이라고 한다.

정리 7.50. *$n \times n$ 행렬 $A = (a_{ij})$에 대해서, 라플라스 전개(Laplace expansion) 혹은 여인수 전개(cofactor expansion)는 다음과 같이 행렬식을 계산하는 것을 말한다:*

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik} = \sum_{k=1}^{n} a_{ik} C_{ik}$$

예시. 다음은 3 행렬의 2행을 중심으로 여인수 전개를 한 것이다:

$$\begin{aligned}
\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= (-1)^{2+1} d \det \begin{pmatrix} b & c \\ h & j \end{pmatrix} + (-1)^{2+2} e \det \begin{pmatrix} a & c \\ g & j \end{pmatrix} \\
&\quad + (-1)^{2+3} f \det \begin{pmatrix} a & b \\ g & h \end{pmatrix} \\
&= -d \det \begin{pmatrix} b & c \\ h & j \end{pmatrix} + e \det \begin{pmatrix} a & c \\ g & j \end{pmatrix} - f \det \begin{pmatrix} a & b \\ g & h \end{pmatrix}
\end{aligned}$$

정리 7.51. *행렬 A가 가역행렬일 때, 다음과 같이 역행렬을 구할 수 있다:*

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

이때 (C_{ij})는 여인수들의 행렬이다.

8 Linear Algebra: Matrix Eigenvalue Problems

정의 8.1. $\mathbb{R}^n$의 두 벡터 $\boldsymbol{a} = (a_1, \ldots, a_n)$과 $\boldsymbol{b} = (b_1, \ldots, b_n)$에 대해서 스칼라곱(scalar product), 점곱(dot product), 혹은 단순히 내적(inner product)은 다음을 말한다:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^{n} a_i b_i$$

정리 8.2. *$\mathbb{R}^n$의 벡터 $\boldsymbol{a}, \tilde{\boldsymbol{a}}, \boldsymbol{b}$와 실수 $c, \tilde{c}$에 대해서 다음이 성립한다:*

1. $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$ *(대칭성(symmetricity))*

2. $(c\boldsymbol{a} + \tilde{c}\tilde{\boldsymbol{a}}) \cdot \boldsymbol{b} = c\boldsymbol{a} \cdot \boldsymbol{b} + \tilde{c}\tilde{\boldsymbol{a}} \cdot \boldsymbol{b}$ *(쌍선형성(bilinearlity))*

3. $\boldsymbol{a} \cdot \boldsymbol{a} \geq 0$ *(등호 성립 조건은* $\boldsymbol{a} = \boldsymbol{0}$*) (양의 정부호성 positive definiteness)*

정의 8.3. 체 F에 대한 벡터공간 V에 속하는 원소 $\boldsymbol{a}, \tilde{\boldsymbol{a}}, \boldsymbol{b}$와 스칼라 $c, \tilde{c}$에 대해, 내적(inner product) $\langle \cdot, \cdot \rangle : V \times V \to F$은 다음을 말한다:

1. $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{b}, \boldsymbol{a} \rangle$ (대칭성(symmetricity))

2. $\langle c\boldsymbol{a} + \tilde{c}\tilde{\boldsymbol{a}}, \boldsymbol{b} \rangle = \langle c\boldsymbol{a}, \boldsymbol{b} \rangle + \langle \tilde{c}\tilde{\boldsymbol{a}}, \boldsymbol{b} \rangle$ (쌍선형성(bilinearlity))

3. $\langle \boldsymbol{a}, \boldsymbol{a} \rangle \geq 0$ (등호 성립 조건은 $\boldsymbol{a} = \boldsymbol{0}$) (양의 정부호성(positive definiteness))

또한 내적이 정의되는 벡터공간을 내적공간(inner product space)이라고 부른다.

예시.

1. $\langle (a, b), (c, d) \rangle = ab + cd$는 쌍선형성을 만족하지 않으므로 내적이 아니다.

2. $\langle (a, b), (c, d) \rangle = ad + bc$는 양의 정부호성을 만족하지 않으므로 내적이 아니다.

3. $\langle (a, b, c), (d, e, f) \rangle = ad + be + kcf$에서 $k = 3$일 경우 내적이고, $k = -1$일 경우 양의 정부호성을 만족하지 않으므로 내적이 아니다.

4. 실함수 $f, g \in \mathcal{C}^0$에 대해서 $\langle f, g \rangle_I = \int_I fg \, \mathrm{d}x$는 내적이다.

5. 실함수 $f, g \in \mathcal{C}^0$에 대해서 $\langle f, g \rangle = \int_{-1}^{1} e^x fg \, \mathrm{d}x$는 내적이다. 이때 e^x를 $[-1, 1]$에서 내적의 무게함수(weight function)이라고 한다.

6. 실함수 $f, g \in \mathcal{C}^0$에 대해서 $\langle f, g \rangle_I = \int_I rfg \, \mathrm{d}x$는 내적이다. 이때

$$\forall x \in I \quad r(x) \geq 0$$

이며 $r(x) = 0$이 되는 x들은 서로 떨어져 있다.

정리 8.4. *함수* $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$*가 대칭 쌍선형형식(symmetric bilinear form)이라는 것은 벡터* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ *에 대해 어떤* $n \times n$ *대칭행렬* A*가 존재해서 다음을 만족한다는 것과 동치이다:*

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{b}^t A \boldsymbol{a}$$

증명. 먼저 $\langle\cdot,\cdot\rangle$가 대칭 쌍선형형식이면 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$에 대해 $\langle\boldsymbol{a},\boldsymbol{b}\rangle=\boldsymbol{b}^tA\boldsymbol{a}$인 $n\times n$ 대칭행렬 A가 존재한다는 것을 보이자.

$\boldsymbol{a}=(a_1,\dots,a_n),\boldsymbol{b}=(b_1,\dots,b_n)$이라고 하자. 쌍선형성에 의해 $\langle\boldsymbol{a},\boldsymbol{b}\rangle$은 다음과 같이 정리할 수 있다:

$$\langle\boldsymbol{a},\boldsymbol{b}\rangle=\langle a_1\boldsymbol{e}_1+\cdots+a_n\boldsymbol{e}_n,b_1\boldsymbol{e}_1+\cdots+b_n\boldsymbol{e}_n\rangle=\sum_{i,j=1}^n a_ib_j\left\langle\boldsymbol{e}_i,\boldsymbol{e}_j\right\rangle$$

대칭성에 따라 $\left\langle\boldsymbol{e}_i,\boldsymbol{e}_j\right\rangle=\left\langle\boldsymbol{e}_j,\boldsymbol{e}_i\right\rangle$이며, 따라서 $A_{ij}=\left\langle\boldsymbol{e}_i,\boldsymbol{e}_j\right\rangle$일 때 $A=(A_{ij})=(A_{ji})$는 대칭행렬이다. 이때

$$\begin{aligned}\langle\boldsymbol{a},\boldsymbol{b}\rangle=\sum_{i,j=1}^n a_ib_jA_{ij}&=\sum_{j=1}^n b_j\left(\sum_{i=1}^n A_{ij}a_i\right)=\sum_{j=1}^n b_j\left(\sum_{i=1}^n A_{ji}a_i\right)\\&=\begin{pmatrix}b_1&\dots&b_n\end{pmatrix}\begin{pmatrix}\sum_{i=1}^n A_{1i}a_i\\\vdots\\\sum_{i=1}^n A_{ni}a_i\end{pmatrix}\\&=\begin{pmatrix}b_1&\dots&b_n\end{pmatrix}\begin{pmatrix}A_{11}&\dots&A_{1n}\\\vdots&\ddots&\vdots\\A_{n1}&\dots&A_{nn}\end{pmatrix}\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}\end{aligned}$$

이므로 대칭행렬 A에 대해 $\langle\boldsymbol{a},\boldsymbol{b}\rangle=\boldsymbol{b}^tA\boldsymbol{a}$임을 알 수 있다.

이제 대칭행렬 A에 대해 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$일 때 $\langle\boldsymbol{a},\boldsymbol{b}\rangle=\boldsymbol{b}^tA\boldsymbol{a}$이면 $\langle\cdot,\cdot\rangle$이 대칭 쌍선형형식임을 보이자.

$\boldsymbol{a}=(a_1,\dots,a_n),\boldsymbol{b}=(b_1,\dots,b_n)$이라고 하자. 그러면

$$\begin{aligned}\langle\boldsymbol{a},\boldsymbol{b}\rangle&=\boldsymbol{b}^tA\boldsymbol{a}\\&=\begin{pmatrix}b_1&\dots&b_n\end{pmatrix}\begin{pmatrix}A_{11}&\dots&A_{1n}\\\vdots&\ddots&\vdots\\A_{n1}&\dots&A_{nn}\end{pmatrix}\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}\\&=\sum_{i,j=1}^n a_ib_jA_{ij}=\sum_{i=1}^n a_i\left(\sum_{j=1}^n A_{ij}b_j\right)\\&=\begin{pmatrix}a_1&\dots&a_n\end{pmatrix}\begin{pmatrix}A_{11}&\dots&A_{1n}\\\vdots&\ddots&\vdots\\A_{n1}&\dots&A_{nn}\end{pmatrix}\begin{pmatrix}b_1\\\vdots\\b_n\end{pmatrix}\\&=\boldsymbol{a}^tA\boldsymbol{b}=\langle\boldsymbol{b},\boldsymbol{a}\rangle\end{aligned}$$

이므로 대칭성을 만족한다. 또한 벡터 $\tilde{\boldsymbol{a}} \in \mathbb{R}^n$ 을 $(\tilde{a}_1, \ldots, \tilde{a}_n)$ 으로 놓고 스칼라 $c, \tilde{c} \in \mathbb{R}$ 에 대해서 쌍선형성을 만족한다:

$$\langle c\boldsymbol{a} + \tilde{c}\tilde{\boldsymbol{a}}, \boldsymbol{b}\rangle = \boldsymbol{b}^t A(c\boldsymbol{a} + \tilde{c}\tilde{\boldsymbol{a}}) = c\boldsymbol{b}^t A\boldsymbol{a} + \tilde{c}\boldsymbol{b}^t A\tilde{\boldsymbol{a}}$$

따라서 $\langle \cdot, \cdot \rangle$ 이 쌍선형형식이라는 것과 $\langle \boldsymbol{a}, \boldsymbol{b}\rangle = \boldsymbol{b}^t A\boldsymbol{a}$ 인 대칭행렬 A 가 존재한다는 것은 동치이다. □

예시. 두 벡터 (a, b, c) 와 (d, e, f) 에 대한 내적 $ae + bd + 4cf$ 에 대해서,

$$\langle (a,b,c), (d,e,f)\rangle = ae + bd + 4cd = \begin{pmatrix} d & e & f \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

이다.

