프리드버그 선형대수학 요약본

Junhyeok Lee (wnsx0000@gmail.com)

August 18, 2024

목차

Part I

벡터공간

1장은 벡터공간에 대한 이야기임.

1. 벡터공간

1. 벡터공간(vector space)

1) 정의

Definition 1. *체 F에서의 벡터공간(vector space) 또는 선형공간(linear space) V는 다음 8가지 조건을 만족시키는 두 연산, 합과 스칼라 곱을 가지는 집합임.*

- *합(sum)은 V의 두 원소 x, y에 대하여 유일한 원소 $x+y \in V$를 대응하는 연산임.*
- *스칼라 곱(scalar multiplication)은 체 F의 원소 a와 벡터공간 V의 원소 x마다 유일한 원소 $ax \in V$를 대응하는 연산이다. 이때 ax는 a와 x의 스칼라 곱(product)이라 함.*

(VS1) 모든 $x, y \in V$에 대하여 $x+y = y+x$임. (덧셈의 교환법칙)
(VS2) 모든 $x, y, z \in V$에 대하여 $(x+y)+z = x+(y+z)$임. (덧셈의 결합법칙)
(VS3) 모든 $x \in V$에 대하여 $x+0 = x$인 $0 \in V$가 존재함. (덧셈에 대한 항등원, 즉 영벡터 존재)
(VS4) 각 $x \in V$마다 $x+y = 0$인 $y \in V$가 존재함. (덧셈에 대한 역원, 즉 역벡터 존재)
(VS5) 각 $x \in V$에 대하여 $1x = x$임. (스칼라 곱에 대한 항등원 존재)
(VS6) 모든 $a, b \in F$와 모든 $x \in V$에 대하여 $(ab)x = a(bx)$임. (스칼라 곱에 대한 결합법칙)
(VS7) 모든 $a \in F$와 모든 $x, y \in V$에 대하여 $a(x+y) = ax+ay$임. (분배법칙)
(VS8) 모든 $a, b \in F$와 모든 $x \in V$에 대하여 $(a+b)x = ax+bx$임. (분배법칙)

2) 벡터공간의 표기

체 F에서의 벡터공간 V는 $F-$벡터공간 V라고 표기함.

혼동의 여지가 없는 경우 체 F는 생략하기도 함.

3) 벡터와 스칼라

벡터공간 V의 원소를 벡터, 체 F의 원소를 스칼라라 함.

여기서의 벡터는 벡터공간의 원소를 가리키는 일반적인 개념임.
지금까지 단순 화살표로 표현해 온 벡터를 포괄하는 의미.

(VS3)을 만족하는 유일한 벡터 0은 V의 영백터(zero vector)라 함.

(VS4)를 만족하는 유일한 벡터 $y(-x)$는 덧셈에 대한 x의 역벡터(additive inverse)라고 함.

2. 관련 정리

1) 벡터 합의 소거법칙

Theorem 1.1 $x, y, z \in V$이고 $x + z = y + z$ 일 때, $x = y$ 임.

Corollary 1 (VS3)을 만족하는 벡터 0(영백터)은 유일함.

Corollary 2 (VS4)를 만족하는 벡터 y(역벡터)는 유일함.

2) 스칼라 곱 관련 성질[1]

Theorem 1.2 모든 벡터공간 V에 대해서 다음이 성립함.

1. 모든 벡터 x에 대하여 $0x = \vec{0}$임.
2. 모든 스칼라 a와 모든 벡터 x에 대하여 $(-a)x = -(ax) = a(-x)$임.
3. 모든 스칼라 a에 대하여 $a\vec{0} = \vec{0}$임.

3. 벡터공간의 예시[2]

1) n순서쌍(n-tuple)의 집합(F^n)

체 F에서 성분을 가져온 n순서쌍의 집합을 F^n이라 표기함.

$u = (a_1, a_2, ..., a_n) \in F^n$, $v = (b_1, b_2, ..., b_n) \in F^n$, $c \in F$ 일 때, 합과 스칼라 곱을 아래와 같이 정의하면 이 집합은 F-벡터공간임.

$$u + v = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n),\ cu = (ca_1, ca_2, ..., ca_n)$$

F^n의 벡터는 행벡터(row vector)보다 열벡터(column vector)로 주로 표현함.

F^1은 그냥 F로 표현하는 경우가 많음.

2) 행렬(matrix)의 집합($M_{m\times n}(F)$)

성분이 체 F의 원소인 모든 $m \times n$ 행렬의 집합은 $M_{m\times n}(F)$라 정의함.

$A, B \in M_{m\times n}(F)$, $c \in F$ 일 때, 합과 스칼라 곱을 아래와 같이 정의하면 이 집합은 F-벡터공간임.

$$(A + B)_{ij} = A_{ij} + B_{ij},\ (cA)_{ij} = cA_{ij}$$

3) 함수(function)의 집합($\mathcal{F}(S, F)$)

체 F의 공집합이 아닌 부분집합 S가 있을 때, $\mathcal{F}(S, F)$는 S에서 F로 가는 모든 함수의 집합임.

$\mathcal{F}(S, F)$에서 모든 $s \in S$에 대하여 $f(s) = g(s)$일 때, 두 함수 f, g는 같다고 함.

$f, g \in \mathcal{F}(S, F)$, $c \in F$, $s \in S$ 일 때, 합과 스칼라 곱을 아래와 같이 정의하면 이 집합은 F-벡터공간임.

$$(f + g)(s) = f(s) + g(s),\ (cf)(s) = c(f(s))$$

실수집합 R에서 R로 가는 모든 연속함수의 집합을 $C(R)$이라 함.

---

[1]스칼라, 벡터, 영벡터 사이의 곱에 대한 정리.
[2]아래의 예시들은 각 성분별로(component-wise) 계산이 가능하기 때문에 매번 8가지 조건들을 검토할 필요가 없음.

4) 다항식(polynominal)의 집합($P(F)$)

체 F에서 계수를 가져온 모든 다항식의 집합을 $P(F)$라 함.

두 다항식의 합과 스칼라 곱을 아래와 같이 정의하면 이 집합은 F-벡터공간임.

$$f(x) + g(x) = (a_n + b_n)x_n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$
$$cf(x) = ca_n x_n + ca_{n-1}x^{n-1} + \cdots + ca_1 x + ca_0$$

음이 아닌 정수 n에 대하여 $P_n(F)$는 n 이하의 차수를 가진 다항식의 집합임.

5) 수열의 집합

체 F에서 정의된 모든 수열의 집합을 V라 할 때, $t \in F$와 두 수열 (a_n), (b_n)에 대해서 합과 스칼라 곱을 정의하면 이 집합은 F-벡터공간임.

$$(a_n) + (b_n) = (a_n + b_n),\ t(a_n) = (ta_n)$$

2. 부분공간(subspace)

1. 부분공간[3]

1) 정의

> **Definition 2.** *F-벡터공간 V의 부분집합 W가 V에서 정의한 합과 스칼라 곱을 가진 F-벡터공간일 때, W를 V의 부분공간이라 함.*

즉, 벡터공간인 부모의 연산을 그대로 물려받은 벡터공간인 부분집합.

임의의 벡터공간 V에 대하여 V와 $\{0\}$은 V의 부분공간임.
특히 $\{0\}$은 점공간인 부분공간(zero subspace)이라고 함.

2) 부분공간 판별

어떤 부분집합이 부분공간이기 위한 필요충분조건은 아래 4가지 성질을 만족하는 것임.
벡터공간의 8가지 조건을 생각해 보면 당연한 이야기.

1. 모든 $x \in W$, $y \in W$에 대하여 $x + y \in W$임. (W는 합에 대하여 닫혀 있음)
2. 모든 $c \in F$와 모든 $x \in W$에 대하여 $cx \in W$임. (W는 스칼라 곱에 대하여 닫혀 있음)
3. W는 영벡터를 포함함. (영벡터 존재)
4. W에 속한 모든 벡터 각각의 덧셈에 대한 역벡터는 W의 원소임. (역벡터 존재)

Theorem 1.3에 따르면 W와 V의 영벡터는 반드시 같고, 4번 성질은 굳이 확인할 필요가 없음(항상 성립)

2. 관련 정리

1) 부분공간 판별

Theorem 1.3 벡터공간 V와 그 부분집합 W에 대하여, W가 V의 부분공간이기 위한 필요충분조건은 아래의 세 가지 조건을 만족하는 것임.

1. $0 \in W$
2. 모든 $x \in W$, $y \in W$에 대하여 $x + y \in W$
3. 모든 $c \in F$와 모든 $x \in W$에 대하여 $cx \in W$

2) 부분공간의 교집합

Theorem 1.4 벡터공간 V의 부분공간들의 임의의 교집합은 V의 부분공간임.

[3]부분공간은 벡터공간의 일부분을 살펴보기 위해 존재함.

3. 일차결합과 생성공간

1. 일차결합(linear combination)

1) 정의

Definition 3. *벡터공간 V의 공집합이 아닌 부분집합 S가 있다고 하자. 유한개의 벡터 $u_1, u_2, \cdots, u_n \in S$와 스칼라 $a_1, a_2, \cdots, a_n$에 대하여 아래를 만족하는 벡터 $v \in V$를 S의 일차결합(linear combination)이라 함.*

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

이때, v는 벡터 $u_1, u_2, \cdots, u_n$의 일차결합이라 하고, $a_1, a_2, \cdots, a_n$은 이 일차결합의 계수(coefficient)라고 함.

영벡터는 공집합이 아닌 모든 부분집합의 일차결합임.

2. 생성공간(span)

1) 정의

Definition 4. *벡터공간 V의 공집합이 아닌 부분집합 S에 대해서, S의 벡터를 사용하여 만든 모든 일차결합의 집합을 S의 생성공간(span)이라 하고, $span(S)$라 표기함.*

편의를 위해 $span(\emptyset) = \{0\}$으로 정의함.[4]

2) 생성

벡터공간 V의 부분집합 S에 대하여 $span(S) = V$이면 S는 V를 생성한다(generate, span)고 하고, S를 V의 생성집합이라고 함.

S의 벡터가 V를 생성한다고 하기도 함.

어떤 집합의 생성공간이 어떤 벡터공간인지 확인하려면, 벡터공간의 임의의 벡터를 해당 집합의 원소들로 표현할 수 있는지 확인하면 됨.

4. 관련 정리

1) 생성공간과 부분공간 사이의 관계

Theorem 1.5 벡터공간 V의 임의의 부분집합 S의 생성공간은 S를 포함하는 V의 부분공간임. 또한 S를 포함하는 V의 부분공간은 반드시 S의 생성공간을 포함함.

즉, 벡터공간의 부분집합으로 만든 생성공간(span)은 부분공간임.
또한 부분공간의 부분집합으로 만든 생성공간(span)은 해당 부분공간에 포함됨.

[4] 점공간의 기저를 설정하기 위한 것.

4. 일차종속과 일차독립

1. 일차종속(linearly dependent)/일차독립(linearly independent)

1) 정의

Definition 5. *벡터공간 V의 부분집합 S에 대하여 $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$을 만족시키는 유한계의 서로 다른 벡터 $u_1, u_2, \cdots, u_n \in S$와 적어도 하나는 0이 아닌 스칼라 $a_1, a_2, \cdots, a_n$이 존재하면 집합 S는 일차종속(linearly dependent)라 함. 이때, S의 벡터 또한 일차종속이라 함.*

벡터공간 V의 부분집합 S가 일차종속이 아니면 일차독립(linearly independent)라 함. 이때. S의 벡터 또한 일차독립이라 함.

2) 영벡터의 자명한 표현

임의의 벡터 $u_1, u_2, \cdots, u_n$에 대하여 $a_1 = a_2 = \cdots = a_n = 0$이면 $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$임. 이를 $u_1, u_2, \cdots, u_n$의 일차결합에 대한 영벡터의 자명한 표현(trivial representation of 0)이라 함.

즉, 일차결합에서 스칼라에 전부 0을 넣어 표현한 영벡터를 의미함.

3) 일차독립인 집합에 대한 명제

일차독립인 집합에 대한 아래의 명제들은 모든 벡터공간에서 참임.

1. 공집합은 일차독립임.[5]
2. 영이 아닌 벡터 하나로 이루어진 집합은 일차독립임.
3. 어떤 집합이 일차독립이기 위한 필요충분조건은 영벡터를 주어진 집합에 대한 일차결합으로 표현하는 방법이 자명한 방법 뿐인 것임.

즉, 3번 명제를 사용하여 유한집합이 일차독립인지 확인할 수 있음.

4) 일차종속/일차독립이 가지는 의미

일차독립
= 영벡터를 자명한 표현으로만 나타낼 수 있음.
= 해당 집합의 모든 벡터가 다른 벡터들의 일차결합으로 표현되지 않음.

5) 일차종속/일차독립의 판정

1. 어떤 벡터가 다른 벡터들의 일차결합으로 표현되는지 확인.

2. 영벡터가 자명한 표현으로만 나타내지는지 확인.
순서쌍들의 집합 등에서는 연립일차방정식을 풀어서 확인할 수 있음.
3장의 행간소사다리꼴(RREF)에 대한 해석을 활용하여 확인할 수 있음.