정의 8.5. 내적공간에 속하는 벡터 $\boldsymbol{v}$ 에 대해

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$$

를 $\boldsymbol{v}$ 의 노름(norm)이라고 한다.

정리 8.6 (코시-부냐콥스키-슈바르츠(Cauchy-Bunyakovsky-Schwarz) 부등식). *내적공간에 속하는 두 벡터 $\boldsymbol{a}$ 와 $\boldsymbol{b}$ 에 대해 다음 부등식이 성립한다:*

$$|\langle \boldsymbol{a}, \boldsymbol{b}\rangle| \leq \sqrt{\langle \boldsymbol{a}, \boldsymbol{a}\rangle}\sqrt{\langle \boldsymbol{b}, \boldsymbol{b}\rangle} = \|\boldsymbol{a}\|\|\boldsymbol{b}\|$$

등호 성립 조건은 $\boldsymbol{a}$ 와 $\boldsymbol{b}$ 가 일차종속일 경우이다. 위 부등식을 코시-부냐콥스키-슈바르츠(Cauchy-Bunyakovsky-Schwarz) 부등식 혹은 CBS 부등식이라고 한다.

증명. 벡터 $\boldsymbol{a}$ 와 $\boldsymbol{b}$ 에 대해 다음과 같은 함수 f 를 정의하자:

$$f(t) = \langle \boldsymbol{a} + t\boldsymbol{b}, \boldsymbol{a} + t\boldsymbol{b}\rangle$$

내적은 양의 정부호성을 가지므로 $f(t) \geq 0$ 임을 알 수 있다. 또한 내적의 쌍선형성과 대칭성에 따라 $f(t)$ 는 다음과 같이 정리할 수 있다:

$$f(t) = \langle \boldsymbol{b}, \boldsymbol{b}\rangle t^2 + 2\langle \boldsymbol{a}, \boldsymbol{b}\rangle t + \langle \boldsymbol{a}, \boldsymbol{a}\rangle$$

따라서 $f(t)$ 는 t 에 관한 이차함수이며, 항상 0 이상이므로 판별식이 $D/4 \leq 0$ 를 만족한다. 즉,

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle^2 \leq \langle \boldsymbol{a}, \boldsymbol{a}\rangle \langle \boldsymbol{b}, \boldsymbol{b}\rangle$$

이다. □

정의 8.7. 내적공간의 두 벡터 $\boldsymbol{a}$와 $\boldsymbol{b}$가 이루는 각도 θ를 다음과 같이 정의한다:

$$\cos\theta = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$$

이때 $-1 \leq \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} \leq 1$임은 정리 8.6의 CBS 부등식에 의해 보장된다.

정의 8.8. 내적공간에 속하는 두 벡터 $\boldsymbol{a}$와 $\boldsymbol{b}$가 서로 수직이라는 것은 $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$이라는 것이다. 나아가 $\boldsymbol{a}$와 $\boldsymbol{b}$가 단위 벡터이면 둘이 정규직교(orthonormal) 한다고 한다.

벡터들의 집합 S의 모든 원소들이 서로 직교하면 S를 직교 집합(orthogonal set), 서로 정규직교하면 S를 정규직교 집합(orthonormal set)이라고 부른다.

정리 8.9. *내적공간에 속하는 정규직교 집합은 일차독립이다.*

증명. 내적공간 V에 속하는 벡터들 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$의 집합이 정규직교 집합이라고 하자. 이때 $\boldsymbol{a}_i$들의 선형결합

$$c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n = 0$$

의 양변에 $i \in \{1, \ldots, n\}$인 $\boldsymbol{a}_i$를 내적하면

$$\langle (c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n), \boldsymbol{a}_i \rangle = \sum_{j \in \{1,\ldots,n\} \setminus \{i\}} c_j \langle \boldsymbol{a}_j, \boldsymbol{a}_i \rangle + c_i \langle \boldsymbol{a}_i, \boldsymbol{a}_i \rangle$$

인데, 모든 벡터들은 서로 직교하므로 $i \neq j$이면 $\langle \boldsymbol{a}_j, \boldsymbol{a}_i \rangle = 0$이다. 나아가 모든 벡터들은 단위 벡터이므로 $\langle \boldsymbol{a}_i, \boldsymbol{a}_i \rangle = 1$이다. 따라서

$$\langle (c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n), \boldsymbol{a}_i \rangle = c_i = 0$$

이며, $c_i = 0$이 된다. 임의의 $i \in \{1, \ldots, n\}$에 대해서 위 과정을 반복하면, 모든 i에 대해서 $c_i = 0$임을 알 수 있다. 따라서 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$은 일차독립이다. □

정의 8.10. 내적공간 V와 V의 부분공간 W에 대해서, $\boldsymbol{a} \in V$를 잡자. 이때 $\boldsymbol{a}$의 W 위로의 정사영 $\boldsymbol{v}$는

$$\forall \boldsymbol{w} \in W \qquad \langle \boldsymbol{a} - \boldsymbol{v}, \boldsymbol{w} \rangle = 0$$

을 만족한다. 이때 $\boldsymbol{v} = p_W(\boldsymbol{a})$로 표기한다.

$\boldsymbol{a} \in V$의 어떤 벡터 $\boldsymbol{u} \in V$가 생성하는 벡터공간 $\langle \boldsymbol{u} \rangle$로의 정사영을 $p_{\boldsymbol{u}}(\boldsymbol{a})$로도 표기한다.

정리 8.11. *내적공간 V 와 부분공간 W 에 대해서 W 의 직교 기저가 $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ 일 때, $\boldsymbol{v} \in V$ 의 W 위로의 정사영 $p_W(\boldsymbol{v})$ 는 다음과 같다:*

$$p_W(\boldsymbol{v}) = \frac{\langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle} + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{w}_n \rangle}{\langle \boldsymbol{w}_n, \boldsymbol{w}_n \rangle}$$

증명. $\boldsymbol{v} \in V$ 의 W 위로의 정사영 $p_W(\boldsymbol{v})$ 는 W 의 원소이므로, W 의 직교 기저 $\{\boldsymbol{w}_1, \ldots \boldsymbol{w}_n\}$ 의 선형결합으로 표현할 수 있다:

$$p_W(\boldsymbol{v}) = c_1 \boldsymbol{w}_1 + \cdots + c_n \boldsymbol{w}_n$$

그리고 정사영의 정의에 따라

$$\forall i \in \{1, \ldots, n\} \qquad \langle \boldsymbol{v} - p_W(\boldsymbol{v}), \boldsymbol{w}_i \rangle = 0$$

를 만족한다. 정리하면,

$$\begin{aligned}
\langle \boldsymbol{v} - p_W(\boldsymbol{v}), \boldsymbol{w}_i \rangle &= \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle p_W(\boldsymbol{v}), \boldsymbol{w}_i \rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle c_1 \boldsymbol{w}_1 + \cdots + c_n \boldsymbol{w}_n, \boldsymbol{w}_i \rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \sum_{j \in \{1, \ldots, n\} \setminus \{i\}} c_j \langle \boldsymbol{w}_j, \boldsymbol{w}_i \rangle - c_i \langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - c_i \langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle = 0
\end{aligned}$$

이므로

$$c_i = \frac{\langle \boldsymbol{v}, \boldsymbol{w}_i \rangle}{\langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle}$$

이다. 따라서 정사영 $p_W(\boldsymbol{v})$ 는

$$p_W(\boldsymbol{v}) = \frac{\langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle}{\langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle} + \cdots + \frac{\langle \boldsymbol{v}, \boldsymbol{w}_n \rangle}{\langle \boldsymbol{w}_n, \boldsymbol{w}_n \rangle}$$

이다. □

정리 8.12 (그람-슈미트 (Gram-Schmidt) 과정)**.** *내적공간 W 의 기저 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$*

에 대해서, 다음의 과정을 그람-슈미트(Gram-Schmidt) 과정이라고 부른다:

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{v}_1 \\
\boldsymbol{w}_2 &= \boldsymbol{v}_2 - p_{\boldsymbol{w}_1}(\boldsymbol{v}_2) \\
\boldsymbol{w}_3 &= \boldsymbol{v}_3 - p_{\boldsymbol{w}_1}(\boldsymbol{v}_3) - p_{\boldsymbol{w}_2}(\boldsymbol{v}_3) \\
&\vdots \\
\boldsymbol{w}_n &= \boldsymbol{v}_n - \sum_{i=1}^{n-1} p_{\boldsymbol{w}_i}(\boldsymbol{v}_n)
\end{aligned}$$

이렇게 구한 $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ 의 집합은 W 의 직교 기저이다.

예시.

1. $\mathbb{R}^3 = \langle (1,0,0), (0,1,1), (1,2,3) \rangle$ 와 스칼라곱으로 주어진 내적에 그람-슈미트 과정을 다음과 같이 실행한다:

$$\begin{aligned}
\boldsymbol{w}_1 &= (1,0,0) \\
\boldsymbol{w}_2 &= (0,1,1) - \frac{(1,0,0)\cdot(0,1,1)}{(1,0,0)\cdot(1,0,0)}(1,0,0) = (0,1,1) \\
\boldsymbol{w}_3 &= (1,2,3) - \frac{(1,0,0)\cdot(1,2,3)}{(1,0,0)\cdot(1,0,0)}(1,0,0) - \frac{(0,1,1)\cdot(1,2,3)}{(0,1,1)\cdot(0,1,1)}(0,1,1) \\
&= (1,2,3) - (1,0,0) - \frac{5}{2}(0,1,1) = \left(0, -\frac{1}{2}, \frac{1}{2}\right)
\end{aligned}$$

   따라서 $\mathbb{R}^3$는 직교 기저 $\left\{(1,0,0), (0,1,1), \left(0, -\frac{1}{2}, \frac{1}{2}\right)\right\}$로 생성된다.

2. 벡터공간 $W = \langle 1, x, x^2 \rangle$에 대해서 $I = [-1,1]$이고 $r(x) = 1$인 내적 $\langle \cdot, \cdot \rangle_{I, r(x)}$가 주어졌다고 하자. 즉,

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, \mathrm{d}x$$

   이다. 이때 $\{1, x, x^2\}$에 대해서 그람-슈미트 과정을 다음과 같이 실행한다:

$$\begin{aligned}
\boldsymbol{w}_1 &= 1 \\
\boldsymbol{w}_2 &= x - \frac{\langle 1, x \rangle}{\langle 1, 1 \rangle} 1 = x - \frac{\left.\frac{x^2}{2}\right|_{-1}^{1}}{\left.x\right|_{-1}^{1}} = x \\
\boldsymbol{w}_3 &= x^2 - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle} 1 - \frac{\langle x^2, x \rangle}{\langle x, x \rangle} x = x^2 - \frac{\left.\frac{x^3}{3}\right|_{-1}^{1}}{\left.x\right|_{-1}^{1}} - \frac{\left.\frac{x^4}{4}\right|_{-1}^{1}}{\left.\frac{x^3}{3}\right|_{-1}^{1}} x = x^2 - \frac{1}{3}
\end{aligned}$$

따라서 $\langle 1, x, x^2 \rangle$는 직교 기저 $\left\{1, x, x^2 - \frac{1}{3}\right\}$로 생성된다.

3. 벡터공간 $W = \langle 1, x, x^2, x^3 \rangle$에 대해서 $I = [-1, 1]$이고 $r(x) = 1$인 내적 $\langle \cdot, \cdot \rangle_{I, r(x)}$가 주어졌다고 하자. 이때 $\{1, x, x^2, x^3\}$에 대해서 그람-슈미트 과정을 실행할 때, 세 번째 직교 기저 벡터까지는 위에서 $\{1, x, x^2\}$에 대해서 그람-슈미트 과정을 실행한 것과 동일하다:

$$
\begin{aligned}
\boldsymbol{w}_1 &= 1 \\
\boldsymbol{w}_2 &= x \\
\boldsymbol{w}_3 &= x^2 - \frac{1}{3} \\
\boldsymbol{w}_4 &= x^3 - \frac{\langle x^3, 1 \rangle}{\langle 1, 1 \rangle} 1 - \frac{\langle x^3, x \rangle}{\langle x, x \rangle} x - \frac{\left\langle x^3, x^2 - \frac{1}{3} \right\rangle}{\left\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \right\rangle} \left(x^2 - \frac{1}{3}\right) \\
&= x^3 - \frac{\frac{x^4}{4}\Big|_{-1}^{1}}{x\Big|_{-1}^{1}} - \frac{\frac{x^5}{5}\Big|_{-1}^{1}}{\frac{x^3}{3}\Big|_{-1}^{1}} x - \frac{\frac{x^6}{6} - \frac{x^4}{12}\Big|_{-1}^{1}}{\frac{x^5}{5} - \frac{2x^3}{9} + \frac{x}{9}\Big|_{-1}^{1}} \left(x^2 - \frac{1}{3}\right) \\
&= x^3 - \frac{3}{5} x
\end{aligned}
$$

따라서 $\langle 1, x, x^2, x^3 \rangle$는 직교 기저 $\left\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\right\}$로 생성된다.

정리 8.13. *$n \times n$ 행렬 A와 $\mathbb{R}^n$에 스칼라곱이 주어져 있을 때, 다음의 명제들은 동치이다:*

1. $A^t A = I$
2. $A A^t = I$
3. $\forall \boldsymbol{x} \in \mathbb{R}^n \quad \|A\boldsymbol{x}\| = \|\boldsymbol{x}\|$
4. $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \quad A\boldsymbol{x} \cdot A\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}$
5. *A의 행벡터들이 정규직교한다.*
6. *A의 열벡터들이 정규직교한다.*

증명. 먼저 1과 2가 동치임은 $A^t = A^{-1}$이기에 자명하다. 1, 2와 5, 6이 동치임을 보이자.

$n \times n$ 행렬 A의 열벡터들을 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, 행벡터들을 $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$이라 하자. 즉,

$$
A = \begin{pmatrix} \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_n \end{pmatrix}
$$

이다. 그렇다면 A의 전치행렬 A^t는

$$A^t = \begin{pmatrix} \boldsymbol{a}_1^t \\ \vdots \\ \boldsymbol{a}_n^t \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1^t & \dots & \boldsymbol{b}_n^t \end{pmatrix}$$

이다. 따라서 A와 A^t의 곱은 다음과 같다:

$$\begin{aligned} A^t A = & \begin{pmatrix} \boldsymbol{a}_1^t \\ \vdots \\ \boldsymbol{a}_n^t \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{a}_1 & \dots & \boldsymbol{a}_1 \cdot \boldsymbol{a}_n \\ \vdots & \ddots & \vdots \\ \boldsymbol{a}_n \cdot \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \cdot \boldsymbol{a}_n \end{pmatrix} \\ = AA^t = & \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{b}_1^t & \dots & \boldsymbol{b}_n^t \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \cdot \boldsymbol{b}_1 & \dots & \boldsymbol{b}_1 \cdot \boldsymbol{b}_n \\ \vdots & \ddots & \vdots \\ \boldsymbol{b}_n \cdot \boldsymbol{b}_1 & \dots & \boldsymbol{b}_n \cdot \boldsymbol{b}_n \end{pmatrix} \end{aligned}$$

따라서 $A^t A$ 혹은 AA^t가 I라는 것은

$$\boldsymbol{a}_i \cdot \boldsymbol{a}_j = \boldsymbol{b}_i \cdot \boldsymbol{b}_j = \delta_{ij}$$

이라는 것과 동치이다. 이때 δ_{ij}는 크로넥커(Kronecker) 델타이며, 이는 $i = j$일 때 1, $i \neq j$일 때 0을 가지는 함수이다. 그러므로 1, 2와 5, 6은 동치이다.

3과 4가 동치임을 보이자. 4에서 $\boldsymbol{x} = \boldsymbol{y}$인 경우에 3이 되는 것은 자명하다. 따라서 3에서 4가 성립함을 보이자.

임의의 $\boldsymbol{x} \in \mathbb{R}^n$에 대해서 $\|A\boldsymbol{x}\| = \|\boldsymbol{x}\|$라고 하자. 벡터 $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$을 고르면,

$$\|A(\boldsymbol{a} + \boldsymbol{b})\| = \|\boldsymbol{a} + \boldsymbol{b}\|$$

가 성립한다. 이때

$$\begin{aligned} \|A(\boldsymbol{a} + \boldsymbol{b})\| &= \sqrt{A(\boldsymbol{a} + \boldsymbol{b}) \cdot A(\boldsymbol{a} + \boldsymbol{b})} \\ &= \sqrt{(A\boldsymbol{a} + A\boldsymbol{b}) \cdot (A\boldsymbol{a} + A\boldsymbol{b})} \\ &= \sqrt{A\boldsymbol{a} \cdot A\boldsymbol{a} + 2A\boldsymbol{a} \cdot A\boldsymbol{b} + A\boldsymbol{b} \cdot A\boldsymbol{b}} \\ &= \sqrt{\|A\boldsymbol{a}\|^2 + 2A\boldsymbol{a} \cdot A\boldsymbol{b} + \|A\boldsymbol{b}\|^2} \\ &= \sqrt{\|\boldsymbol{a}\|^2 + 2A\boldsymbol{a} \cdot A\boldsymbol{b} + \|\boldsymbol{b}\|^2} \end{aligned}$$

이고

$$\begin{aligned}\|\boldsymbol{a}+\boldsymbol{b}\| &= \sqrt{(\boldsymbol{a}+\boldsymbol{b}\cdot\boldsymbol{a}+\boldsymbol{b})}\\ &= \sqrt{\boldsymbol{a}\cdot\boldsymbol{a}+2\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{b}\cdot\boldsymbol{b}}\\ &= \sqrt{\|\boldsymbol{a}\|^2+2\boldsymbol{a}\cdot\boldsymbol{b}+\|\boldsymbol{b}\|^2}\end{aligned}$$

이므로,

$$A\boldsymbol{a}\cdot A\boldsymbol{b} = \boldsymbol{a}\cdot\boldsymbol{b}$$

를 만족한다. 따라서 3이면 4가 성립하며, 3과 4는 동치이다.

마지막으로 4와 6이 동치임을 보이자. 먼저 4이면 6임을 보이자.

A의 열벡터들 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$은 각각 $A\boldsymbol{e}_1,\ldots,A\boldsymbol{e}_n$와 같다. 따라서 각 열벡터들이 정규직교한다는 것은

$$\|A\boldsymbol{e}_i\| = 1$$

이고

$$A\boldsymbol{e}_i\cdot A\boldsymbol{e}_j = \delta_{ij}$$

라는 것이다. 그런데 $\boldsymbol{e}_i\cdot\boldsymbol{e}_j = \delta_{ij}$이며 4에 의해 $A\boldsymbol{e}_i\cdot A\boldsymbol{e}_j = \boldsymbol{e}_i\cdot\boldsymbol{e}_j$이므로, $A\boldsymbol{e}_i\cdot A\boldsymbol{e}_j = \delta_{ij}$이다. 나아가 4와 동치인 3에 의해 $\|A\boldsymbol{e}_i\| = \|A\boldsymbol{e}_i\| = 1$이다. 따라서 4이면 6이다.

이제 6이면 4를 보이자. 벡터 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$에 대해서,

$$\begin{aligned}\boldsymbol{a} &= (a_1,\ldots,a_n)\\ \boldsymbol{b} &= (b_1,\ldots,b_n)\end{aligned}$$

라고 하면

$$\begin{aligned}A\boldsymbol{a}\cdot A\boldsymbol{b} &= A(a_1\boldsymbol{e}_1+\cdots+a_n\boldsymbol{e}_n)\cdot A(b_1\boldsymbol{e}_1+\cdots+b_n\boldsymbol{e}_n)\\ &= \sum_{i,j=1}^{n} Aa_i\boldsymbol{e}_i\cdot Ab_j\boldsymbol{e}_j = \sum_{i,j=1}^{n} a_ib_jA\boldsymbol{e}_i\cdot A\boldsymbol{e}_j\\ &= \sum_{i,j=1}^{n} a_ib_j\delta_{ij} = \sum_{i=1}^{n} a_ib_i = \boldsymbol{a}\cdot\boldsymbol{b}\end{aligned}$$

이므로 4가 성립한다. 따라서 4와 6이 동치이며, 1, 2, 3, 4, 5, 6 모두가 동치인 명제임을 알 수 있다. □

정의 8.14. $n\times n$ 행렬 A에 대해서 정리 8.13의 조건들을 만족하는 행렬을 직교 (orthogonal) 행렬이라고 한다.