[5] $span(\emptyset) = \{0\}$인 것과 마찬가지로, 점공간의 기저를 설정하기 위한 것.

2. 관련 정리

1) 일차종속/일차독립과 집합관계

Theorem 1.6 V는 벡터공간이고 $S_1 \subseteq S_2 \subseteq V$일 때, S_1이 일차종속이면 S_2도 일차종속임.

Corollary 1 V는 벡터공간이고 $S_1 \subseteq S_2 \subseteq V$일 때, S_2가 일차독립이면 S_1도 일차독립임.

2) 일차독립과 벡터의 유일한 표현

Theorem 1.7 벡터공간 V과 일차독립인 V 부분집합 S가 있음. S에 포함되지 않는 벡터 $v \in V$에 대하여, $S \cup \{v\}$가 일차종속이기 위한 필요충분조건은 $v \in span(S)$임.

즉, 일차독립인 집합 S의 원소들을 일차결합하여 만들 수 있는 벡터가 S에 추가된 집합은 일차종속임. 반대로, 일차결합하여 만들 수 없는 벡터가 추가되면 여전히 일차독립임.

Proof. 1. $S \cup \{v\}$가 일차종속 $\rightarrow v \in span(S)$
$S \cup \{v\}$가 일차종속이면 $u_1, u_2, \cdots, u_n \in S$와 스칼라 $a_1, a_2, \cdots, a_{n+1}$에 대해서 $a_1u_1 + a_2u_2 + \cdots + a_nu_n + a_{n+1}v = 0$, $a_{n+1} \neq 0$이 성립함. 정리하면 $v = -\frac{1}{a_{n+1}}(a_1u_1 + a_2u_2 + \cdots + a_nu_n + a_{n+1}v = 0)$이므로 $v \in span(S)$임.

2. $v \in span(S) \rightarrow S \cup \{v\}$가 일차종속 $v \in span(S)$이므로 $v = a_1u_1 + a_2u_2 + \cdots + a_nu_n + a_{n+1}v = 0$으로 표현할 수 있고, 정리하면 $a_1u_1 + a_2u_2 + \cdots + a_nu_n + (a_{n+1} - 1)v = 0$으로 자명적이지 않은 표현이 존재하므로 $S \cup \{v\}$는 일차종속임. □

5. 기저와 차원

1. 기저(basis)

1) 정의

Definition 6. *벡터공간 V와 부분집합 β에 대해서, β가 일차독립이고 V를 생성하면 β를 V의 기저 (basis)라 함.*

즉, V의 기저인 β는 V를 생성하고 일차독립임.

기저는 유한집합이 아닐 수 있음.[6]

2) 표준기저

벡터공간 F^n의 벡터 $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, \cdots, 0), \cdots, e_n = (0, \cdots, 1)$에 대하여, 집합 $\{e_1, e_2, \cdots, e_n\}$은 F^n의 표준기저임.
여기서의 e_n은 임의로 잡은 기호가 아니라, 벡터공간 F^n의 표준기저를 일반적으로 나타내는 기호임.

표준기저는 기본적으로 n순서쌍(n-tuple)에 대한 개념이지만 다른 벡터공간에서도 비슷한 기저를 생각해 볼 수 있음. 집합 $\{1, x, x^2, \cdots, x^n\}$은 벡터공간 $P_n(F)$의 기저임. 행렬 $E^{ij} \in M_{m\times n}(F)$는 i행 j열 성분만 1이고 나머지 성분은 0인 행렬이라 하면 이는 $M_{m\times n}(F)$의 기저임.

2. 차원(dimension)

1) 정의

Definition 7. *V의 기저가 n개의 벡터로 이루어질 때, 유일한 자연수 n은 V의 차원(dimension)이고, $dim(V)$라 표기함. 기저가 유한집합인 벡터공간을 유한차원(finite dimension)이라 하고, 유한차원이 아닌 벡터공간을 무한차원(infinite dimension)이라 함.*

대체정리의 Corollary 1에서 알 수 있듯이, 기저를 형성하는 벡터의 개수는 벡터공간 V의 본질적 성질임.

2) 예시

벡터공간 $\{0\}$의 차원은 0임.
벡터공간 F^n의 차원은 n임.
벡터공간 $M_(m \times n)(F)$의 차원은 mn임.
벡터공간 $P_n(F)$의 차원은 $n+1$임.
벡터공간 $P(F)$는 무한차원임.

3. 관련 정리

1) 부분집합이 기저가 되기 위한 필요충분조건

Theorem 1.8 벡터공간 V의 부분집합 $\beta = \{u_1, u_2, \cdots, u_n\}$가 V의 기저이기 위한 필요충분조건은, 임의의 벡터 $v \in V$를 β에 속한 벡터의 일차결합으로 나타낼 수 있고 그 표현이 유일한 것임.

즉, 기저 β는 유일한 일차결합으로 V의 벡터[7]를 표현할 수 있고, β의 유일한 일차결합으로 V의 벡터가 표현된다면 β는 기저인 것.

[6] ex. 집합 $\{1, x, x^2, \cdots\}$은 $P(F)$의 기저임.
[7] β의 벡터가 아닌 V의 벡터인 것 유의.

2) 생성집합과 기저의 관계[8]

Theorem 1.9 유한집합 S가 벡터공간 V를 생성하면, S의 부분집합 중 V의 (유한집합인) 기저가 존재함.

Proof. $S = \emptyset$ 또는 $S = \{0\}$인 경우, $V = \{0\}$임. $\emptyset$은 일차독립이므로, S의 부분집합이면서 V의 기저임.

S가 영벡터가 아닌 벡터 u_1을 가지고 있다고 가정하면, $\{u_1\}$은 일차독립인 S의 부분집합임. 집합 $\{u_1, u_2, \cdots, u_k\}$가 일차독립이 되도록 S에서 순차적으로 $u_2, \cdots, u_k$를 꺼내서 추가하는 것을 유한 번 반복함. 최종적으로 얻은 일차독립인 집합을 $\beta = \{u_1, u_2, \cdots, u_k\}$라 함.

S에서 일차독립인 부분집합을 추출했으니, 이제 이 부분집합이 V를 생성하는지 확인해야 함.
(i) $\beta = S$인 경우. S는 일차독립이고 V를 생성하므로 V의 기저임.
(ii) β가 S의 일차독립인 진부분집합인 경우. $span(\beta)$는 V의 부분공간이고, S를 포함하는 V의 부분공간은 $span(S)$(즉, V.) 또한 포함하므로[9], $S \subseteq span(\beta)$인지 증명하면 β는 V를 생성한다고 할 수 있음.

$v \in S$에 대하여 $v \in \beta$이면 당연히 $v \in span(\beta)$임. $v \notin \beta$이면 $\beta \cap \{v\}$는 β의 구성 방식에 의해 일차종속임. $\beta \cap \{v\}$가 일차종속이므로 $v \in span(\beta)$임[10]. 따라서 $S \subseteq span(\beta)$임. □

3) 대체정리(replacement theorem)[11]

Theorem 1.10 n개의 벡터로 이루어진 집합 G가 벡터공간 V를 생성한다고 하자. L이 m개의 벡터로 이루어진 일차독립인 V의 부분집합이면, $m \leq n$임. 또한 다음 조건을 만족시키는 $H \subseteq G$가 존재함. H는 $n - m$개의 벡터로 이루어졌으며, $L \cup H$는 V를 생성함.

Corollary 1 벡터공간 V가 유한집합인 기저를 포함한다고 가정하면, V의 모든 기저는 유한집합이며 같은 개수의 벡터로 이루어져 있음.

Proof. β가 n개의 벡터로 이루어진 V의 기저이고, γ가 또 다른 V의 기저라고 하자. γ가 $n+1$개의 벡터로 이루어져 있다고 하면, γ는 일차독립이고 β는 V를 생성하므로 대체정리에 의해 $n+1 < n$이 성립해야 하는데 이는 모순임. 즉, γ가 m개의 벡터로 이루어져 있다면 $m \leq n$임. β와 γ를 바꾸어 똑같은 논리를 반복하면 $n \leq m$이므로 $m = n$임. 즉, V의 모든 기저는 같은 개수의 벡터로 이루어져 있음. □

Corollary 2 V를 차원이 n인 벡터공간이라 하자.
(1) V의 유한 생성집합에는 반드시 n개 이상의 벡터가 있음. 또한 n개의 벡터로 이루어진 V의 생성집합은 V의 기저임.
(2) 일차독립이고 n개의 벡터로 이루어진 V의 부분집합은 V의 기저임.
(3) 집합 $L \subset V$가 일차독립이면 $L \subseteq \beta$인 기저 β가 존재함. 즉, 일차독립인 V의 부분집합을 확장시켜 기저를 만들 수 있음.

3번 정리는 일차독립인 집합으로 원하는 기저를 생성할 수 있다는 의미.

3장에서 등장하는 행간소사다리꼴(RREF)의 해석을 사용하면 일차독립인 집합으로 기저를 실제로 생성할 수 있음.

정리하면, 유한차원 벡터공간 V에서 $dim(V) = n$이라 할 때, 일차독립인 부분집합은 벡터의 개수가 n보다 클 수 없고, V를 생성하는 부분집합은 n보다 작을 수 없음. 기저는 일차독립인 집합의 집합과 생성집합의 집합의 교집합으로, 그 크기가 n임.

---

[8] 아래의 증명 방식을 눈여겨보자. 특히 일차독립인 집합을 생성하고, 그 집합이 벡터공간을 생성하는지 확인하는 방식에 집중할 것.
[9] 정리 1.5
[10] 정리 1.7
[11] 대체정리는 수학적 귀납법으로 증명하지만 그 내용은 정리하지 않음. 여기서는 Corollary 1에 집중하자.

4) 부분공간의 차원

Theorem 1.11 유한차원 벡터공간 V에 대하여 부분공간 W는 유한차원이고, $dim(W) \le dim(V)$임. 특히, $dim(W) = dim(V)$이면 $V = W$임.

Proof. $dim(V) = n$이라 하자. $W = \{0\}$이면 W는 유한차원이고 $dim(W) = 0 \le n$임. $W \ne \{0\}$ 인 경우, W는 영벡터가 아닌 벡터 x_1을 가지고, $\{x_1\}$은 일차독립임. $\{x_1, x_2, \cdots x_k\}$이 일차독립이 되도록 W에서 $x_1, x_2, \cdots, x_k$를 꺼내자. V의 일차독립인 부분집합은 n개 이상의 벡터를 가질 수 없으므로, 이 과정은 $k \le n$인 범위 안에서 끝남. 이때, $\{x_1, x_2, \cdots x_k\}$는 일차독립이고 W에서 벡터를 더 추가하면 일차종속이 되고, 해당 벡터는 $span(\{x_1, x_2, \cdots x_k\})$의 원소임. 즉, $\{x_1, x_2, \cdots x_k\}$는 W의 기저이고, $dim(W) = k \le n$임.

$dim(W) = n$이면, W의 기저는 n개의 벡터로 이루어지고 일차독립인 V의 부분집합임. 대체정리의 Corollary 2에 의해 이 집합은 V의 기저임. 즉, $W = V$임. □

Corollary 1 유한차원 벡터공간 V의 부분공간 W에 대해서, W의 임의의 기저를 확장하여 V의 기저를 얻을 수 있음.

Proof. W의 임의의 기저는 일차독립인 V의 부분집합이므로, 대체정리의 Corollary 2에 의해 이를 확장시켜 V의 기저를 만들 수 있음. □

Part II

선형변환과 행렬

2장에서는 처음부터 끝까지 선형변환이 곧 행렬이고 행렬이 곧 선형변환이라는 이야기를 함.

1. 선형변환

1. 선형변환(linear transformation)

벡터공간의 벡터 합과 스칼라 곱을 보존하는 함수.[12]

1) 정의

> **Definition 8.** *F-벡터공간 V와 W가 있다고 하자. 모든 $x, y \in V$, $c \in F$에 대하여 아래의 두 조건을 만족하는 함수 $T : V \to W$를 V에서 W로 가는 선형변환(linear transformation) 또는 선형(linear), 선형사상(linear map)이라 함.*
>
> 1. $T(x+y) = T(x) + T(y)$
> 2. $T(ax) = aT(x)$

2) 선형의 성질

성질1. T가 선형이면 $T(0) = 0$임.

성질2. T가 선형이기 위한 필요충분조건은 모든 $x, y \in V$, $c \in F$에 대해서 $T(cx+y) = cT(x) + T(y)$인 것임.

성질3. T가 선형이면 모든 $x, y \in V$에 대해서 $T(x-y) = T(x) - T(y)$임.

성질4. T가 선형이기 위한 필요충분조건은 모든 $x_1, x_2, \cdots, x_n \in V$와 $a_1, a_2, \cdots, a_n \in F$에 대해서 아래의 식을 만족하는 것임.

$$T(\sum_{i=1}^{n} a_i x_i) = \sum_{i=1}^{n} a_i T(x_i)$$

어떤 함수가 선형인지 보일 때, 주로 성질2를 사용함.