정리 8.15. *선형 등거리 변환(linear isometry)* $L_A : \mathbb{R}^n \to \mathbb{R}^n$ *에 대응되는 행렬* A *는 직교행렬이다.*

예시.

1. $\mathbb{R}^2 \to \mathbb{R}^2$ 에서 원점을 기준으로 θ만큼 회전시키는 회전변환

$$\boldsymbol{x} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \boldsymbol{x}$$

   의 행렬은 직교행렬이다.

2. $\mathbb{R}^2 \to \mathbb{R}^2$ 에서 원점을 지나고 x축과 θ의 각을 이루는 직선에 대한 대칭변환

$$\boldsymbol{x} \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \boldsymbol{x}$$

   의 행렬은 직교행렬이다.

2018.10.8. *참고.* $\mathbb{R}^3 \to \mathbb{R}^3$ 에서 벡터 $\boldsymbol{v}$를 기준으로 α만큼 회전시키는 회전변환 $R_{\boldsymbol{v}}(\alpha)$를 구하자.

우선 z 축을 중심으로 α만큼 회전하는 $R_{e_3}(\alpha)$는 $\mathbb{R}^2$ 에서의 회전변환을 통해

$$R_{e_3}(\alpha) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

임을 쉽게 알 수 있다.

이제 $\boldsymbol{v}$와 직교를 이루는 $\boldsymbol{u}, \boldsymbol{w}$를 잡자. 이때 $\boldsymbol{w} \times \boldsymbol{u}$가 $\boldsymbol{v}$와 평행이 되도록 하자. 이들을 정규화한 $w/\|w\|, u/\|u\|, v/\|v\|$를 각각 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$로 변환하는 선형변환 L을 잡자. 즉,

$$L\left(\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}\right) = \boldsymbol{e}_1$$
$$L\left(\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\right) = \boldsymbol{e}_2$$
$$L\left(\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}\right) = \boldsymbol{e}_3$$

이다. 그런데 L은 회전변환이므로 반드시 역함수 L^{-1}가 존재한다. 따라서

$$\begin{aligned}\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} &= L^{-1}(\boldsymbol{e}_1)\\ \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} &= L^{-1}(\boldsymbol{e}_2)\\ \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} &= L^{-1}(\boldsymbol{e}_3)\end{aligned}$$

이고, L에 대응되는 행렬은

$$\begin{pmatrix} \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} & \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} & \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \end{pmatrix}$$

이다. 그런데 $\boldsymbol{w}/\|\boldsymbol{w}\|, \boldsymbol{u}/\|\boldsymbol{u}\|, \boldsymbol{v}/\|\boldsymbol{v}\|$는 정규직교하므로 이는 직교행렬이고, 따라서 역행렬은

$$\begin{pmatrix} \frac{\boldsymbol{w}^t}{\|\boldsymbol{w}\|} \\ \frac{\boldsymbol{u}^t}{\|\boldsymbol{u}\|} \\ \frac{\boldsymbol{v}^t}{\|\boldsymbol{v}\|} \end{pmatrix}$$

이다.

$\boldsymbol{v}$를 기준으로 α만큼의 회전은 $\boldsymbol{v}/\|\boldsymbol{v}\|$가 $\boldsymbol{e}_3$가 되도록 회전한 후, z 축을 중심으로 α만큼의 회전, 즉 $R_{\boldsymbol{e}_3}(\alpha)$를 한 후, 다시 원래대로 $\boldsymbol{e}_3$를 $\boldsymbol{v}/\|\boldsymbol{v}\|$의 위치로 옮기는 것이다. 따라서 $R_{\boldsymbol{v}}(\alpha)$는 아래와 같은 행렬에 대응된다:

$$\begin{pmatrix} \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} & \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} & \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\boldsymbol{w}^t}{\|\boldsymbol{w}\|} \\ \frac{\boldsymbol{u}^t}{\|\boldsymbol{u}\|} \\ \frac{\boldsymbol{v}^t}{\|\boldsymbol{v}\|} \end{pmatrix}$$

벡터함수 $\boldsymbol{F} = (f_1, \ldots, f_n)$가 $\mathbb{R}^m \to \mathbb{R}^n$으로 갈 때, 점 P에서 다음과 같이 근사할 수 있다:

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{x}) &= \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{pmatrix} \approx \begin{pmatrix} f_1(P) + f_1'(P)(\boldsymbol{x} - P) \\ \vdots \\ f_n(P) + f_n'(P)(\boldsymbol{x} - P) \end{pmatrix} \\ &= \begin{pmatrix} f_1'(P) \\ \vdots \\ f_n'(P) \end{pmatrix} (\boldsymbol{x} - P) + \boldsymbol{F}(P)\end{aligned}$$

이때 $\boldsymbol{F}$의 P에서의 야코비 행렬(Jacobian matrix)을 다음과 같이 정의하여 더 간단하게 표현할 수 있다.

정의 8.16. 벡터함수 $\boldsymbol{F} : \mathbb{R}^m \to \mathbb{R}^n$에 대해 야코비 행렬 (Jacobian matrix)을 다음과 같이 정의한다:

$$\boldsymbol{F}' = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_m)} = \begin{pmatrix} f_1' \\ \vdots \\ f_n' \end{pmatrix}$$

이때 $\boldsymbol{F} = (f_1, \ldots, f_n)$이다.

따라서 $\boldsymbol{F}$의 P에서의 근사는

$$\boldsymbol{F}(\boldsymbol{x}) \approx \boldsymbol{F}'(P)(\boldsymbol{x} - P) + \boldsymbol{F}(P)$$

로 쓸 수 있다.

예시. 각도를 보존하는 사상을 등각사상 (conformal mapping)이라고 하는데, 어떤 사상의 야코비 행렬이 직교행렬의 상수 배이면 해당 사상은 등각사상임이 알려져 있다. $\boldsymbol{F}(x, y) = (f(x, y), g(x, y)) = (x^2 - y^2, -2xy)$로 정의된 벡터함수의 야코비 행렬을 구하면

$$\boldsymbol{F}' = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ -2y & -2x \end{pmatrix}$$

이다. 그런데 $\boldsymbol{F}'$의 열벡터 $2(x, -y)$와 $2(-y, -x)$의 내적이 0이므로 두 열벡터는 직교한다. 따라서 $\boldsymbol{F}'$는 직교행렬의 상수배임을 알 수 있고, $\boldsymbol{F}$는 등각사상이다.

정의 8.17. $n \times n$ 행렬 A에 대해서 $A\boldsymbol{x} = \lambda\boldsymbol{x}$가 되는 $\boldsymbol{x} \neq \boldsymbol{0}$가 존재한다. 이때 λ를 특성값 (eigenvalue), $\boldsymbol{x}$를 λ-특성벡터 (eigenvector) 혹은 단순히 특성벡터라고 부르며, λ-특성벡터들과 $\boldsymbol{0}$의 집합 E_λ를 λ-특성공간 (eigenspace)이라고 부른다. 또한 $\det(A - \lambda I)$을 특성다항식 (characteristic polynomial)이라고 부르고, $\operatorname{char} A$라고 쓴다.

참고. $n \times n$ 행렬 A의 특성값 λ와 특성벡터 $\boldsymbol{x}$를 구하기 위해서,

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

에서 우변을 좌변으로 이항한다:

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

그런데 $\boldsymbol{x} \neq \boldsymbol{0}$이므로 $A - \lambda I$는 정리 7.16와 7.45에 의해 $\det(A - \lambda I) = 0$이어야 한다. 따라서 특성다항식 $\operatorname{char} A = \det(A - \lambda I) = 0$을 풀면 특성값 λ를 구할 수

있다. λ-특성벡터 $\boldsymbol{x}$의 경우, $(A-\lambda I)\boldsymbol{x}=\mathbf{0}$이므로 $\mathrm{N}(A-\lambda I)$의 원소를 구하면 된다. 이는 $A-\lambda I$를 행사다리꼴 형태로 바꾸어 구할 수 있다.

예시.

1. 행렬 $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$의 특성값들과 각각의 특성벡터들을 찾기 위해 특성다항식을 푼다:

$$\det\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = (\lambda-3)(\lambda-1) = 0$$

   따라서 $\lambda=3$일 경우와 $\lambda=1$일 경우를 나누어 특성벡터를 찾는다.

   a) $\lambda=1$일 때
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\boldsymbol{x}=\mathbf{0}$$
   이므로 $k\in\mathbb{R}$에 대해 1-특성벡터 $\boldsymbol{x}=k(1,-1)$이며 1-특성공간 $E_1 = \langle 1,-1\rangle$이다.

   b) $\lambda=3$일 때
$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\boldsymbol{x}=\mathbf{0}$$
   이므로 $k\in\mathbb{R}$에 대해 3-특성벡터 $\boldsymbol{x}=k(1,1)$이며 3-특성공간 $E_3 = \langle 1,1\rangle$이다.

2. $0<\alpha<\pi$에 대해 행렬 $A=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$의 특성값은 실수 범위에서 존재하지 않는다. 이는

$$\operatorname{char} A = \lambda^2 - 2\lambda\cos\alpha + 1$$

   인데, λ에 대한 이차방정식의 판별식에서 $D/4 = \cos^2\alpha - 1 < 0$이기 때문에 실근을 가지지 않는다. 하지만 복소수 범위에서 $\lambda = e^{i\alpha}$ 혹은 $\lambda = e^{-i\alpha}$임을 알 수 있다.

2018.10.10.

3. 행렬 $\begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$의 특성값들과 각각의 특성벡터들을 찾기 위해 특성

다항식을 푼다:

$$\det\begin{pmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{pmatrix} = -(\lambda-5)(\lambda+3)^2$$

따라서 $\lambda = 5$일 경우와 $\lambda = -3$일 경우를 나누어 특성벡터를 찾는다.

a) $\lambda = 5$일 때

$$\begin{pmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{pmatrix} \to \begin{pmatrix} -7 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

이므로 $k \in \mathbb{R}$에 대해 5-특성벡터 $\boldsymbol{x} = k(-1,-2,1)$이며 5-특성공간 $E_5 = \langle(-1,-2,1)\rangle$이다.

b) $\lambda = -3$일 때

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

이므로 $k, l \in \mathbb{R}$에 대해 -3-특성벡터 $\boldsymbol{x} = k(-2,1,0) + l(3,0,1)$이며 -3-특성공간 $E_{-3} = \langle(-2,1,0),(3,0,1)\rangle$이다.

여기서 $\lambda = 5$일 때 $\operatorname{char} A$에서는 $\lambda - 5$ 항이 하나이고 $\dim E_5 = 1$로 같고, $\lambda = -3$일 때 $\operatorname{char} A$에서 $\lambda + 3$ 항이 두 개, $\dim E_{-3} = 2$로 동일하다. 그러나 이렇게 특성방정식에서 중근의 개수와 해당 특성값에 대응되는 특성공간의 차원이 항상 같은 것은 아니다.

정의 8.18. 행렬 A에 대해서, $\operatorname{char} A$의 어떤 해, 즉 특성값이 나타나는 횟수를 대수적 중복도 (algebraic multiplicity), 해당 특성값에 대응되는 특성공간의 차원을 기하적 중복도 (geometric multiplicity) 라고 한다.

정리 8.19. *행렬 A의 기하적 중복도는 대수적 중복도를 초과하지 못한다.*

예시. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$의 특성방정식은 $(\lambda-1)^2 = 0$이다. 따라서 $\lambda = 1$의 대수적 중복도는 1이고, $A - \lambda I = O$이므로 $\mathbb{R}^2$의 임의의 벡터가 특성벡터이다. 그러므로 특성공간은 $\mathbb{R}^2$이고, 기하적 중복도는 2이다.

$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 의 특성방정식은 $\lambda^2 = 0$이다. 따라서 $\lambda = 0$의 대수적 중복도는 2이고, $k \in \mathbb{R}$에 대해 $(k, 0)$이 특성벡터이므로, 특성공간은 $\langle(1, 0)\rangle$로 기하적 중복도가 1이다.

위의 두 경우 모두 기하적 중복도가 대수적 중복도 이하임을 확인할 수 있다.

정의 8.20. 벡터공간 V와 부분공간 U, W에 대해서 $U \cap W = \{\mathbf{0}\}$일 때 U와 W의 합을 직합(direct sum)이라고 하고 $U \oplus W$라고 쓴다.

참고. 그라스만 공식(정리 7.10)에서 $\dim(U + W) = \dim U + \dim W$임을 알 수 있다.

정리 8.21. *두 벡터공간 U와 W의 교집합이 $\{\mathbf{0}\}$일 때, $\boldsymbol{x} \in U \oplus W$에 대해서*

$$\boldsymbol{x} = \boldsymbol{u}_1 + \boldsymbol{w}_1 = \boldsymbol{u}_2 + \boldsymbol{w}_2$$

이고 $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$, $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$일 때, $\boldsymbol{u}_1 = \boldsymbol{u}_2$이고 $\boldsymbol{w}_1 = \boldsymbol{w}_2$이다. 즉, $U \oplus W$의 원소의 합 표현은 유일하다.

증명. $\boldsymbol{x} \in U \oplus W$에 대해서 $\boldsymbol{u}_1 + \boldsymbol{w}_1 = \boldsymbol{u}_2 + \boldsymbol{w}_2 = \boldsymbol{x}$이고 $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$, $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$라고 하자. 그러면 $\boldsymbol{u}_1 - \boldsymbol{u}_2 = \boldsymbol{w}_2 - \boldsymbol{w}_1$이므로 $U \cap W = \{\mathbf{0}\}$의 원소이다. 따라서 $\boldsymbol{u}_1 - \boldsymbol{u}_2 = \boldsymbol{w}_2 - \boldsymbol{w}_1 = \mathbf{0}$이어서 $\boldsymbol{x}$의 합 표현은 유일하다. □

정의 8.22. 벡터공간 V의 부분공간 $U_1, \ldots, U_k$에 대해서 어떤 벡터 $\boldsymbol{x} \in U_1 + \cdots + U_k$의 합 표현이 유일할 때

$$U_1 + \cdots + U_k = U_1 \oplus \cdots \oplus U_k$$

로 쓴다.

정리 8.23. *벡터공간 $U_1, \ldots, U_k$에 대해서 $U_1 \oplus \cdots \oplus U_k$가 존재할 때,*

$$\{\boldsymbol{u}_i \mid \boldsymbol{u}_i \in U_i \setminus \{\mathbf{0}\},\ i \in \{1, \ldots, k\}\}$$

은 일차독립이다.

증명. 스칼라 $c_1, \ldots, c_k$에 대해서

$$c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k = \mathbf{0}$$

라고 가정하자. $\boldsymbol{x}$를

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + \cdots + c_{k-1}\boldsymbol{u}_{k-1} = -c_k\boldsymbol{u}_k$$

로 두면,

$$\begin{aligned}\boldsymbol{x} &= c_1\boldsymbol{u}_1 + \cdots + c_{k-1}\boldsymbol{u}_{k-1} + 0\boldsymbol{u}_k \\ &= 0\boldsymbol{u}_1 + \cdots + 0\boldsymbol{u}_{k-1} + (-c_k)\boldsymbol{u}_k\end{aligned}$$

이고 합 표현의 유일성에 따라 $c_1 = \cdots = c_k = 0$이다. 그러므로 $\{\boldsymbol{u}_i \mid \boldsymbol{u}_i \in U_i \setminus \{\boldsymbol{0}\},\ i \in \{1, \ldots, k\}\}$은 일차독립이다. □

정리 8.24. *$n \times n$ 행렬 A에 대해서 서로 다른 특성값 $\lambda_1, \ldots, \lambda_k$에 대응되는 특성공간 $E_1, \ldots, E_n$에 대해, 다음이 성립한다:*

$$E_1 + \cdots + E_k = E_1 \oplus \cdots \oplus E_k$$

증명. k에 대한 수학적 귀납법으로 증명한다.

먼저 $k = 2$일 경우, 어떤 벡터 $\boldsymbol{v} \in E_1 \cap E_2$를 고르자. 그러면

$$A\boldsymbol{v} = \lambda_1\boldsymbol{v} = \lambda_2\boldsymbol{v}$$

이므로,

$$(\lambda_1 - \lambda_2)\boldsymbol{v} = \boldsymbol{0}$$

이다. 그런데 $\lambda_1 \neq \lambda_2$이므로 $\boldsymbol{v} = \boldsymbol{0}$이다. 따라서 $E_1 + E_2 = E_1 \oplus E_2$이다.

이제 $k > 2$의 경우에, $E_1 + \cdots + E_{k-1} = E_1 \oplus \cdots \oplus E_{k-1}$가 성립한다고 가정하자. $\boldsymbol{x} \in E_1 + \cdots + E_k$에 대해서 $\boldsymbol{v}_i, \boldsymbol{u}_i \in E_i$일 때

$$\boldsymbol{x} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = \boldsymbol{u}_1 + \cdots + \boldsymbol{u}_k \in E_1 + \cdots + E_k$$

라고 하자. 그러면

$$(\boldsymbol{v}_1 - \boldsymbol{u}_1) + \cdots + (\boldsymbol{v}_{k-1} - \boldsymbol{u}_{k-1}) = \boldsymbol{u}_k - \boldsymbol{v}_k$$

이고, $\boldsymbol{v}_i - \boldsymbol{u}_i = \boldsymbol{w}_i$로 놓자. 따라서

$$\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_{k-1} = -\boldsymbol{w}_k$$

이고 $\boldsymbol{w}_i \in E_i$이다. 양변에 A를 왼쪽에 곱하면

$$\begin{aligned} & A(\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_{k-1}) = A(-\boldsymbol{w}_k) \\ \Leftrightarrow\ & A\boldsymbol{w}_1 + \cdots + A\boldsymbol{w}_{k-1} = -A\boldsymbol{w}_k \\ \Leftrightarrow\ & \lambda_1\boldsymbol{w}_1 + \cdots + \lambda_{k-1}\boldsymbol{w}_{k-1} = -\lambda_k\boldsymbol{w}_k\end{aligned}$$

이다. 또한 양변에 λ_k를 곱한 것은

$$\lambda_k \boldsymbol{w}_1 + \cdots + \lambda_k \boldsymbol{w}_{k-1} = -\lambda_k \boldsymbol{w}_k$$

이다. 둘을 양변에서 빼면

$$(\lambda_1 - \lambda_k)\boldsymbol{w}_1 + \cdots + (\lambda_{k-1} - \lambda_k)\boldsymbol{w}_{k-1} = \mathbf{0}$$

이다. 따라서 $i \in \{1, \ldots, k-1\}$에 대해 $\lambda_i \neq \lambda_k$이므로 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}$는 일차종속이다. 그런데 귀납 가정에 따라 $E_1 + \cdots + E_{k-1} = E_1 \oplus \cdots \oplus E_{k-1}$이며 정리 8.23에 따라 $\mathbf{0}$가 아닌 벡터들은 일차독립이어야 하므로, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{k-1}$은 $\mathbf{0}$이다. 그러므로 $\boldsymbol{v}_i = \boldsymbol{u}_i$이고, 벡터 $\boldsymbol{x} \in E_1 + \cdots + E_k$의 합 표현은 유일하고, $E_1 + \cdots + E_k = E_1 \oplus \cdots \oplus E_k$이다. 이로써 귀납 증명이 완료된다. □

참고. 벡터공간 $U_1, \ldots, U_k$에 대해서

$$U_1 + \cdots + U_k = U_1 \oplus \cdots \oplus U_k \Rightarrow \big(\forall i, j \in \{1, \ldots, k\} \quad i \neq j \Rightarrow U_i \cap U_j = \{\mathbf{0}\}\big)$$

이지만 그 역은 성립하지 않는다. 예를 들어, $\mathbb{R}^2$ 좌표평면에서 원점을 지나는 세 직선으로 나타내어지는 벡터공간의 교집합은 원점뿐이지만, $\mathbb{R}^2$의 벡터들은 세 직선에 속하는 벡터들의 합 표현으로 유일하게 나타내지지 않는다.

그러나 벡터공간 $U_1, \ldots, U_k$들의 집합 $\mathcal{U}$에 대해서,

$$\forall \mathcal{P}_1, \mathcal{P}_2 \subseteq \mathcal{U} \qquad \mathcal{P}_1 \cap \mathcal{P}_2 = \varnothing \Rightarrow \sum_{U \in \mathcal{P}_1} U \ \cap \sum_{\tilde{U} \in \mathcal{P}_2} \tilde{U} = \{\mathbf{0}\}$$

인 것은 $U_1 + \cdots + U_k = U_1 \oplus \cdots \oplus U_k$와 동치이다.