3) 항등변환(identity transformation)과 영변환(zero transformation)

항등변환 $I_V : V \to V$는 모든 $x \in V$에 대해서 $I_V(x) = x$라 정의되는 선형임. 간단히 I라 표기하기도 함.
영변환 $T_0 : V \to W$는 모든 $x \in V$에 대해서 $T_0(x) = 0$이라 정의되는 선형임.

4) 선형연산자(linear operator)

정의역과 공역이 같은 선형변환. 다시 말해, 벡터공간 V에서 V로 가는 선형변환.

5) 선형의 예시

대표적인 선형 기하 변환으로는 회전, 대칭, 사영(한 값을 0으로 만드는 것)이 있음.
행렬을 전치행렬로 변환하는 함수, 미분, 적분 등 또한 선형임.

[12] 연산을 하고 보내나 보내고 연산을 하나 똑같으므로, 보내기 전의 연산과 보낸 후의 연산이 동일한 기능을 하고 있다는 것임.

2. 영공간(null space, kernel)과 상공간(range, image)

1) 정의

Definition 9. *벡터공간 V, W와 선형변환 $T : V \to W$가 있다고 하자.*

영공간(null space, kernel)은 $T(x) = 0$인 $x \in V$를 원소로 가지는 집합임. $N(T)$로 표기함.
$N(T) = \{x \in V \,|\, T(x) = 0\}$임.

상공간(range, image)은 T의 함숫값을 원소로 가지는 W의 부분집합임. $R(T)$라 표기함.
$R(T) = \{T(x) \,|\, x \in V\}$

3. nullity와 rank

1) 정의

Definition 10. *벡터공간 V, W와 선형변환 $T : V \to W$에 대해서 $N(T)$와 $R(T)$가 유한차원이라 가정하자.*

- *$N(T)$의 차원을 nullity(영공간의 차원)라 하고, nullity(T)라 표기함.*
- *$R(T)$의 차원을 rank(랭크, 계수)라 하고, rank(T)라 표기함.*

4. 관련 정리

1) 영공간/상공간은 부분공간임

Theorem 2.1 벡터공간 V, W와 선형변환 $T : V \to W$에 대하여 $N(T)$, $R(T)$는 각각 V, W의 부분공간임.

Proof. $T(0) = 0$이므로 $0 \in N(T)$임. $x, y \in N(T)$, $c \in F$에 대해서, $T(x + y) = 0$, $T(cx) = 0$이므로 합과 스칼라 곱에 대해 닫혀 있음. 그러므로 $N(T)$는 V의 부분공간임.

$T(0) = 0$이므로 $0 \in R(T)$임. $x, y \in N(T)$, $c \in F$에 대해서, $T(v) = x, T(w) = y$인 $v, w \in V$가 존재함. $T(v + w) = x + y$, $T(cv) = cx$이므로 합과 스칼라 곱에 대해서 닫혀 있음. 그러므로 $R(T)$는 V의 부분공간임. □

2) 선형변환의 상공간 생성 방법

Theorem 2.2 벡터공간 V, W와 선형변환 $T : V \to W$, 그리고 V의 기저 $\beta = \{v_1, v_2, \cdots, v_n\}$에 대해서 아래의 식이 성립함.

$$R(T) = span(T(\beta)) = span(\{T(v_1), T(v_2), \cdots, T(v_n)\})$$

즉, V의 기저를 보내서 span하면 전체가 생성됨.

Proof. 모든 $T(\beta)$는 W의 부분집합이고 $R(T)$는 부분공간이므로, $span(T(\beta)) \subseteq R(T)$.[13]

$w \in R(T)$이면 $w = T(v)$인 $v \in V$가 존재함. β는 V의 기저이므로, 아래와 같이 쓸 수 있음.

$$v = \sum_{i=1}^{n} a_i v_i \, (a_1, a_2, \cdots, a_n \in F)$$

T가 선형이므로, $w = T(v) = T(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i T(v_i) \in span(T(\beta))$임. 즉, $R(T) \subseteq span(T(\beta))$임.

$span(T(\beta)) \subseteq R(T)$이고 $R(T) \subseteq span(T(\beta))$이므로 $R(T) = span(T(\beta))$임. □

[13]정리 1.5

3) 차원정리(dimension theorem)

Theorem 2.3 벡터공간 V, W와 선형변환 $T : V \to W$에 대하여 V가 유한차원이면 아래의 식이 성립함.

$$nullity(T) + rank(T) = dim(V)$$

선형변환 T에 대해서, V의 기저 중 일부는 0으로 가고 나머지는 상공간으로 간다는 의미.

아래의 다이어그램은 벡터공간 V, W와 선형변환 $T : V \to W$에 대해서, dimension theorem에서 기저가 어떻게 이동하는지를 보여줌.

$$\begin{tikzcd}
\mathrm{V} \arrow[d, dash] \arrow[dr, "T"] & \mathrm{W} \arrow[d, dash] \\
\mathrm{N(T)} \arrow[d, dash] \arrow[dr, "T"] & \mathrm{R(T)} \arrow[d, dash] \\
0 \arrow[r, "\mathrm{T}"] & 0
\end{tikzcd}$$

Proof. 차원정리의 증명은 4가지 단계로 이루어짐.

1. $kernel$의 기저 구성
$dim(V) = n$, $nullity(T) = k$라 하고, $N(T)$의 기저를 $\alpha = \{v_1, v_2, \cdots v_k\}$라 하자.

2. V의 기저 구성
Theorem 1.11의 Corollary에 의해 $N(T)$의 기저를 확장하여 V의 기저를 만들 수 있음. V의 기저는 $\beta = \{v_1, v_2, \cdots v_k, v_{k+1}, \cdots, v_n\}$이라 하자.

3. 주장
$S = \{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}$이 $R(T)$의 기저임을 보이면 증명이 완료됨.

4-1. 일차독립인지 확인
$a_1, a_2, \cdots, a_n \in F$에 대해서 $\sum_{i=k+1}^{n} a_i T(v_i) = 0$이 성립한다고 하자. $\sum_{i=k+1}^{n} a_i T(v_i) = T(\sum_{i=k+1}^{n} a_i v_i) = 0$이므로 $\sum_{i=k+1}^{n} a_i v_i \in N(T)$임. 즉, 이는 $N(T)$의 기저로 나타낼 수 있음. $b_1, b_2, \cdots, b_n \in F$에 대해서 $\sum_{i=k+1}^{n} a_i v_i = \sum_{i=1}^{k} b_i v_i$, $-\sum_{i=1}^{k} b_i v_i + \sum_{i=k+1}^{n} a_i v_i = 0$인데, β가 일차독립이므로 자명한 표현만이 성립함.

4-2. $R(T)$를 생성하는지 확인
상공간은 정의역 벡터공간의 기저를 보내서 확장하면 얻을 수 있고, $T(v_1) = 0, T(v_2) = 0, \cdots, T(v_k) = 0$임. 그러므로 $R(T) = span(T(\beta)) = span(\{T(v_1), T(v_2), \cdots, T(v_n)\}) = span(\{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}) = span(S)$임. 즉, S는 $R(T)$를 생성함. □

4) 단사함수의 확인

Theorem 2.4 벡터공간 V, W와 선형변환 $T : V \to W$에 대해서, T가 단사함수이기 위한 필요충분조건은 $N(T) = \{0\}$인 것임.

Proof. T가 단사함수인 경우, $T(x) = 0$인 x는 0밖에 없으므로 $N(T) = \{0\}$임.

$N(T) = \{0\}$인 경우, $T(x) = T(y)$라고 가정하자. $T(x) - T(y) = 0$, $T(x - y) = 0$인데 $N(T) = \{0\}$이므로 $x - y = 0$, $x = y$임. 즉, T는 단사함수임. □

5) 차원이 같은 경우 동치인 명제들

Theorem 2.5 유한차원 벡터공간 V, W의 차원이 같을 때, 선형변환 $T : V \to W$에 대해서 아래의 세 명제는 동치임.

1. T는 단사(one-to-one, injection)임.
2. T는 전사(onto, surjection)임.
3. $rank(T) = dim(V)$

즉, 두 벡터공간의 차원이 같을 때 세 가지 명제 중 하나만 밝히면 나머지도 밝혀짐.

추가로, 이는 곧 해당 선형변환이 가역임을 의미함.

공역과 치역이 같은지를 확인함으로써 전사인지를 알 수 있음.

Proof. T가 단사임 $\Leftrightarrow N(T) = \{0\} \Leftrightarrow nullity(T) = 0 \Leftrightarrow rank(T) = dim(V) = dim(W) \Leftrightarrow dim(R(T)) = dim(W)$

$R(T)$는 W의 부분공간이므로 Theorem 1.11에 의해 차원이 같으면 $R(T) = W$임. 즉, T는 전사임. □

6) linear extension theorem

Theorem 2.6 F-벡터공간 V, W와 V의 기저 $\{v_1, v_2, \cdots, v_n\}$을 생각하자. 벡터 $w_1, w_2, \cdots, w_n \in W$에 대하여 아래의 조건을 만족시키는 선형변환 $T : V \to W$가 유일하게 존재함.

$$i = 1, 2, \cdots, n\text{에 대하여 } T(v_i) = w_i$$

즉, 정의역의 모든 값들을 보낼 필요 없이 기저의 원소만 보내 봐도 선형변환을 확정지을 수 있음. 기저에서 보낸 함숫값들의 조합은 선형변환에 따라 여러 가지가 있을 수 있는데, 각 조합에 대해서 선형변환이 유일하게 존재한다는 것.

선형변환을 정의하는 기본적인 방법. (기저만 보내면 충분함.)

선형변환의 행렬표현에서 기저만으로 선형변환을 정의하는 것은 linear extension theorem이 그래도 된다는 것을 보장하기 때문임.

Proof. $x \in V$는 $a_1, a_2, \cdots, a_n \in F$에 대해서 $x = \sum_{i=1}^{n} a_i v_i$로 유일하게 표현할 수 있음.

선형변환 $T : V \to W$를 $T(v_i) = w_i$, $T(x) = \sum_{i=1}^{n} a_i w_i$라 정의하자.[14]

(1) T는 선형인가?

$cT(v) + T(u) = T(cv + u)$가 성립함.

(2) $i = 1, 2, \cdots, n$에 대하여 $T(v_i) = w_i$임.

(3) T는 유일한가?[15]

선형변환 $U : V \to W$가 $i = 1, 2, \cdots, n$에 대하여 $U(v_i) = w_i$를 만족한다고 가정하자. $x = \sum_{i=1}^{n} a_i v_i \in V$에 대해서, $U(x) = \sum_{i=1}^{n} a_i w_i = T(x)$이므로 $U = T$임. □

Corollary 두 벡터공간 V, W에 대하여 V가 유한집합인 기저 $\{v_1, v_2, \cdots, v_n\}$를 포함한다고 가정하자. 두 선형변환 $U, T : V \to W$가 $i = 1, 2, \cdots, n$일 때, $U(v_i) = T(v_i)$를 만족하면 $U = T$임.

즉, 기저에서 보낸 함숫값들의 조합이 같으면 같은 선형변환임.

[14]이렇게 정의했을 때 해당 선형변환에 대해서 linear extension theorem이 성립하는지를 보는 것.
[15]정의한 $T(x)$가 유일한 선형변환임을 알 수 있음.

2. 선형변환의 행렬표현

1. 순서기저(ordered basis)

1) 정의

Definition 11. *유한차원 벡터공간 V의 순서기저(ordered basis)는 순서가 주어진 기저임. 즉, 순서기저는 일차독립이며 V를 생성하는 벡터들로 이루어진 수열임.*

같은 벡터로 이루어져 있더라도 순서가 다르면 다른 순서기저임.

2) 표준 순서기저(standard ordered basis)

벡터공간 F^n의 표준 순서기저는 $\{e_1, e_2, \cdots, e_n\}$임.

표준 순서기저는 기본적으로 F^n(순서쌍)에 대한 개념이지만, 다른 벡터공간에 대해서도 유사한 기저를 생각해 볼 수 있음. $P_n(F)$에서는 $\{1, x, \cdots, x_n\}$이, R^n에서는 $\{(1, 0, \cdots), \cdots, (\cdots, 0, 1, 0, \cdots), \cdots, (\cdots, 0, 1)\}$이 표준스러운 순서기저임.

2. 좌표벡터(coordinate vector)

1) 정의

Definition 12. *유한차원 벡터공간 V의 순서기저를 $\beta\{u_1, u_2, \cdots, u_n\}$이라 하고, $x \in V$에 대해서 $a_1, a_2, \cdots, a_n$은 $x = \sum_{i=1}^{n} a_i u_i$를 만족시키는 유일한 스칼라라 하자. β에 대한 좌표벡터(coordinate vector) $[x]_\beta$는 아래와 같음.*

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

순서기저를 사용해 벡터를 n순서쌍으로 나타낸 것.

벡터를 좌표벡터로 표현하면 여러 계산에서 일종의 행렬로서 다룰 수 있게 됨. 벡터를 행렬의 세계로 가져온 것. 표준표현 관련 내용에 삽입되어 있는 그림을 보면 쉽게 이해할 수 있음.