정리 8.25. *행렬 A의 서로 다른 특성공간 $E_1, \ldots E_k$와 각각의 기저가 $\mathcal{B}_1, \ldots, \mathcal{B}_k$이라고 하자. 그러면 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$는 $E_1 \oplus \cdots \oplus E_k$의 기저이다.*

증명. $i \in \{1, \ldots, k\}$에 대해 E_i의 기저 $\mathcal{B}_i = \{\boldsymbol{v}_{i1}, \ldots, \boldsymbol{v}_{ij_i}\}$라고 하자. 먼저 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$가 $E_1 \oplus \cdots \oplus E_k$를 생성하는 것은 각각의 $\mathcal{B}_i$가 E_i를 생성하므로 당연하다. 이제 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$가 일차독립임을 보이자.

기저 $\mathcal{B}_i$에 속하는 기저 벡터들의 선형결합들의 합을

$$\sum_{i=1}^{k} (c_{i1}\boldsymbol{v}_{i1} + \cdots + c_{ij_i}\boldsymbol{v}_{ij_i}) = \mathbf{0}$$

라고 하자. $l \in \{1, \ldots, k\}$를 하나 골라서 우변으로 이항하면,

$$\sum_{i \neq l} (c_{i1} v_{i1} + \cdots + c_{ij_i} v_{ij_i}) = -(c_{l1} v_{l1} + \cdots + c_{lj_l} v_{lj_l})$$

이다. 그런데 정리 8.24에 따라 양변은 $E_1 \oplus \cdots \oplus E_k$의 원소이므로 합 표현이 유일하고, 따라서 모든 c_{ij_i}는 0이다. 따라서 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$가 일차독립이다. □

정의 8.26. $L_A : \mathbb{R}^n \to \mathbb{R}^n$에 대응되는 $n \times n$ 행렬 A가 있을 때, $v_1, \ldots, v_n$이 $\mathbb{R}^n$에 대한 행렬 A의 특성기저(eigenbasis)라는 것은 $v_1, \ldots, v_n$이 A의 특성벡터이면서 $\mathbb{R}^n$의 기저벡터라는 것을 말한다.

참고. $\mathbb{R}^n$에 대한 $n \times n$ 행렬 A의 특성기저가 존재한다는 것은 char A의 중근을 포함한 실근의 개수가 n개라는 것이고, 이는 대수적 중복도의 합과 기하적 중복도의 합이 n으로 같다는 것이다.

2018.10.15 **정의 8.27.** $n \times n$ 행렬 A와 B에 대해서 A와 B가 닮았다(similar)는 것은 $n \times n$ 가역행렬 P가 존재해서

$$A = P^{-1}BP$$

라는 것이다. 이 때 $A \sim B$라고 표기한다.

정리 8.28. *$n \times n$ 행렬 A와 B에 대해서 다음이 성립한다:*

1. *$A \sim A$ (반사성(reflexivity))*
2. *$A \sim B \Rightarrow B \sim A$ (대칭성(symmetricity))*
3. *$A \sim B \wedge B \sim A \Rightarrow A \sim C$ (추이성(transivity))*

증명.

1. 반사성은 자명하게 성립한다.
2. 대칭성은 $A = P^{-1}BP$이면 $B = (P^{-1})^{-1}A(P^{-1})$이므로 성립한다.
3. 추이성은 $A = P^{-1}BP$이고 $B = Q^{-1}CQ$일 때 $A = (QP)^{-1}C(QP)$이므로 성립한다.

□

정의 8.29. $n \times n$ 행렬의 주대각선을 제외한 곳의 원소가 모두 0이면 대각행렬(diagonal matrix)이라고 부른다. 또한 대각행렬과 닮은 행렬을 대각화 가능 행렬(diagonalizable matrix)이라고 부른다.

정리 8.30. *$n \times n$ 행렬 A가 대각화 가능 행렬일 필요충분조건은 A이 $\mathbb{R}^n$에 대해 특성기저를 가지는 것이다.*

증명. 먼저 $\mathbb{R}^n$에 대한 $n \times n$ 행렬 A의 특성기저가 존재하면 A가 대각화 가능하다는 것을 보이자.

$\mathbb{R}^n$에 대한 $n \times n$ 행렬 A의 특성기저를 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, 각각에 대응되는 특성값을 $\lambda_1, \ldots, \lambda_n$이라고 하자. 즉,

$$\begin{aligned} A\boldsymbol{v}_1 &= \lambda_1 \boldsymbol{v}_1 \\ &\vdots \\ A\boldsymbol{v}_n &= \lambda_n \boldsymbol{v}_n \end{aligned}$$

이며, 이는 다시

$$A \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

로 쓸 수 있다. 따라서

$$A = \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix}^{-1}$$

이므로 A는 대각화 가능 행렬이다.

이제 A가 대각화 가능하면 $\mathbb{R}^n$에 대해 $n \times n$ 행렬 A의 특성기저가 존재한다는 것을 보이자.

$n \times n$ 가역행렬 P과 대각행렬 D에 대해서 $A = PDP^{-1}$이라고 하자. P의 열벡터들을 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$으로 놓고, $i \in \{1, \ldots, n\}$일 때 D의 i행 i열의 원소를 λ_i로 놓자. λ_i들이 A의 특성값이고, $\boldsymbol{v}_i$가 λ_i에 대응되는 특성벡터임을 보이자.

$$A \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

이므로 모든 $\boldsymbol{v}_i$에 대해 $A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$임을 알 수 있다. 따라서 λ_i들은 A의 특성값이고 $\boldsymbol{v}_i$는 λ_i-특성벡터이다. 이제 $\boldsymbol{v}_i$들이 $\mathbb{R}^n$의 기저벡터임을 보이자. P는 가역행렬이므로 열벡터들이 일차독립이다 (쪽 27의 7과 9). 나아가 정리 7.31에 따라 $\boldsymbol{v}_i$들은 기저를 이룬다. □

참고. 위의 증명 과정에서 임의의 특성기저를 상수배하여도 대각화할 때 사용되는 가역행렬의 열벡터로 사용할 수 있다는 것을 알 수 있다. 또한, 두 열벡터를 교환하면 각각의 특성벡터에 해당하는 특성값을 교환하면 여전히 대각화가 성립한다. 즉,

$$\begin{aligned} A &= \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2 & 4 & 7 \\ 4 & 5 & 8 \\ 6 & 6 & 10 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 & 7 \\ 4 & 5 & 8 \\ 6 & 6 & 10 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 4 & 2 & 7 \\ 5 & 4 & 8 \\ 6 & 6 & 10 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 4 & 2 & 7 \\ 5 & 4 & 8 \\ 6 & 6 & 10 \end{pmatrix}^{-1} \end{aligned}$$

이다.

정리 8.31. *$n \times n$ 행렬 A가 대각화 가능하면, 즉 어떤 $n \times n$ 가역행렬 P와 대각행렬 D에 대해서 $A = PDP^{-1}$ 일 때,*

$$A^k = PD^kP^{-1}$$

이다.

증명. $A = PDP^{-1}$일 때

$$A^k = (PDP^{-1})^k = \underbrace{(PDP^{-1})\dots(PDP^{-1})}_{k\text{ 개}} = P\underbrace{D\dots D}_{k\text{ 개}}P^{-1} = PD^kP^{-1}$$

이다. □

정리 8.32. *대칭 실행렬 A는 대각화 가능 행렬이다. 즉, $A = A^t$ 인 것은 char A의 모든 근이 실수이고 모든 근에 대해 대수적 중복도와 기하적 중복도가 같은 것과 동치이다.*

도움정리 8.33. *$n \times n$ 행렬 A가 대칭행렬인 것은 A가 스칼라곱에 대해 자기 수반 (self-adjoint) 이라는 것이다. 즉,*

$$A = A^t \qquad \Longleftrightarrow \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \quad A\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot A\boldsymbol{y}$$

이다.

증명. 우선 A가 대칭행렬이면 스칼라곱에 대해 자기 수반이라는 것을 보이자.

어떤 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$을 고르자. 그러면

$$\boldsymbol{x} \cdot A\boldsymbol{y} = (A\boldsymbol{y})^t \boldsymbol{x} = \boldsymbol{y}^t A^t \boldsymbol{x} = \boldsymbol{y}^t A\boldsymbol{x} = A\boldsymbol{x} \cdot \boldsymbol{y}$$

이므로 A는 스칼라곱에 대해 자기 수반이다.

이제 A가 스칼라곱에 대해 자기 수반이면 A가 대칭행렬임을 보이자.

$A = (\boldsymbol{v}_1 \quad \dots \quad \boldsymbol{v}_n) = (a_{ij})$라고 하자. $A\boldsymbol{e}_i = \boldsymbol{v}_i$이고, $\boldsymbol{v}_i \cdot \boldsymbol{e}_j = a_{ji}$이다. 그런데

$$A\boldsymbol{e}_i \cdot e_j = e_i \cdot A\boldsymbol{e}_j$$

이므로,

$$a_{ji} = a_{ij}$$

이다. 따라서 $A = A^t$이다. 그러므로 A가 대칭행렬인 것은 A가 스칼라곱에 대해 자기 수반인 것과 동치이다. □

참고. $L : \mathcal{C}^\infty \to \mathcal{C}^\infty$에 대해서도 내적 $\langle \cdot, \cdot \rangle$이 주어졌을 때 임의의 x, y에 대해 $\langle Lx, y \rangle = \langle x, Ly \rangle$인 변환을 생각할 수 있다.

정리 8.34. *대칭 실행렬 A의 서로 다른 특성값 λ와 μ에 대해서, 대응되는 특성벡터 $\boldsymbol{x}$와 $\boldsymbol{y}$는 직교한다.*

증명. A가 대칭 실행렬인 것은 도움정리 8.33에 따라 A가 스칼라곱에 대해 자기 수반인 것을 의미한다. 따라서 λ-특성벡터 $\boldsymbol{x}$와 μ-특성벡터 $\boldsymbol{y}$에 대해

$$A\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot A\boldsymbol{y}$$

이며, 따라서

$$\lambda\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot \mu\boldsymbol{y}$$

이다. 우변을 좌변으로 이항하면

$$(\lambda - \mu)\boldsymbol{x} \cdot \boldsymbol{y} = 0$$

이고, $\lambda \neq \mu$이므로 $\boldsymbol{x} \cdot \boldsymbol{y} = 0$이다. □

예시. $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 8 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^{-1}$는 대칭행렬이 아니다. $(1,2) \cdot (3,4) \neq 0$이기 때문이다.

참고. 정리 8.34에 따라 대칭행렬은 가역 직교행렬을 통해 대각화할 수 있음을 알 수 있다. 예를 들어, 어떤 행렬 A가 다음과 같이 대각화된다고 하자:

$$A = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{w}_1 & \boldsymbol{w}_2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{w}_1 & \boldsymbol{w}_2 \end{pmatrix}^{-1}$$

이 때 A의 특성값 $2, 3$과 2-특성공간 $E_2 = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \rangle$과 3-특성공간 $E_3 = \langle \boldsymbol{w}_1, \boldsymbol{w}_2 \rangle$을 알 수 있다. 각각의 특성공간의 기저에 그람-슈미트 과정을 통해서 직교 기저를 얻을 수 있고, 정규화하여 정규직교 기저를 구할 수 있다. E_2와 E_3의 기저들을 정규직교화하여 $\{\tilde{\boldsymbol{v}}_1, \tilde{\boldsymbol{v}}_2, \tilde{\boldsymbol{v}}_3\}$와 $\{\tilde{\boldsymbol{w}}_1, \tilde{\boldsymbol{w}}_2\}$를 구했을 때, A는 다음과 같이 직교 대각화할 수 있다:

$$A = \begin{pmatrix} \tilde{\boldsymbol{v}}_1 & \tilde{\boldsymbol{v}}_2 & \tilde{\boldsymbol{v}}_3 & \tilde{\boldsymbol{w}}_1 & \tilde{\boldsymbol{w}}_2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{v}}_1^t \\ \tilde{\boldsymbol{v}}_2^t \\ \tilde{\boldsymbol{v}}_3^t \\ \tilde{\boldsymbol{w}}_1^t \\ \tilde{\boldsymbol{w}}_2^t \end{pmatrix}$$

즉, 대칭행렬은 직교 대각화 가능하다: 대각행렬 A에 대해서

$$A = PDP^t \quad \wedge \quad PP^t = I$$

인 P가 존재한다.

2018.10.17 **정리 8.35.** *함수* $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *가 내적이라는 것은 벡터* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ *에 대해 특성값이 모두 양수인 어떤* $n \times n$ *대칭행렬* A *가 존재해서 다음을 만족한다는 것과 동치이다:*

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{b}^t A \boldsymbol{a}$$

증명. 정리 8.4에 의해 어떤 $n \times n$ 대칭행렬 A가 존재하여 $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{b}^t A \boldsymbol{a}$인 것은 $\langle \cdot, \cdot \rangle$이 대칭 쌍선형형식인 것과 동치이다. 따라서 $n \times n$ 대칭행렬 A의 특성값이 모두 양수이면 $\langle \cdot, \cdot \rangle$이 양의 정부호성을 가지는 것과 동치인 것만을 보이면 된다.

$A = PDP^t$와 같이 대각화를 하고, D의 i행 i열의 원소를 특성값 λ_i라 하자. 어떤 벡터 $\boldsymbol{x} \in \mathbb{R}^n$에 대해 $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$은 다음과 같이 정리할 수 있다:

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^t A \boldsymbol{x} = \boldsymbol{x}^t PDP^t \boldsymbol{x} = (P^t \boldsymbol{x})^t D (P^t \boldsymbol{x})$$

이때 $P^t\boldsymbol{x} = \boldsymbol{y} = (y_1, \dots, y_n)$이라 두면,

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{x} \rangle &= \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 y_1 & \dots & \lambda_n y_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2\end{aligned}$$

그런데 $P^t\boldsymbol{x} = \boldsymbol{y}$은 $\boldsymbol{x} = P\boldsymbol{y}$인 것과 동치이므로, $\boldsymbol{x}$와 $\boldsymbol{y}$는 P에 의해 일대일 대응된다. 그러므로 임의의 $\boldsymbol{x}$에 대해 $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$이 양의 정부호를 가진다는 것은 임의의 $\boldsymbol{y}$에 대해 $\sum_{i=1}^n \lambda_i y_i^2$이 0 이상이고, 0일 경우는 $\boldsymbol{y} = \mathbf{0}$이라는 것과 동치이다. 임의의 $\boldsymbol{y}$에 대해 $\sum_{i=1}^n \lambda_i y_i^2$이 0 이상이고 $\boldsymbol{y} = \mathbf{0}$일 때만 0이라는 것은 $\lambda_i > 0$이라는 것이다. 그러므로 $\langle \cdot, \cdot \rangle$이 양의 정부호를 가질 필요충분조건은 모든 특성값이 양수인 것이다. □

정의 8.36. 대칭행렬 A의 모든 특성값이 양수이면 양의 정부호(positive-definite)를 가진다고 하고, 음수가 아니면 양의 준정부호(positive semi-definite)를 가지며, 양의 특성값과 음의 특성값 모두를 가지면 부정부호(indefinite)를 가진다고 한다.

예시. $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$가 양의 정부호를 가질 조건을 구하여라.

풀이. 특성방정식을 구하면

$$\operatorname{char} A = \det(A - \lambda I) = (a-\lambda)(c-\lambda) - b^2 = \lambda^2 - (\operatorname{tr} A)\lambda + \det A$$

이므로, $\operatorname{char} A = 0$의 모든 해가 양수이려면 아래로 볼록인 이차함수가 원점 오른쪽에서 x축과 두 번 교차할 조건에 따라

$$ac - b^2 > 0 \quad \wedge \quad \frac{a+c}{2} > 0 \quad \wedge \quad (a+c)^2 - 4(ac - b^2) \geq 0$$

이다. 그런데 마지막 식은 $(a-c)^2 + 4b^2 \geq 0$으로 정리되어 항등식이므로,

$$\det A > 0 \quad \wedge \quad \operatorname{tr} A > 0$$

가 A가 양의 정부호를 가질 조건이다. ■

1 First-Order ODEs

미분방정식

$$y' = y$$

를 만족하는 $y(x)$를 구하면,

$$y = ke^x \quad (k \in \mathbb{R})$$

임을 알 수 있다. 이때 위 해들의 집합 $\{ke^x \mid k \in \mathbb{R}\}$을 해집합 (solution set) 이라고 한다. 또한, $y' = y$와 같이 하나의 독립변수만을 가지는 함수를 구하는 미분방정식을 상미분방정식 (ordinary differential equation, ODE) 이라고 한다. 특히 $y' = y$는 1계 제차 선형 상미분방정식의 예시이다.

정의 1.1. 상미분방정식이 y의 도함수의 일차결합으로 나타낼 수 있는 경우, 선형 상미분방정식 (linear ODE) 이라고 한다. 즉,

$$L[y] = q_n(x)y^{(n)} + q_{n-1}(x)y^{(n-1)} + \cdots + q_1(x)y' + q_0(x)y = r(x)$$

와 같이 나타낼 수 있는 상미분방정식을 선형 상미분방정식이라고 한다. 이때 $q_i(x)$를 계수 (coefficient) 함수, $r(x)$를 초항 (source term) 이라고 한다.

$r(x) = 0$일 경우를 제차 선형 상미분방정식 (homogeneous linear ODE), $r(x) \neq 0$일 경우를 비제차 선형 상미분방정식 (nonhomogeneous linear ODE) 이라고 한다. 특히 $q_n(x) \neq 0$인 경우 n계 선형 상미분방정식 (n-th order linear ODE) 이라고 부른다.

반면

$$L[y] = y'' + \cos y = x$$

와 같이 y의 도함수의 일차결합이 아닌 경우 비선형 상미분방정식 (nonlinear ODE) 이라고 부른다.

정리 1.2. *제차 선형 상미분방정식 $L[y] = 0$의 두 해 y_1과 y_2에 대해서, 둘의 일차결합 $c_1y_1 + c_2y_2$도 $L[y] = 0$의 해이다. 즉, 제차 선형 상미분방정식의 해집합은 벡터공간을 이루어 해공간이라고 부를 수 있다.*

증명. 제차 선형 상미분방정식

$$L[y] = q_n(x)y^{(n)} + \cdots + q_1(x)y' + q_0(x)y = 0$$

의 두 해 y_1, y_2에 대해서

$$
\begin{aligned}
&L[c_1y_1 + c_2y_2] \\
&= q_n(x)(c_1y_1 + c_2y_2)^{(n)} + \cdots + q_1(x)(c_1y_1 + c_2y_2)' + q_0(x)(c_1y_1 + c_2y_2) \\
&= q_n(x)\left(c_1y_1^{(n)} + c_2y_2^{(n)}\right) + \cdots + q_1(x)\left(c_1y_1' + c_2y_2'\right) + q_0(x)\left(c_1y_1 + c_2y_2\right) \\
&= \left(c_1q_n(x)y_1^{(n)} + \cdots + c_1q_1(x)y_1' + c_1q_0(x)y_1\right) + \left(c_2q_n(x)y_2^{(n)} + \dots\right. \\
&\qquad \left. + c_2q_1(x)y_2' + c_2q_0(x)y_2\right) \\
&= c_1L[y_1] + c_2L[y_2] = 0
\end{aligned}
$$

이므로 $c_1y_1 + c_2y_2$ 또한 $L[y] = 0$의 해이다. □

정리 1.3. *비제차 선형 상미분방정식 $L[y] = r(x)$의 해는 $L[y] = 0$의 일반해(general solution) y_h와 $L[y] = r(x)$를 만족하는 특수해(particular solution) y_p의 합으로 나타낼 수 있다.*

예시. 미분방정식

$$y' = y$$

를 풀기 위해, y에 관한 항을 좌변으로 이동한다:

$$\frac{y'}{y} = 1$$

이때 $y \neq 0$을 가정하였다. 양변을 x에 대해서 적분하면

$$\ln|y| = x + C \quad C \in \mathbb{R}$$

이 되고, e의 지수를 취하면

$$|y| = e^C e^x = ke^x \qquad k \in \mathbb{R}_{>0}$$

이 된다. 절댓값을 없애주면

$$y = ke^x \qquad k \in \mathbb{R} \setminus \{0\}$$

가 되는데, 위에서 $y \neq 0$일 경우를 제외하였다. $y = 0$의 경우도 원래 미분방정식 $y' = y$의 해를 만족하므로, 최종적으로 해는

$$y = ke^x \qquad k \in \mathbb{R}$$

이 된다. 이처럼 한 변수를 한쪽에 모두 몰아 미분방정식을 푸는 방식을 변수분리법 (separation of variables) 이라고 하며, 이러한 방식으로 풀 수 있는 미분방정식을 분리 가능하다 (separable) 고 한다.