3. 선형변환의 행렬표현(matrix representation)

선형변환을 행렬표현으로, 행렬표현을 선형변환으로 생각할 수 있음.

1) 정의[16]

> **Definition 13.** *유한차원 벡터공간 V, W와 각각의 순서기저 $\beta = \{v_1, \cdots, v_n\}$, $\gamma = \{w_1, \cdots, w_m\}$, 선형변환 $T : V \to W$가 있음. $j = 1, 2, \cdots, n$일 때, j마다 다음을 만족하는 유일한 스칼라 $a_{ij} \in F$가 존재함.*
>
> $$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$
>
> *성분이 $A_{ij} = a_{ij}$인 $m \times n$ 행렬 A를 순서기저 β, γ에 대한 선형변환 T의 행렬표현(matrix representation)이라 하고, $A = [T]_\beta^\gamma$라 표기함. $V = W$, $\beta = \gamma$이면 간단히 $A = [T]_\beta$라 표기함.*

선형변환 $T : V \to W$에 대해서 V의 기저의 함숫값을 W의 기저로 나타낸 것.

j번째 열은 γ에 대한 $T(v_j)$의 좌표벡터라고 볼 수 있음.

linear extension theorem의 Corollary에 의해, 선형변환 $U : V \to W$가 $[U]_\beta^\gamma = [T]_\beta^\gamma$를 만족하면 $U = T$임.

2) 영변환의 행렬표현

$T_0(v_j) = 0 = 0w_1 + 0w_2 + \cdots + 0w_m$이므로 $[T_0]_\beta^\gamma = O(m \times n$영행렬$)$임.

즉, 영변환의 행렬표현은 영행렬임.

3) 항등변환의 행렬표현[17]

$I_{v_j} = v_j = 0v_1 + 0v_2 + \cdots + 1v_j + \cdots + 0v_n$이므로 I_V의 j열은 e_j이고, $[I_V]_\beta^\beta = [I_V]_\beta$는 아래와 같음.

$$[I_V]_\beta^\beta = [I_V]_\beta = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

즉, 항등변환의 행렬표현은 $n \times n$ 항등행렬임.

[16] linear extension theorem이 이 정의가 가진 논리의 바탕이 됨.
[17] 항등변환은 그 정의상 정의역과 공역이 동일한 벡터공간임.

4. 선형변환의 집합 $\mathcal{L}(V, W)$[18]

1) 정의

Definition 14. *F-벡터공간 V, W에 대하여 V에서 W로 가는 모든 선형변환의 모임으로 이루어진 벡터공간을 $\mathcal{L}(V, W)$라 표기함. $V = W$이면 $\mathcal{L}(V, V)$를 간단히 $\mathcal{L}(V)$라 표기함.*

Theorem 2.7에 의하면 $\mathcal{L}(V, W)$는 F-벡터공간임.

2) $[\cdot]_\beta^\gamma$는 선형변환인가?

선형변환을 행렬표현으로 나타낼 수 있으므로, $T \in \mathcal{L}(V, W)$에 대해서 $U(T) = [T]_\beta^\gamma$인 $U : \mathcal{L}(V, W) \to M_{dim(W) \times dim(V)}(F)$라는 함수를 생각할 수 있음.

Theorem 2.8에 의하면 $U(T) = [T]_\beta^\gamma$는 선형변환임.

3) $\mathcal{L}(V, W)$의 기저

$\mathcal{L}(V, W)$의 기저는 아래와 같음. (증명은 여종헌 프린트 참고.)

$$T_{ij}(v_k) = \begin{cases} w_i & (k = j) \\ 0 & (k \neq j) \end{cases}$$

5. 관련 정리

1) $\mathcal{L}(V, W)$은 벡터공간인가?

Theorem 2.7 F-벡터공간 V, W와 선형변환 $T, U : V \to W$에 대하여 아래가 성립함.

먼저, 선형변환의 합과 스칼라 곱을 정의하자.
선형변환의 합과 스칼라 곱은 보편적인 함수의 집합($\mathcal{F}$)의 것과 같음.

1. 임의의 $a \in F$에 대하여, $aT + U$는 선형임.
2. 위 정의와 같이 선형변환의 합과 스칼라 곱을 정의할 때, V에서 W로 가는 모든 선형변환의 집합은 F-벡터공간임.

선형변환의 집합($\mathcal{L}$)은 함수의 집합($\mathcal{F}$)의 부분공간임.

Proof. 미리 정의한 함수의 합과 곱을 사용해 $(aT + U)(cx + y) = c(aT + U)(x) + (aT + U)(y)$임을 보이면 됨.

합과 곱에 대해서 닫혀 있고 T_0이 $\mathcal{L}(V, W)$에서 영벡터이므로 $\mathcal{L}(V, W)$는 벡터공간임. □

2) $U(T) = [T]_\beta^\gamma$는 선형변환인가?

Theorem 2.8 유한집합 벡터공간 V, W와 각각의 순서기저 β, γ, 선형변환 $T, U : V \to W$에 대해서 아래가 성립함.

1. $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$
2. 모든 스칼라 a에 대해서 $[aT]_\beta^\gamma = a[T]_\beta^\gamma$

즉, $[\cdot]_\beta^\gamma$, 다시 말해 $U(T) = [T]_\beta^\gamma$는 선형변환임.

Proof. $\beta = \{v_1, v_2, \cdots, v_n\}$, $\gamma = \{w_1, w_2, \cdots, w_n\}$라 하자. $a_1, \cdots, a_n, b_1, \cdots, b_n \in F$, $j = 1, 2, \cdots, n$에 대해서, $T(v_j) = \sum_{i=1}^m a_{ij} w_i$, $U(v_j) = \sum_{i=1}^m b_{ij} w_i$이고 $([T]_\beta^\gamma)_{ij} = a_{ij}$, $([U]_\beta^\gamma)_{ij} = b_{ij}$임. 이를 이용해 $(T + U)(v_j)$, $(aT)(v_j)$를 정리할 수 있음. □

[18]새로운 집합과 함수가 등장했으므로, 지금까지 배운 내용을 생각해 보면 $\mathcal{L}(V, W)$가 벡터공간인지, $U(T) = [T]_\beta^\gamma$가 선형변환인지는 당연히 생각해봐야 하는 주제임.

3. 선형변환의 합성과 행렬 곱

1. 선형변환의 합성

1) 새로운 연산의 필요성

Theorem 2.9에 의하면 선형변환의 합성 또한 선형변환임.

지금까지 설명한 선형변환의 행렬표현은 선형변환의 합성에 대한 연산이 존재하지 않으므로 이를 정의해줘야 함.

2. 행렬 곱

1) 정의

Definition 15. *$m \times n$ 행렬 A와 $n \times p$ 행렬 B에 대하여 두 행렬 A, B의 곱(product) AB는 아래와 같이 정의된 $m \times p$ 행렬임.*

$$1 \leq i \leq m,\, 1 \leq j \leq p,\, (AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_k j$$

이 행렬 곱은 두 선형변환의 합성을 행렬로서 나타낸 연산임.

2) 유도과정

유한차원 벡터공간 V, W, Z와 선형변환 $T : V \to W$, $U : W \to Z$가 있음. V의 순서기저 $\{v_1, \cdots, v_n\}$, W의 순서기저 $\{w_1, \cdots, w_m\}$, Z의 순서기저 $\{z_1, \cdots, z_p\}$에 대하여 $A = [U]_\beta^\gamma$, $B = [T]_\alpha^\beta$라 하자.

$AB = [UT]_\alpha^\gamma$가 되도록 하는 행렬 곱을 정의할 것임. 선형변환의 행렬표현은 $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$꼴에 의해 결정되므로, $(UT)(v_j)$를 $\sum_{i=1}^{p} C_{ij} z_i$꼴로 나타내 C_{ij}에 대해 살펴볼 것임.

$$\begin{aligned}(UT)(v_j) &= U(T(v_j)) = U(\sum_{k=1}^{m} B_{kj} w_k) = \sum_{k=1}^{m} B_{kj} U(w_k) = \sum_{k=1}^{m} B_{kj} (\sum_{i=1}^{p} A_{ik} z_i) = \sum_{i=1}^{p} (\sum_{k=1}^{m} A_{ik} B_{kj}) z_i \\ &= \sum_{i=1}^{p} C_{ij} z_i \, (C_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj})\end{aligned}$$

즉, 성분이 $c_{ij} = C_{ij}$인 $q \times n$ 행렬 c는 선형변환 UT의 행렬표현 $[UT]_\alpha^\gamma$임.

여기서의 $C_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj}$를 행렬 곱으로 정의함.

3) 행렬 곱과 행렬의 크기[19]

행렬 곱을 하는 두 행렬의 내부 차원이 같아야 행렬 곱이 정의됨.

행렬 곱을 하는 두 행렬의 외부 차원은 결과로 만들어지는 행렬의 크기를 결정함.

4) 연속적인 행렬 곱의 표현

$n \times n$ 행렬 A에 대해서, $A^1 = A, A^2 = AA, A^3 = A^2 A, \cdots$로 정의함. 즉, 2 이상의 자연수 k에 대해서, $A^k = A^{k-1} A$라 정의함. 이때, $A^0 = I_n$임.

[19] 이는 기저의 개수를 생각하면 당연함.

3. 선형변환의 합성과 행렬 곱 사이의 관계

1) 대응관계

행렬 곱의 유도과정과 Theorem 2.9에서 확인할 수 있듯이, 선형변환의 합성과 행렬 곱은 서로 완전히 대응됨.

2) 동일한 성질

Theorem 2.10, 2.12, 2.16 등에서 확인할 수 있듯이, 선형변환의 합성과 행렬 곱은 아래의 성질들을 공유함.

1. 분배법칙 성립.
2. 결합법칙 성립.
3. 항등원 존재. (항등변환, 항등행렬)
4. 스칼라 곱의 분배법칙 성립.
5. 교환법칙 성립하지 않음.[20]
6. 곱셈의 소거법칙 성립하지 않음.

4. 좌측 곱 변환

행렬로 선형변환의 성질을, 또는 선형변환으로 행렬의 성질을 유추할 때 가장 유용하게 사용할 수 있는 도구.

1) 정의

Definition 16. *A는 $m \times n$ 행렬이고, 성분은 체 F의 원소임. 아래의 선형변환을 간단히 L_A라 표기하자.*

$$L_A : F^n \to F^m,\ L_A(x) = Ax$$

L_A를 좌측 곱 변환이라 함. 이때 x는 F^n의 열벡터이고, Ax는 A와 x의 행렬 곱임.

2) L_A와 A의 관계

L_A에 대해서, 수많은 가짓수의 행렬표현들 중에서 표준 순서기저로 만든 행렬표현이 A임.

3) 성질

Theorem 2.15에서 알 수 있듯이, 좌측 곱 변환은 선형이고 해당 Theorem에 서술한 여러 성질들을 가짐.

이 성질들을 이용해 행렬의 결합법칙 등 행렬의 성질들을 증명할 수 있음.

[20]즉, $TU \neq UT$, $AB \neq BA$임

4. 관련 정리

1) 선형변환의 합성은 선형변환인가?

Theorem 2.9 F-벡터공간 V, W, Z와 선형변환 $T : V \to W$, $U : W \to Z$를 생각하자. 두 선형변환의 합성 $UT : V \to Z$는 선형변환임.

2) 선형변환의 합성이 가지는 성질

Theorem 2.10 벡터공간 V와 선형변환 $T, U_1, U_2 \in \mathcal{L}(V)$에 대해서 아래가 성립함.

1. $T(U_1 + U_2) = TU_1 + TU_2$이고, $(U_1 + U_2)T = U_1T + U_2T$ (분배법칙)
2. $T(U_1U_2) = (TU_1)U_2$ (결합법칙[21])
3. $TI = IT = T$
4. 모든 스칼라 a에 대해서 $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$

3) 선형변환의 합성과 행렬 곱

Theorem 2.11 유한차원 벡터공간 V, W, Z와 각각의 순서기저 α, β, γ, 선형변환 $T : V \to W$, $U : W \to Z$에 대해서 아래가 성립함.

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$$

행렬 곱의 유도과정과 정의를 생각하면 당연함.

4) 행렬 곱의 성질

Theorem 2.12 A가 $m \times n$ 행렬, B와 C가 $n \times p$ 행렬, D와 E가 $q \times m$ 행렬일 때, 다음이 성립함.

1. $A(B + C) = AB + AC$, $(D + E)A = DA + EA$ (분배법칙)
2. 임의의 스칼라 a에 대하여 $a(AB) = (aA)B = A(aB)$
3. $I_mA = A = AI_n$

이 성질은 각각 Theorem 2.10의 1, 4, 3과 대응됨.

5) 특정 열벡터 구하는 방법

Theorem 2.13 $m \times n$ 행렬 A와 $n \times p$ 행렬 B가 있음. $j = 1, 2, \cdots, p$인 j에 대해서 AB의 j열을 u_j, B의 j열을 v_j라 표기하면 아래가 성립함.

1. $u_j = Av_j$
2. $v_j = Be_j$ (이때, e_j는 F^p의 j번째 표준 벡터[22])

즉, AB의 특정 열벡터를 구하려면 A에다가 B의 (구하려는 순번의) 열벡터를 곱하면 됨.

B의 특정 열벡터를 구하려면 B 뒤에 e_j를 곱하면 됨.

[21] 함수의 합성에서 결합법칙이 성립하는 것은 책의 부록에서 확인할 수 있음.
[22] j번째만 1이고 나머지는 0인 tuple.

6) 좌표벡터의 행렬 곱

Theorem 2.14 V, W는 유한차원 벡터공간이고, 순서기저는 각각 β, γ임. 선형변환 $T : V \to W$와 $u \in V$에 대해서 아래가 성립함.

$$[T(u)]_\gamma = [T]_\beta^\gamma [u]_\beta$$

즉, 선형변환에 벡터를 대입하고 행렬로 표현하나, 선형변환의 행렬표현에 해당 벡터의 좌표벡터를 곱하나 동일함.

좌표벡터의 행렬 곱은 해당 선형변환에 대입하는 것과 같음.

Proof. $u \in V$를 고정하고 아래와 같은 선형변환 $f : F \to V, g : F \to W$를 정의하자.

$$\text{모든 } a \in F\text{에 대해서 } f(a) = au,\ g(a) = aT(u)$$

$\alpha = \{1\}$을 F의 표준 순서기저라 하자. $g = Tf$이므로 아래가 성립함.

$$[T(u)]_\gamma = [g(1)]_\gamma = [g]_\alpha^\gamma = [Tf]_\alpha^\gamma = [T]_\beta^\gamma [f]_\alpha^\beta = [T]_\beta^\gamma [f(1)]_\beta = [T]_\beta^\gamma [u]_\beta$$

□

7) 좌측 곱 변환의 성질

Theorem 2.15 A는 $m \times n$ 행렬이고, 성분은 체 F의 원소임. 좌측 곱 변환 $L_A : F^n \to F^m$은 선형임.

또한 임의의 $m \times n$ 행렬 B(성분은 체 F의 원소)와 F^n의 표준 순서기저 β, F^m의 표준 순서기저 γ에 대해서 아래가 성립함.

1. $[L_A]_\beta^\gamma = A$ (표준 순서기저에 대해서만 성립함.)
2. $L_A = L_B \iff A = B$
3. $L_{A+B} = L_A + L_B$
4. 모든 $a \in F$에 대하여 $L_{aA} = aL_A$
5. $T : F^n \to F^m$가 선형이면 $T = L_C$가 되도록 하는 $m \times n$ 행렬 C가 유일하게 존재함. $C = [T]_\beta^\gamma$임.
6. $L_{AE} = L_A L_E$ (E는 $n \times p$ 행렬)
7. $m = n$이면 $L_{I_n} = I_{F^n}$임.
8. $a_i (1 \le i \le n)$가 A의 i번째 열일 때, $R(L_A) = span(\{a_1, \cdots, a_n\})$임.

Proof. (1) $[L_A]_\beta^\gamma$의 j열은 $L_A(e_j)$와 같음. $L_A(e_j) = Ae_j$는 A의 j열이므로 $[L_A]_\beta^\gamma = A$임.

(2) $A = [L_A]_\beta^\gamma = [L_B]_\beta^\gamma = B$

(5) 모든 $x \in F^n$에 대해서 $T(x) = Cx = L_C(x)$이므로 $T = L_C$임. 2번 성질에 의해 C는 유일함. □

8) 행렬 곱의 결합법칙

Theorem 2.16 행렬 곱에서 결합법칙이 성립함. 즉, $A(BC)$를 정의할 수 있는 행렬 A, B, C는 $(AB)C$도 정의할 수 있고 $A(BC) = (AB)C$임.

이는 Theorem 2.10의 2와 대응됨. (선형변환의 합성의 결합법칙 성립과 대응됨.)

Proof. $L_{A(BC)} = L_A L_{BC} = L_A(L_B L_C) = (L_A L_B)L_C = L_{AB}L_C = L_{(AB)C}$

Theorem 2.15의 2번 성질에 의해 $A(BC) = (AB)C$임. □

4. 가역성과 동형사상

1. 함수의 역함수(inverse)와 가역(invertible)

1) 정의

> **Definition 17.** *벡터공간 V, W와 선형변환 $T : V \to W$를 생각하자. $TU = I_W$이고 $UT = I_V$인 함수 U를 T의 역함수(inverse)라 함. 이 역함수는 T^{-1}로 표기함. T가 가역이면 T^{-1}은 유일함.*
>
> *역함수가 존재하는지를 가역(invertible)이라 함.*

2) 사실

$(TU)^{-1} = U^{-1}T^{-1}$ (사실1)

$(T^{-1})^{-1} = T$임. 특히, T^{-1}은 가역임. (사실2)

V, W가 차원이 같은 벡터공간일 때, 선형변환 $T : V \to W$가 가역일 필요충분조건은 $rank(T) = dim(V)$인 것임. (사실3)

즉, 함수가 가역이기 위한 필요충분조건은 해당 함수가 단사(injection)이고 전사(subjection)인 것임.

3) 역함수(inverse)의 선형성

Theorem 2.17에 의하면 선형변환의 역함수 또한 선형변환임.

2. 행렬의 역행렬(inverse)과 가역(invertible)

1) 정의

> **Definition 18.** *$n \times n$ 행렬 A에 대해서 $AB = BA = I$인 $n \times n$ 행렬 B를 A의 역행렬(inverse)이라고 함. 이 역행렬은 A^{-1}로 표기함.*
>
> *역행렬이 존재하는지를 가역(invertible)이라고 함.*

정의에서 알 수 있듯이, 정사각행렬(정방행렬)이 아니면 가역일 수 없음.

행렬의 역행렬을 구하는 방법은 Part3에서 다룸.

2) 성질

A가 가역이면 A의 역행렬은 유일함.

3. 가역성의 판별

1) 선형변환의 가역성 판별

선형변환이 전단사인지를 확인함(Dimension Theorem, Theorem 2.4, Theorem 2.5). 또는 그 행렬표현이 가역인지를 확인함(Theorem 2.18).

2) 행렬의 가역성 판별

$n \times n$ 행렬인 경우 랭크가 n인지 확인하거나, 행렬식이 0이 아닌지 확인함. 또는 해당하는 선형변환이 가역인지를 확인함(Theorem 2.18).

4. 동형(isomorphic)과 동형사상(isomorphism)

구체적으로는 다르지만 구조적으로는 같은 것을 동형(isomorphic)이라 함.

동형사상은 벡터공간의 모든 구조를 보존하는 함수임.

1) 정의

> **Definition 19.** *두 벡터공간 V, W 사이에 가역인 선형변환 $T : V \to W$가 존재하면 V와 W는 동형(isomorphic)임. 이때 가역인 선형변환을 V에서 W로 가는 동형사상(isomorphism)이라 함.*

무엇이 무엇과 동형이라는 개념은 동치관계임.[23]

2) 동형이기 위한 필요충분조건

Theorem 2.19에 의하면, 같은 체 위의 두 벡터공간의 차원이 같다면 두 벡터공간은 동형임.

즉, 차원이 같음 = 동형임 = 가역인 선형변환이 존재함.

3) 동형사상의 역할

동형사상은 구조적으로 동일한 두 벡터공간을 이어준다는 점에서 벡터공간의 '이름만 바꾸는' 역할을 한다고 이해할 수 있음.

일례로, 선형변환의 행렬표현이 가능했던 것 또한 사실 선형변환의 집합과 행렬의 집합이 동형이고, $[\,]_\beta^\gamma$가 '선형변환의 이름을 행렬로 바꾸는' 동형사상이기 때문임. 이는 Theorem 2.20에서 Φ_β^γ로 정의하고 있음.

아래의 그림은 벡터공간 V, W에 대해 동형사상(isomorphism) $T : V \to W$이 어떤 식으로 '이름을 바꾸는지'를 나타낸 다이어그램임. V의 연산을 W의 연산으로 바꾸고 있음.

$$\begin{tikzcd}
\mathrm{V} \arrow[r, "+\times"] \arrow[d, "\mathrm{T}"'] & \mathrm{V} \arrow[d, "\mathrm{T}"] \\
\mathrm{W} \arrow[r, "+'\times'"] & \mathrm{W}
\end{tikzcd}$$

5. 표준표현(standard representation)

1) 정의

> **Definition 20.** *체 F에서의 n차원 벡터공간 V의 순서기저를 β라 하자. β에 대한 V의 표준표현(standard representation)은 아래와 같이 정의된 함수 $\phi_\beta : V \to F^n$임.*
>
> $$x \in V,\ \phi_\beta(x) = [x]_\beta$$

즉, 벡터를 좌표벡터로 나타내는 함수.

2) 선형성과 동형성

ϕ_β는 선형변환임.

Theorem 2.21에 의하면 ϕ_β는 벡터공간 V와 그 순서기저 β에 대해서 동형사상임.
이는 곧 n차원 벡터공간과 F^n이 동형이라는 것을 의미함.

[23] 즉, V는 V와 동형이고, V가 W와 동형이면 W가 V와 동형이고, V가 W와 동형이고 W가 S와 동형이면 V가 S와 동형임. 간단히 정리하면 'V와 W는 그저 동형'인 것.

3) 선형변환 합성의 가환적 그래프

차원이 각각 n, m인 벡터공간 V, W와 각각의 순서기저 β, γ, 선형변환 $T : V \to W$가 있다고 하자. ϕ_β를 이용하면 V와 F^n을 동일시할 수 있고, ϕ_γ를 이용하면 W와 F^m을 동일시할 수 있음. 즉, T와 $L_{[T]_\beta^\gamma}$또한 동일시할 수 있음. 이를 아래와 같이 수식으로 나타낼 수 있음.

$$L_{[T]_\beta^\gamma}\phi_\beta = \phi_\gamma T$$

이것을 이용하면 임의의 두 벡터공간 사이에 정의된 연산을 F^n과 F^m 사이에 정의된 연산으로 나타낼 수 있음.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, shift right, "\phi_\beta"'] & W \arrow[d, shift right, "\phi_\gamma"'] \\
\mathbb{F}^n \arrow[r, "L_{[T]_\beta^\gamma}"'] \arrow[u, shift right, "\phi_\beta^{-1}"'] & \mathbb{F}^m \arrow[u, shift right, "\phi_\gamma^{-1}"']
\end{tikzcd}$$

6. 관련 정리

1) 역함수의 선형성

Theorem 2.17 벡터공간 V, W와 가역인 선형변환 $T : V \to W$에 대해서 역함수 $T^{-1} : V \to W$ 또한 선형임.

Corollary 선형변환 $T : V \to W$가 가역이라고 하자. V가 유한차원이기 위한 필요충분조건은 W가 유한차원인 것임. 특히 이때 $dim(V) = dim(W)$임.

즉, 가역인 선형변환이 존재하면 두 벡터공간 모두가 유한차원이거나 무한차원이고, 유한차원인 경우 두 벡터공간의 차원도 같음.

2) 선형변환의 역함수와 행렬의 역행렬의 관계

Theorem 2.18 유한차원 벡터공간 V, W와 각각의 순서기저 β, γ, 선형변환 $T : V \to W$에 대해서 T가 가역이기 위한 필요충분조건은 $[T]_\beta^\gamma$가 가역인 것임. 특히, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$임.

이를 다이어그램으로 생각해 보면 이해가 편함.

Corollary $n \times n$ 행렬 A가 가역이기 위한 필요충분조건은 L_A가 가역인 것임. 특히, $(L_A)^{-1} = L_{A^{-1}}$임.

3) 동형이기 위한 필요충분조건

Theorem 2.19 같은 체 위에서 정의된 유한차원 벡터공간 V, W에 대해서 V와 W가 동형이기 위한 필요충분조건은 $dim(V) = dim(W)$임.

Proof. 1. V, W가 동형임. $\to dim(V) = dim(W)$
V, W 사이에는 가역인 선형변환이 존재하는데, 두 벡터공간의 차원이 다르다면 모순임.

2. $dim(V) = dim(W) \to V, W$가 동형임.
두 벡터공간의 차원이 같으므로, 각 기저를 일대일대응시키는 선형변환을 생각할 수 있음. 이 선형변환은 동형사상임. □

4) Φ_β^γ의 동형성

Theorem 2.20 차원이 각각 n, m인 F-벡터공간 V, W를 생각하자. V, W의 순서기저를 각각 β, γ라 할 때, 아래와 같이 정의한 함수 $\Phi_\beta^\gamma : \mathcal{L}(V, W) \to M_{m\times n}(F)$는 동형사상임.

$$T \in \mathcal{L}(V, W),\ \Phi_\beta^\gamma(T) = [T]_\beta^\gamma$$

즉, 선형변환의 집합과 행렬의 집합은 동형이고, 선형변환의 행렬표현을 나타내는 $[]_\beta^\gamma$, 즉 Φ_β^γ는 동형사상임. 이를 Fundamental Theorem of Linear Algebra라고 함. (선형사상=행렬)

Proof. 동형사상은 가역인 선형변환임. Theorem 2.8에 의하면 Φ_β^γ는 선형이므로 전단사인 것을 보여야 함. 임의의 $m \times n$ 행렬 A에 대해서 $\Phi_\beta^\gamma(T) = A$인 선형변환 $T : V \to W$가 유일하게 존재함을 보이면 됨.