정의 1.4. 함수 $u(x, y)$ 의 전미분 (total derivative) $\mathrm{d}u$ 를 제1미분형식 (differential 1-form) 으로 나타내면

$$\mathrm{d}u = u_x \, \mathrm{d}x + u_y \, \mathrm{d}y$$

와 같다. 이때 u_x 와 u_y 는 각각 x 와 y 에 대한 u 의 편미분 (partial derivative) 이다.

참고. 함수 $u(x, y)$ 의 전미분은 벡터장

$$\nabla u = (u_x, u_y)$$

와 대응된다.

정의 1.5. 함수 $u(x, y)$ 의 전미분

$$\mathrm{d}u = u_x \, \mathrm{d}x + u_y \, \mathrm{d}y$$

에 대해서 $u(x, y)$ 가 상수 함수이면 $\mathrm{d}u = 0$ 이 되어,

$$P(x, y) \, \mathrm{d}x + Q(x, y) \, \mathrm{d}y = 0$$

의 꼴로 나타낼 수 있다. 이를 완전미분방정식 (exact differential equation) 이라고 한다. 이때 $u(x, y)$ 를 잠재함수 (potential function) 라고 한다.

예시. $u : \mathbb{R}^2 \to \mathbb{R}$ 를

$$u(x, y) = x^2 + xy - \cos y = c$$

로 정의하자. 이때 y 는 x 의 함수이다. 이를 x 로 미분하면,

$$2x + y + x\frac{\mathrm{d}y}{\mathrm{d}x} + \sin y \frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

이 되고, 정리하여

$$(2x + y) + (x + \sin y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

이 된다. 미분형식으로 나타낸다면

$$(2x + y) \, \mathrm{d}x + (x + \sin y) \, \mathrm{d}y = 0$$

으로 쓸 수 있다.

역으로

$$(2x+y)\,\mathrm{d}x + (x+\sin y)\,\mathrm{d}y = 0$$

를 풀고자 한다면,

$$\begin{aligned} u_x &= 2x+y \\ u_y &= x+\sin y \end{aligned}$$

인 $u(x,y)$를 찾으면 된다. u_x를 x로 적분하면

$$u = x^2 + xy + h(y)$$

가 된다. 이때 $h(y)$는 y에 대한 함수로, 적분상수이다. 이를 다시 y로 미분하여 원래 식의 u_y랑 비교하자.

$$\frac{\partial u}{\partial x} = x + h'(y)$$

에서 $h'(y) = \sin y$가 되므로,

$$h(y) = -\cos y + C \qquad C \in \mathbb{R}$$

이다. 따라서 해 $u(x,y)$는

$$u(x,y) = x^2 + xy - \cos y + C = 0 \qquad C \in \mathbb{R}$$

꼴임을 알 수 있다. 처음의 식과 상수 차이만 나는 것을 볼 수 있다.

위 예시의 미분방정식이 잘 풀린 이유는

$$\frac{\partial}{\partial y}(2x+y) = \frac{\partial}{\partial x}(x+\sin y)$$

이기 때문이다. 아래의 예시의 미분방정식은 원래 꼴로는 위 방법처럼 풀 수 없다.

예시. 미분방정식

$$-y\,\mathrm{d}x + x\,\mathrm{d}y = 0$$

은

$$\frac{\partial}{\partial y}(-y) \neq \frac{\partial x}{\partial x}$$

이므로 완전미분방정식이 아니다. 그러나 아래와 같이

$$\frac{-y}{x^2+y^2}\,\mathrm{d}x + \frac{x}{x^2+y^2}\,\mathrm{d}y = 0$$

과 같이 x^2+y^2으로 나누어주면 완전미분방정식의 꼴이 되어

$$u = \arctan\frac{y}{x} = c \qquad c \in \mathbb{R}$$

로 풀 수 있다. 이는 곧

$$\frac{y}{x} = c,$$

즉

$$y = kx \qquad k \in \mathbb{R}$$

이다.

위 예시의 $1/(x^2+y^2)$와 같이 곱하여 완전미분방정식의 꼴이 되도록 하는 항을 적분인자 (integrating factor) 라고 한다. 미분방정식

$$P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y = 0$$

의 적분인자를 찾기 위해서는 먼저

$$P_y - Q_x = 0$$

인지를 확인해야 한다. 만약 $P_y - Q_x$가 0이라면 완전미분방정식이므로 바로 잠재함수 $u(x,y)$를 찾을 수 있다. 그렇지 않은 경우에만 적분인자를 찾으면 된다.

어떤 적분인자 F에 대해서,

$$FP(x,y)\,\mathrm{d}x + FQ(x,y)\,\mathrm{d}y = 0$$

이 완전미분방정식, 즉

$$\frac{\partial(FP)}{\partial y} - \frac{\partial(FQ)}{\partial x} = 0$$

이 되어야 한다. 전개하면

$$F_yP + FP_y = F_xQ + FQ_x$$

가 된다. 이때 F가 x만의 함수인 경우와 y만의 함수인 경우로 나눌 수 있다.

1. F가 x만의 함수인 경우:
   $F_y = 0$이므로
   $$FP_y = F'Q + FQ_x$$

가 된다. 이는 변수분리를 통해

$$\frac{F'}{F} = \frac{1}{Q}(P_y - Q_x)$$

로 변형할 수 있고[11], 적분하면

$$\ln|F| = \int \frac{1}{Q}(P_y - Q_x)\,\mathrm{d}x$$

가 되어 F는

$$F = \exp\left(\int \frac{1}{Q}(P_y - Q_x)\,\mathrm{d}x\right)$$

로 하나를 정할 수 있다.

2. F가 y만의 함수인 경우:
   $F_x = 0$이므로

$$\frac{\mathrm{d}F}{\mathrm{d}y}P + FP_y = FQ_x$$

가 된다. 이는 변수분리를 통해

$$\frac{1}{F}\frac{\mathrm{d}F}{\mathrm{d}y} = \frac{1}{P}(Q_x - P_y)$$

로 변형할 수 있고, 적분하면

$$\ln|F| = \int \frac{1}{P}(Q_x - P_y)\,\mathrm{d}y$$

가 되어 F는

$$F = \exp\left(\int \frac{1}{P}(Q_x - P_y)\,\mathrm{d}y\right)$$

로 하나를 정할 수 있다.

예시. 미분방정식

$$(e^{x+y} + ye^y)\,\mathrm{d}x + (xe^y - 1)\,\mathrm{d}y = 0$$

의 적분 인자를 구하자. 우선

$$\frac{\partial}{\partial y}\left(e^{x+y} + ye^y\right) - \frac{\partial}{\partial x}(xe^y - 1) = e^{x+y} + e^y + ye^y - e^y = e^{x+y} + ye^y \neq 0$$

[11] u를 구하려면 당연히 $F \neq 0$이어야 한다.

이므로, 완전상미분방정식이 아니다.

$$\frac{e^{x+y} + ye^y}{xe^y - 1}$$

은 x만의 함수가 아니다.

$$\frac{e^{x+y} + ye^y}{e^{x+y} + ye^y} = -1$$

은 상수 함수이므로 y만의 함수라고 할 수 있다. 따라서 적분 인자 F는

$$F = \exp\left(\int -1 \, \mathrm{d}y\right) = e^{-y}$$

가 되어, 원래 미분방정식을 완전상미분방정식

$$(e^x + y)\, \mathrm{d}x + (x - e^{-y})\, \mathrm{d}y = 0$$

으로 변형할 수 있다. 따라서 잠재함수 $u(x, y)$는

$$u(x, y) = e^x + xy + e^{-y}$$

가 된다.

1계 선형 상미분방정식

$$y' + p(x)y = r(x)$$

을 생각해보자. 이때

1. 제차 미분방정식인 경우, 즉 $r(x) = 0$인 경우 변수분리법을 통해

   $$\frac{y'}{y} = -p(x)$$

   로 변형할 수 있다. 이를 풀면

   $$y = k \exp\left(\int -p(x)\, \mathrm{d}x\right) \qquad k \in \mathbb{R}$$

   이다.

2. 비제차 미분방정식인 경우, 즉 $r(x) \neq 0$인 경우, 완전상미분방정식인 경우 바로 잠재함수를 구한다. 그렇지 않을 경우, 적분인자를 구하여 완전상미분방정식으로 꼴을 변형할 수 있다.

예시. 1계 선형 비제차 상미분방정식

$$y' + (\tan x)y = \sin(2x)$$

는 정리 1.3 제차 상미분방정식

$$y' + (\tan x)y = 0$$

의 해 y_h와 원래 미분방정식의 특수해 y_p의 합으로 나타낼 수 있다. 만약

$$y(0) = 1$$

와 같은 초기값(initial condition)이 주어진다면, 일반적으로 유일한 해를 구할 수 있다. 이처럼 초기값이 주어지는 문제를 초기값 문제(initial value problem)라고 부른다.

정의 1.6. 1계 비선형 상미분방정식

$$y' + p(x)y = g(x)y^a$$

의 꼴을 가지는 미분방정식을 베르누이 방정식(Bernoulli equation)이라고 부른다.

베르누이 방정식은

$$u(x) = y(x)^{1-a}$$

를 새로 정의하여 선형 상미분방정식으로 변형할 수 있다. 양변을 x에 대해 미분하면

$$u' = (1-a)y^{-a}y'$$

가 되는데, 원래 미분방정식에서 y'을 대입하면

$$u' = (1-a)y^{-a}(gy^a - py) = (1-a)\left(g - py^{1-a}\right) = (1-a)(y - pu)$$

이 되어 정리하면

$$u' + (1-a)pu = (1-a)g$$

로, 선형 상미분방정식이 된다.

정의 1.7. 1. 맬더스(Malthus)는 시각 t에 따른 인구y를 다음과 같이 나타내었다:

$$y' = ky$$

위 미분방정식을 맬더스의 인구방정식이라고 한다. 이를 풀면

$$y = Ae^{kt}$$

이 되는데, 무한히 증가하는 문제가 있다.

2. 벨허스트 (Velhulst) 는 이를 보정하여

$$y' = ky - cy^2$$

꼴로 변형하였다. 위 방정식은 베르누이 방정식의 일종이다. 이처럼 $y' = f(y)$ 의 꼴이 되어 독립변수에 직접적으로 의존하지 않는 미분방정식을 자율미분방정식 (autonomous differential equation) 이라고 부른다.