$\beta = \{v_1, \cdots, c_n\}$, $\gamma = \{w_1, \cdots, w_m\}$이라 하면, linear extension theorem에 의해 아래와 같이 정의된 선형변환 $T : V \to W$가 유일함.

$$T(v_j) = \sum_{i=1}^{m} A_{ij} w_i \, (1 \le j \le n)$$

즉, 행렬 A에 대해 선형변환 T가 유일함.

+ 위 증명법은 좀 미흡함. 여종헌 프린트에 $\mathcal{L}(V, W)$의 차원을 직접 구해 증명하는 방법도 나와있음. □

Corollary 차원이 각각 n, m인 유한차원 벡터공간 V, W에 대해서 $\mathcal{L}(V, W)$는 차원이 nm인 벡터공간임.

5) ϕ_β의 동형성

Theorem 2.21 임의의 유한차원 벡터공간 V와 순서기저 β에 대해서 ϕ_β는 동형사상임.

5. 좌표변환 행렬과 닮음

1. 좌표변환 행렬(change of coordinate matrix)

1) 정의

Definition 21. *Theorem 2.22에 의하면, 유한차원 벡터공간 V의 두 순서기저 β, β'에 대해서, $Q = [I_V]_{\beta'}^{\beta}$라 하자. 아래가 성립함.*

1. *Q는 가역행렬임.*
2. *임의의 $v \in V$에 대해서 $[v]_\beta = Q[v]_{\beta'}$*

Proof. 1. $Q = [I_V]_{\beta'}^{\beta}$인데, I_V가 가역이므로, Q 또한 가역임.

2. $[v]_\beta = [I_V(v)]_\beta = [I_V]_{\beta'}^{\beta}[v]_{\beta'} = Q[v]_{\beta'}$ □

즉, 좌표변환 행렬은 벡터공간의 어떤 벡터에 대해서, 특정 기저로 표현한 좌표벡터를 다른 기저로 표현한 좌표벡터로 만들 수 있는 것.

Q는 β'좌표를 β좌표로 변환함.

이때, Q^{-1}은 β좌표를 β'좌표로 변환함.

2. 선형변환 행렬표현의 전환

1) 선형변환 행렬표현의 전환

Theorem 2.23에 의하면, 한 선형변환의 행렬표현을 다른 기저를 이용한 동일한 선형변환의 행렬표현으로 전환할 수 있음. 아래가 그 식임. 이때 Q는 β'좌표를 β좌표로 옮기는 좌표변환 행렬임.

$$[T]_{\beta'} = [I_V]_{\beta}^{\beta'}[T]_{\beta}^{\beta}[I_V]_{\beta'}^{\beta} = Q^{-1}[T]_\beta Q$$

$$[T]_{\beta} = [I_V]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'}[I_V]_{\beta}^{\beta'} = Q[T]_{\beta'}Q^{-1}$$

$$[T]_{\alpha}^{\beta} = [I_W]_{\beta'}^{\beta}[T]_{\alpha'}^{\beta'}[I_V]_{\alpha}^{\alpha'}$$

3. 닮음(similar)[24]

1) 정의

Definition 22. *A, B가 $M_{n\times n}(F)$의 행렬이라 하자. $B = Q^{-1}AQ$인 가역행렬 Q가 존재하면 B는 A와 서로 닮음(similar)임.*

즉, 동일한 선형변환에 대해 서로 다른 두 행렬표현은 서로 닮음(similar)임.

닮음 관계는 동치관계임.

서로 닮은 행렬들에 대해서 모두 같은 값을 갖는, invariant인 것들이 있음. invariant에는 랭크(rank), 행렬식(determinant), 대각합(trace) 등이 있음.[25]

[24]닮음에 대한 자세한 이야기는 5 7장에서 다룸.
[25]invariant에 속하는 것들은 뒷 장에서 계속해서 등장함.

4. 관련 정리

1) 선형변환 행렬표현의 전환

Theorem 2.23 유한차원 벡터공간 V의 선형연산자 T와 V의 순서기저 β, β'이 있음. Q가 β'좌표를 β좌표로 옮기는 좌표변환 행렬이라 하면 아래가 성립함.

$$[T]_{\beta'} = [I_V]_{\beta}^{\beta'}[T]_{\beta}^{\beta}[I_V]_{\beta'}^{\beta} = Q^{-1}[T]_{\beta}Q$$

즉, T가 유한차원 벡터공간 V의 선형연산자이고 β와 β'이 V의 순서기저일 때, $[T]_{\beta'}$과 $[T]_{\beta}$는 서로 닮음 (similar)임.

이를 더 일반화하여 V, W 사이에 정의된 선형변환 $T : V \to W$에 대해 나타내면 아래와 같음.

$$[T]_{\alpha}^{\beta} = [I_W]_{\beta'}^{\beta}[T]_{\alpha'}^{\beta'}[I_V]_{\alpha}^{\alpha'}$$

Proof. 방법1. 선형변환 행렬표현의 전환을 함수의 합성으로 나타내기.

벡터공간 V의 기저 α, α'와 W의 기저 β, β'에 대해서, 벡터의 이동을 고려하면 아래의 다이어그램을 생각할 수 있음. 즉, $[T]_{\alpha}^{\beta} = [I_W]_{\beta'}^{\beta}[T]_{\alpha'}^{\beta'}[I_V]_{\alpha}^{\alpha'}$임. W 대신 V를 넣으면 위 식을 유도할 수 있음.

$$\begin{tikzcd}
\text{V } \alpha \arrow[r, "{[T]_{\alpha}^{\beta}}"] \arrow[d, "{[I_V]_{\alpha}^{\alpha'}}"'] & \text{W } \beta \\
\text{V } \alpha' \arrow[r, "{[T]_{\alpha'}^{\beta'}}"'] & \text{W } \beta' \arrow[u, "{[I_W]_{\beta'}^{\beta}}"']
\end{tikzcd}$$

방법2. 단순 유도하기.

$Q[T]_{\beta'} = [I]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'} = [IT]_{\beta'}^{\beta} = [TI]_{\beta'}^{\beta} = [T]_{\beta}^{\beta}[I]_{\beta'}^{\beta} = [T]_{\beta}Q$ □

Corollary $A \in M_{n\times n}(F)$와 F^n의 순서기저 γ에 대해서 아래가 성립함.

$$[L_A]_{\gamma} = Q^{-1}AQ$$

이때, $n \times n$ 행렬 Q의 j열은 γ의 j번째 벡터임.

6. 쌍대공간

1. 선형범함수(linear functional)

1) 정의

Definition 23. *벡터공간 V에서 F로 가는 선형변환을 선형범함수(linear functional)라고 함.*

본 필기에서 선형범함수는 $f, g, h, \cdots$ 등으로 표기함.

2) 좌표함수

Definition 24. *순서기저 $\beta = \{x_1, \cdots, x_n\}$을 가지는 유한차원 벡터공간 V에 대해서 각 $i = 1, 2, \cdots, n$ 마다 아래와 같은 함수를 정의함.*

$$\beta\text{에 대한 } x\text{의 좌표벡터가 } [x]_\beta = (a_1\, a_2\, \cdots\, a_n)^t \text{일 때, } f_i(x) = a_i$$

이때, f_i는 V의 선형범함수이고 기저 β에 대한 i번째 좌표함수(coordinate function)라 함.

또한, 좌표함수 f_i에 대해서 $f_i(x_j) = \delta_{ij}$가 성립함.

즉, 좌표함수는 x를 순서기저의 일차결합으로 표현했을 때, 해당 기저에 대해서 i번째 계수를 내보냄.

자연스럽게 좌표함수는 어떤 고정된 기저에 대한 것임.

2. 쌍대공간(dual space)

1) 정의

Definition 25. *F-벡터공간 V에 대해서 벡터공간 $\mathcal{L}(V, F)$를 V의 쌍대공간(dual space)이라 하며, 간단히 V^*라 표기함.*

즉, V^*는 V에서 F로 가는 선형범함수로 이뤄진 벡터공간임.

2) 쌍대공간의 차원

Theorem 2.20에 의해, $dim(V^*) = dim(\mathcal{L}(V, F)) = dim(V)dim(F) = dim(V)$이므로 V와 V^*는 차원이 같음.

즉, V와 V^*은 동형임.

3. 쌍대기저(dual basis)

1) 정의

Definition 26. *Theorem 2.24의 표기법을 그대로 따르자. $f_i(x_j) = \delta_{ij}$ $(1 \le i, j \le n)$를 만족하는 V^*의 순서기저 $\beta^* = \{f_1, \cdots, f_n\}$을 β의 쌍대기저(dual basis)라고 함.*

즉, 쌍대기저는 어떤 고정된 기저에 대한 좌표함수들의 집합임. 기저로부터 쌍대기저가 나온다고 이해할 수 있음.

Theorem 2.24에 의하면 쌍대공간의 임의의 벡터 $f \in V^*$는 $f = \sum_{i=1}^{n} f(x_i) f_i$로 표현됨.

2) 구체적인 쌍대기저 구하기

V의 표준 순서기저를 $\gamma = \{e_1, \cdots, e_n\}$이라 하자.

1. $f_i(x_j) = \delta_{ij}$이므로 각 벡터를 대입해 $f_i(e_i)$에 대한 연립일차방정식을 풂.
2. 임의의 벡터 $x \in V$는 e_i에 대해 간단히 표현되므로, $f_i(x) = \sum_{j=1}^{n} a_j f_j(e_j)$꼴로 f_i를 구할 수 있음.

또는 그냥 단순히 β로 임의의 벡터를 표현한 후, f_i에 대입, 정리하여 구할 수도 있음. 편한 방법으로 하자.

증명 등에서 쌍대기저는 구체적으로 구하기보단 주로 단순하게 $f_1, f_2, \cdots$ 등으로만 표기하여 사용함.

4. 전치(transpose)

1) 정의

Definition 27. *Theorem 2.25에서 정의한 선형변환 T^t를 T의 전치(transpose)라 함.*

5. 이중 쌍대공간

1) 정의

Definition 28. *V^*의 쌍대공간을 V의 이중 쌍대공간(double dual) V^{**}이라 함.*

여기서는 V와 V^{**}의 동일화(identification)에 대해 다룸. 두 벡터공간의 기저를 어떻게 선택하더라도 영향을 받지 않는 동형사상이 존재하면 동일한 것.

2) $\hat{x}$

Definition 29. *벡터 $x \in V$, $f \in V^*$에 대해서 함수 $\hat{x} : V^* \to F$를 $\hat{x}(f) = f(x)$라 정의함.*

$\hat{x}$는 선형범함수이고, $\hat{x} \in V^{**}$임.

$x \in V$와 $\hat{x} \in V^{**}$가 서로 대응된다면 두 벡터공간의 기저를 어떻게 선택하더라도 영향을 받지 않는 동형사상을 정의할 수 있게 됨.

3) V와 V^{**}의 관계

Theorem 2.26에 의하면 V와 V^{**}는 기저의 고정과 관계없이 ψ라는 동형사상이 존재함. 이런 동형성을 Natural Isomorphism이라고 하고, V와 V^{**}는 동일화(identification)된다고 함.[26]

여담으로, 이 성질은 V가 유한차원일 때만 성립함.

[26] 이 부분은 내용이 명확히 정리되지 않아 보충이 필요함.

6. 관련 정리

1) 쌍대기저

Theorem 2.24 순서기저 $\beta = \{x_1, \cdots, x_n\}$을 가지는 유한차원 벡터공간 V와 β에 대한 i번째 좌표함수 f_i $(1 \leq i \leq n)$를 생각하자. 이때 $\beta^* = \{f_1, f_2, \cdots, f_n\}$은 V^*의 순서기저임. 또한 임의의 $f \in V^*$에 대해서 아래가 성립함.

$$f = \sum_{i=1}^{n} f(x_i) f_i$$

Proof. $dim(V^*) = n$이므로 임의의 $f \in V*$에 대해서 $f = \sum_{i=1}^{n} f(x_i) f_i$가 성립함을 보이면 됨. $g = \sum_{i=1}^{n} f(x_i) f_i$이라 하면 $1 \leq j \leq n$에 대해서 아래가 성립함.

$$g(x_j) = \sum_{i=1}^{n} f(x_i) f_i(x_j) = \sum_{i=1}^{n} f(x_i) \delta_{ij} = f(x_j)$$

$g = f$이므로 임의의 $f \in V^*$에 대해서 $f = \sum_{i=1}^{n} f(x_i) f_i$가 성립한다는 것을 알 수 있음. □

2) 선형변환의 전치

Theorem 2.25 V와 W는 F에서의 유한차원 벡터공간이고, 순서기저는 각각 β와 γ임. 임의의 선형변환 $T : V \to W$에 대해서, 함수 $T^t : W^* \to V^*$를 아래와 같이 정의함.

모든 $g \in W^*$에 대해서 $T^t(g) = gT$

T^t는 선형변환이고, $[T^t]_{\gamma^*}^{\beta*} = ([T]_\beta^\gamma)^t$가 성립함.

Proof. T^t가 선형변환임은 자명함.

순서기저를 $\beta = \{x_1, \cdots, x_n\}$, $\gamma = \{y_1, \cdots, y_m\}$이라 하고, 각각의 쌍대기저를 $\beta^* = \{f_1, \cdots, f_n\}$, $\gamma^* = \{g_1, \cdots, g_m\}$이라 하자. T^t의 행렬표현을 구하려 함.

$g_j T \in V^*$이므로 Theorem 2.24에 의해 아래와 같이 정리할 수 있음.

$$T^t(g_j) = g_j T = \sum_{i=1}^{n} (g_j T)(x_i) f_i$$

f_i가 기저이므로 $(g_j T)(x_i)$를 행렬의 원소로서 정리하면 됨. 편의를 위해 $[T]_\beta^\gamma = A$라고 함.

$$(g_j T)(x_i) = g_j(T(x_i)) = g_j(\sum_{k=1}^{m} A_{ki} y_k) = \sum_{k=1}^{m} A_{ki} g_j(y_k) = \sum_{k=1}^{m} A_{ki} \delta_{jk} = A_{ji}$$

즉, $[T^t]_{\gamma^*}^{\beta^*} = A^t$임. □

3) V와 V^{**}의 관계

Lemma 유한차원 벡터공간 V와 $x \in V$를 생각하자. 임의의 $f \in V^*$에 대해서 $\hat{x}(f) = 0$이면 $x = 0$임.

Proof. $x \neq 0$이면 $\hat{x}(f) \neq 0$인 어떤 $f \in V^*$가 존재함을 보이면 됨. V의 순서기저를 $\beta = \{x_1, \cdots, x_n\}$, β의 쌍대기저를 $\beta^* = \{f_1, \cdots, f_n\}$이라고 하자. $x_1 \neq 0$이고 $\hat{x_1}(f_1) = f_1(x_1) = 1 \neq 0$이므로 성립함. □

Theorem 2.26 유한차원 벡터공간 V에 대해서 함수 $\psi : V \to V^{**}$를 $\psi(x) = \hat{x}$이라 정의하면, ψ는 동형사상임.

Proof. 1. ψ는 선형인가?
$x, y \in V$와 $c \in F$를 고정하자. $f \in V^*$에 대해서 아래가 성립함.

$$\psi(cx+y)(f) = f(cx+y) = cf(x) + f(y) = c\hat{x}(f) + \hat{y}(f) = (cx+y)(f) = (c\psi(x) + \psi(y))(f)$$

$\psi(cx+y) = c\psi(x) + \psi(y)$이므로 ψ는 선형임.

2. ψ는 단사인가?
Lemma에 의해 0으로 가는 것은 0밖에 없으므로 단사임.

3. ψ는 동형사상인가?
V와 V^{**}의 차원이 같고, 단사이므로 ψ는 동형사상임. □

Corollary 유한차원 벡터공간 V와 쌍대공간 V^*를 생각하자. V^*의 모든 순서기저는 V의 어떤 기저의 쌍대기저임.

Part III

기본행렬연산과 연립일차방정식

3장에서는 벡터공간, 선형변환과 행렬에 대한 지식을 바탕으로 연립일차방정식을 완전하게 분석함.

1. 기본행렬연산과 기본행렬

1. 기본연산(elementary operation)

1) 정의

Definition 30. *$m \times n$ 행렬 A에 대해서 A의 행[열]에 대한 아래의 세 연산을 기본행[열]연산(elementary row[column] operation)이라 함.*

1. *A의 두 행[열]을 교환하는 것.*
2. *A의 한 행[열]에 영이 아닌 스칼라를 곱하는 것.*
3. *A의 한 행[열]에 다른 행[열]의 스칼라 배를 더하는 것.*

행연산(row operation)과 열연산(column operation)을 통틀어 기본연산(elementary operation)이라 함. 기본연산의 1, 2, 3을 각각 1형(type), 2형, 3형이라 함.

2. 기본행렬(elementary matrix)

1) 정의

Definition 31. *$n \times n$ 기본행렬(elementary matrix)은 항등행렬 I_n에 기본연산을 한 번 적용하여 얻은 행렬임. I_n에 1형, 2형, 3형 연산을 하여 얻은 행렬을 각각 1형, 2형, 3형이라고 함.*

2) 기본연산의 적용

Theorem 3.1에 의하면, 기본연산을 적용하는 것은 그 행렬에 적절한 기본행렬을 곱하는 것과 같음. 이 덕분에 기본연산을 적용하는 것을 수식적으로 나타낼 수 있음.

3. 관련 정리

1) 기본연산과 기본행렬의 곱

Theorem 3.1 행렬 $A \in M_{m\times n}(F)$에 기본행[열]연산을 하여 행렬 B를 얻었다면, $B = EA[B = AE]$가 되는 $m \times m[n \times n]$ 기본행렬 E가 존재함. 이때, A에서 B를 얻은 기본행[열]연산을 $I_m[I_n]$에 똑같이 적용하면 행렬 E가 됨. 역으로 E가 $m \times m[n \times n]$ 기본행렬일 때, $I_m[I_n]$에서 E를 얻은 기본행[열]연산을 A에 똑같이 적용하면 $EA[AE]$가 됨.

즉, 기본행렬을 곱하는 것은 그 기본행렬에 해당하는 기본연산을 적용하는 것과 같음.

2) 기본행렬의 가역성

Theorem 3.2 기본행렬은 가역임. 그 역행렬은 같은 종류의 기본행렬임.

Proof. 기본행렬을 만든 연산을 거꾸로 수행하면(해당하는 기본행렬 곱하면) 항등행렬이 나옴. □

2. 행렬의 랭크

행렬의 랭크를 구할 수 있어야 해당 연립일차방정식을 완전히 이해할 수 있음. 또한, 행렬의 랭크를 통해 역행렬/역변환을 구할 수 있음.

1. 행렬의 랭크

1) 정의

Definition 32. *행렬 $A \in M_{m\times n}(F)$에 대해서 A의 랭크(rank)는 선형변환 $L_A : F^n \to F^m$의 랭크로 정의하고 $rank(A)$라 표기함.*

2) 행렬의 랭크와 가역성

Theorem $n\times n$ 행렬이 가역이기 위한 필요충분조건은 행렬의 랭크가 n인 것임. 행렬이 가역이 아니기 위한 필요충분조건은 행렬의 랭크가 n보다 작은 것임.

그렇기 때문에, 정사각행렬의 행렬의 랭크를 구하면 그 행렬이 가역인지를 판단할 수 있음.

Proof. 1. $n \times n$ 행렬 A가 가역 $\to$ 행렬 A의 랭크가 n.
Theorem 2.18의 Corollary에 의해, A가 가역이면 L_A도 가역임. 즉 $dim(F^n) = rank(L_A) = n$임. L_A의 랭크가 행렬 A의 랭크이므로 성립.

2. 행렬 A의 랭크가 $n \to n \times n$ 행렬 A가 가역.
$rank(L_A) = dim(F^n)$이고, $L_A : F^n \to F^n$으로 정의역과 공역의 차원이 같으므로 행렬 A는 가역임. □

3) 행렬의 랭크와 선형변환의 랭크

Theorem 3.3에 의하면, 선형변환의 랭크와 그 행렬표현의 랭크는 동일함.

그러므로, 선형변환의 랭크를 찾는 문제는 그 행렬표현의 랭크를 찾는 문제와 귀결됨.

2. 행렬의 랭크 구하기

1) 행렬의 랭크를 보존하는 연산

Theorem 3.4에 의하면, 가역행렬의 곱은 행렬의 랭크를 보존하는 연산임.

Theorem 3.4의 Corollary에 의해, 행렬의 기본연산은 랭크를 보존함.
즉, 행렬에 기본연산을 적용하여(기본행렬을 곱하여) 랭크를 구하기 더 쉬운 형태로 바꿀 수 있음.

2) 행렬의 랭크 구하기

정리하면 행렬의 랭크는 아래의 방법으로 구할 수 있음.

1. 기본연산으로 행렬을 정리함. (Theorem 3.4)
2. 일차독립인 열의 개수를 확인함. (Theorem 3.5, Theorem 3.6, Theorem 3.6 Corollary 2)

즉, 행렬의 일차독립인 행 또는 열이 보일 때까지 기본연산을 적용하여 간단히 만드는 것.

4. 관련 정리

1) 행렬의 랭크과 선형변환의 랭크

Theorem 3.3 유한차원 벡터공간 사이에서 정의된 선형변환 $T : V \to W$와 V, W 각각의 순서기저 β, γ에 대해서 $rank(T) = rank([T]_\beta^\gamma)$임.

즉, 선형변환의 랭크과 그 행렬표현의 랭크는 동일함.

2) 행렬의 랭크를 보존하는 연산

Theorem 3.4 $m \times n$ 행렬 A, $m \times m$ 가역행렬 P, $n \times n$ 가역행렬 Q에 대해서 아래가 성립함.

1. $rank(AQ) = rank(A)$
2. $rank(PA) = rank(A)$
3. $rank(PAQ) = rank(A)$

즉, 가역행렬의 곱은 행렬의 랭크를 보존하는 연산임.

Corollary 행렬의 기본행연산과 기본열연산은 랭크를 보존함.

Proof. 기본연산은 기본행렬을 곱하는 것인데, 기본행렬은 정사각행렬인 가역행렬이므로 행렬의 랭크를 보존함. □

3) 행렬의 랭크

Theorem 3.5 임의의 행렬의 랭크는 일차독립인 열의 최대 개수와 같음. 즉, 행렬의 랭크는 그 열에 의해 생성된 부분공간의 차원임.

즉, 행렬의 각 열을 하나의 벡터로 생각했을 때, 일차독립인 열들의 집합을 만들면 그 개수가 곧 랭크임.

행렬의 열은 곧 기저를 보낸 것을 의미하는데, 상공간 생성 방법을 생각해 보면 이 정리는 매우 당연함.

4) 행렬의 랭크를 구하기 위한 구체적 방법

Theorem 3.6 랭크가 r인 $m \times n$ 행렬 A를 생각하자. $r \le m$, $r \le n$이 성립하고 기본행연산과 기본열연산을 유한 번 사용하여 A를 아래와 같은 꼴로 바꿀 수 있음.

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

이때, $i \le r$이면 $D_{ii} = 1$, 그렇지 않으면 $D_{ij} = 0$이고 O_1, O_2, O_3은 영행렬임.

즉, 행렬에 기본연산을 유한 번 사용해 왼쪽 위가 I_r이고 나머지는 0인 행렬로 만들 수 있음. 이 꼴로 만들면 일차독립인지를 확인하는 것이 굉장히 간단해짐.[27]

증명은 프리드버그 p.179에 있지만 굳이 정리하지 않겠음.

[27] 물론 정확히 이렇게 만들 필요는 없고, 랭크를 구할 수 있을 정도까지만 연산을 하면 됨.

Corollary 1 Theorem 3.5의 행렬 A에 대해서, $D = BAC$를 만족하는 $m \times m$ 가역행렬 B, $n \times n$ 가역행렬 C가 존재함.

즉, 행렬에 기본행렬을 곱해 D로 만들 수 있다는 것.

Corollary 2 $m \times n$ 행렬 A에 대해 아래가 성립함.

1. $rank(A^t) = rank(A)$
2. 임의의 행렬의 랭크는 일차독립인 행의 최대 개수와 같음. 행렬의 랭크는 그 행에 의해 생성된 부분공간의 차원임.
3. 임의의 행렬의 행과 열은 차원이 같은 부분공간을 생성함. 각각의 차원은 행렬의 랭크와 같음.

Proof. (1) Corollary 1에 의하면 $D = BAC$인데, $D^t = (BAC)^t = C^t A^t B^t$임. C^t, B^t는 가역이므로 $rank(C^t A^t B^t) = rank(A^t) = rank(D^t)$인데, $rank(D^t) = rank(D) = rank(A)$임.

(2) 전치해 보면 확인할 수 있음.

(3) Theorem 3.5, Theorem 3.6의 Corollary 2 (1), (2)를 보면 알 수 있음. □

Corollary 3 모든 가역행렬은 기본행렬의 곱으로 나타남.

Proof. $n \times n$ 가역행렬 A의 랭크는 n임. $D = I_n = BAC$임. B와 C는 각각 $B = E_p E_{p-1} \cdots E_1$, $C = G_1 G_2 \cdots G_q$를 만족하는 기본행렬 E_i, G_i가 존재함. 정리하면 아래와 같음.

$$A = B^{-1} I_n C^{-1} = B^{-1} C^{-1} = E_1 E_2 \cdots E_p G_q G_{q-1} \cdots G_1$$

즉, 행렬 A는 기본행렬의 곱으로 나타낼 수 있음. □

5) 합성과 행렬 곱에 따른 랭크

Theorem 3.7 유한차원 벡터공간 V, W, Z 사이에 정의된 선형변환 $T : V \to W$, $U : W \to Z$와 행렬 곱 AB를 정의하는 두 행렬 A, B에 대해서 아래가 성립함.

1. $rank(UT) \leq rank(U)$
2. $rank(UT) \leq rank(T)$
3. $rank(AB) \leq rank(A)$
4. $rank(AB) \leq rank(B)$

즉, 선형변환의 합성 또는 행렬의 곱은 랭크를 더 커지게 할 수 없음.

Proof. (1) $R(UT) = (UT)(V) = U(T(V)) \subseteq U(W) = R(U)$임.

(3) (1)이 성립하므로 선형변환을 행렬표현으로 나타내면 성립함을 확인할 수 있음.

(4) (3)이 성립하므로 $rank(AB) = rank(B^{-1}A^{-1}) \leq rank(A^{-1}) = rank(A)$으로 성립함.

(2) (4)가 성립하므로 행렬표현을 선형변환으로 나타내면 성립함을 확인할 수 있음.

□

3. 역행렬 구하기

기본연산과 행렬의 랭크로 역행렬/역변환을 구할 수 있음.

1. 첨가행렬(augmented matrix)

1) 정의

Definition 33. *$m \times n$ 행렬 A와 $m \times p$ 행렬 B에 대해서 첨가행렬(augmented matrix) $(A|B)$는 $m \times (n+p)$ 행렬 $(A\,B)$임. 즉, 처음 n개 열은 A의 열이고, 그 다음 p개 열은 B의 열인 행렬임.*

쉽게 말해, 행렬 A의 오른쪽에 B를 그대로 붙인 행렬을 첨가행렬 $(A|B)$라 하는 것.

2) 행렬과 첨가행렬의 곱

n개의 행을 가지는 행렬 A, B와 $m \times n$ 행렬 M에 대해서 아래가 성립함.

$$M(A|B) = (MA|MB)$$

즉, 왼쪽에 곱한 행렬의 연산이 분배법칙처럼 각각 적용됨.

2. 기본행연산으로 역행렬 구하기

1) 역행렬 구하기

Theorem A가 $n \times n$ 가역행렬이면 행렬 $(A|I_n)$에 기본행연산을 유한 번 적용해서 $(I_n|A^{-1})$로 변형할 수 있음.

즉, $(A|I_n)$에 기본행연산을 하여 $(I_n|A^{-1})$을 만들 수 있다는 것.

Proof. $A^{-1}(A|I_n) = (A^{-1}A|A^{-1}I_n) = (I_n|A^{-1})$임. Theorem 3.6의 Corollary 3에 의해, A^{-1}는 정사각행렬이므로 기본행렬의 곱으로 나타낼 수 있음. 그런데 왼쪽에 곱해진 기본행렬은 기본행연산이므로, $(A|I_n)$에 기본행연산을 하여 $(I_n|A^{-1})$을 만들 수 있음. □

2) 변형이 되면 역행렬임

Theorem $n \times n$ 가역행렬 A에 대해서, 첨가행렬 $(A|I_n)$에 기본행연산을 유한 번 적용하여 $(I_n|B)$로 변형할 수 있으면 $B = A^{-1}$임.

즉, 첨가행렬을 기본행연산으로 변형해서 일단 $(I_n|B)$꼴을 만들면 B가 역행렬임. (다른 이상한 행렬이 나오지 않음.)

Proof. 행렬 $C = E_p E_{p-1} \cdots E_1$일 때, $C(A|I_n) = (CA|C) = (I_n|B)$임. $CA = I_n$, $C = B$이므로 $B = A^{-1}$임. □

3) 가역이 아닌 경우

Theorem 가역인 아닌 $n \times n$ 행렬 A에 대해서, $(A|I_n)$에 기본행연산을 적용하여 $(I_n|B)$ 꼴로 변형을 시도하면 성공하지 못하고 앞쪽 n개 성분이 모두 0인 행을 가진 행렬을 얻게 됨.

Proof. A가 가역이 아니므로 $rank(A) < n$임. 유한 번의 기본연산으로 $(A|I_n)$을 $(I_n|B)$로 바꿀 수 있다고 가정하면, 유한 번의 기본연산으로 A을 I_n로 바꿀 수 있어야 하는데 기본연산은 랭크를 보존하므로 $rank(A) = rank(I_n) = n$으로 모순임. 즉, 유한 번의 기본연산으로 $(A|I_n)$을 $(I_n|B)$로 바꿀 수 없음. □

3. 2×2 행렬에서 역행렬 구하기

1) 2×2 행렬에서 역행렬 구하기

Part 4의 Theorem 4.2에 의하면, 2×2 행렬에서는 특수한 방식으로 역행렬을 구할 수 있음.

4. 역변환 구하기

1) 역변환 구하기

선형변환 행렬표현의 가역성을 확인한 후 역행렬을 구하면, Theorem 2.18에 의해 $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$으로 해당 역행렬은 해당 선형변환 역변환의 행렬표현임.

역행렬에 임의의 벡터를 넣는 방식으로 역변환을 알아낼 수 있음.

아래는 그 예시임. $T : P_2(R) \to P_2(R)$, $P_2(R)$의 표준 순서기저를 β라 함.

$$[T^{-1}(a_0 + a_1x + a_2x^2)]_\beta = [T^{-1}]_\beta[(a_0 + a_1x + a_2x^2)]_\beta = \begin{pmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 \end{pmatrix}$$

즉, $T^{-1}(a_0 + a_1x + a_2x^2) = (a_0 - a_1) + (a_1 - 2a_2)x + a_2x^2$임.

4. 연립일차방정식 : 이론적 측면

연립일차방정식이 나오면 두 가지 질문에 완벽히 답할 수 있어야 함.
1. 주어진 연립방정식에 해가 있는가? (Theorem 3.11)
2. 해가 있다면 모든 해(해집합)를 어떻게 구할 수 있는가? (Theorem 3.8, Theorem 3.9, Theorem 3.10)

1. 연립일차방정식(system of linear equations)

1) 정의

Definition 34. *아래의 형태를 체 F 위 n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식(system of linear equations)이라 함.*

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{gathered}$$

이때, a_{ij}와 b_i는 F의 스칼라이고, x_i는 F에서 값을 가지는 변수임.

2) 계수행렬(coefficient matrix)

Definition 35. *아래의 $m \times n$ 행렬 A를 연립일차방정식 (S)의 계수행렬(coefficient matrix)이라 함.*

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

이때, 아래와 같이 정의하면 연립일차방정식 (S)는 하나의 행렬 식(matrix equation) $Ax = b$로 나타낼 수 있음.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}$$

3) 해집합(solution set)

Definition 36. *$As = b$인 n순서쌍 s를 연립일차방정식 (S)의 해(solution)라 하고, 연립일차방정식 (S)가 가지는 모든 해들의 집합을 해집합(solution set)이라 함.*

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \cdots \\ s_n \end{pmatrix} \in F^n$$

해집합이 공집합이 아니면 이 연립일차방정식을 모순이 없다(consistent) 또는 해가 존재한다고 함.
해집합이 공집합이면 이 연립일차방정식을 모순이 있다(inconsistent) 또는 해가 존재하지 않는다고 함.

연립일차방정식은 해가 하나이거나, 해가 무한히 많거나, 해가 없음.

2. 동차(homogeneous)/비동차(non-homogeneous)[28]

1) 정의

Definition 37. *n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식 $Ax = b$는 $b = 0$일 때, 동차(homogeneous)라 함. 동차가 아닌 연립방정식은 비동차(non-homogeneous)라 함.*

임의의 동차 연립일차방정식에는 적어도 하나의 해가 있음. (영벡터)

2) 동차 연립일차방정식의 해집합 구하기

연립일차방정식의 해집합을 구하는 것은 대응하는 동차 연립일차방정식의 해를 구하는 것에서부터 시작함.

Theorem 3.8에 의하면 동차 연립일차방정식의 해집합은 영공간(부분공간)이므로, 기저 하나만 찾으면 그 해집합을 알 수 있음.

기저를 찾는 가장 간단한 방법은 해집합의 차원을 찾고, $nullity(L_A)$개의 해를 대충 맞춰서 찾아내는 것임.

해집합의 차원을 쉽게 찾는 방법은 아래와 같음.

1. 행렬의 열의 개수를 세면 그게 $dim(V)$임.
2. 행렬의 랭크를 알아냄.
3. dimension theorem에 의해 열의 개수와 랭크를 빼서 해집합의 차원을 구함.

기저를 찾는 더 구체적이고 형식적인 방법은 '5. 연립방정식 : 계산적 측면'에서 다룸.

3) 비동차 연립일차방정식의 해집합 구하기

Theorem 3.9에 의하면, 비동차 연립일차방정식의 해집합은 대응하는 동차 연립일차방정식의 해집합으로 알아낼 수 있음.

$Ax = 0$을 $Ax = b$에 대응하는 동차 연립일차방정식이라고 함.

4) 계수행렬이 가역인 연립일차방정식

Theorem 3.10에 의해, 계수행렬이 가역인 연립일차방정식은 유일한 해를 간단히 구할 수 있음.

5) 연립일차방정식이 해를 가지는지 판별하기

Theorem 3.11에 의해, $Ax = b$에서 $rank(A) = rank(A|b)$인지를 보면 해를 가지는지 판별할 수 있음.

[28] 연립일차방정식의 풀이는 동차 연립일차방정식부터 시작하여, 동차/비동차 연립일차방정식의 해집합을 부분공간으로 묘사하는 것임.

3. 관련 정리

1) 동차 연립일차방정식의 해집합

Theorem 3.8 체 F에서 n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식 $Ax = 0$을 생각하자. 방정식 $Ax = 0$의 해집합을 K라 할 때, $K = N(L_A)$임. 즉, K는 F^n의 부분공간이고 차원은 $n - rank(L_A) = n - rank(A)$임.

즉, 동차 연립일차방정식의 해집합은 L_A의 kernel(null space)이고, 그 차원은 $n - rank(A)$임.

Proof. L_A는 왼쪽에 A를 곱하는 선형변환이므로, $Ax = 0$인 x는 kernel(space)에 속함. 차원은 차원정리로 생각할 수 있음. □

Corollary $m < n$이면 연립일차방정식 $Ax = 0$은 영벡터가 아닌 해가 있음.

Proof. $Ax = 0$의 해집합이 $N(L_A)$이므로, $Ax = 0$이 영벡터가 아닌 해를 가진다는 것은 $N(L_A)$ 영벡터가 아닌 원소를 가진다는 것임. 즉, $nullity(L_A) \neq 0$이어야 함. $rank(A) = rank(L_A) \leq m$이므로, $nullity(L_A) = n - rank(L_A) \geq n - m > 0$으로 $nullity(L_A) \neq 0$임. □

2) 비동차 연립일차방정식의 해집합

Theorem 3.9 모순이 없는 연립일차방정식 $Ax = b$의 해집합을 K, 대응하는 연립일차방정식 $Ax = 0$의 해집합을 K_H라 하자. $Ax = b$의 임의의 해를 s라 하면 아래가 성립함.

$$K = \{s\} + K_H = \{s + k | k \in K_H\}$$

즉, $Ax = b$의 임의의 해 하나를 고정하고, 그 해와 $Ax = 0$의 해집합의 원소들을 각각 더한 벡터들이 $Ax = b$의 해집합이라는 것.

Proof. 1. $K \subseteq \{s\} + K_H$
$w, s \in K$에 대해서, $Aw = b, As = b, A(w - s) = 0$이므로 $k = w - s \in K_H$임. 즉, $w = s + k \in \{s\} + K_H$이고 $K \subseteq \{s\} + K_H$임,

2. $\{s\} + K_H \subseteq K$
$w \in \{s\} + K_H$, $k \in K_H$에 대해서, $w = s + k$, $Aw = As + Ak = b$이므로 $w \in K$임. 즉, $\{s\} + K_H \subseteq K$임.

$K \subseteq \{s\} + K_H$이고 $\{s\} + K_H \subseteq K$이므로 $K = \{s\} + K_H$임. □

3) 행렬의 가역성과 유일한 해

Theorem 3.10 n개의 미지수와 n개의 일차방정식으로 이루어진 연립일차방정식 $Ax = b$를 생각하자. 행렬 A가 가역이면 이 연립일차방정식은 유일한 해 $A^{-1}b$가 있음. 역으로, 이 방정식의 해가 유일하면 행렬 A는 가역임.

즉, $n \times n$ 행렬이 가역이면 연립일차방정식이 유일한 해($A^{-1}b$)를 가지고, 연립일차방정식이 유일한 해를 가지면 $n \times n$ 행렬이 가역임.

Proof. 1. $n \times n$ 행렬이 가역이면 연립일차방정식이 유일한 해($A^{-1}b$)를 가짐. $AA^{-1}b = b$. $s \in K$인 s가 존재한다고 가정하자. $A^{-1}As = A^{-1}b, s = A^{-1}b$이므로 연립일차방정식이 유일한 해를 가짐.

2. 연립일차방정식이 유일한 해를 가지면 $n \times n$ 행렬이 가역임. $As = b$인 $s \in K$가 유일하다고 하자. Theorem 3.9에 의해 $\{s\} = \{s\} + K_H$이므로 $K_H = \{0\}$임. 즉, $n = rank(L_A) + nullity(L_A) = rank(A)$이므로 A는 가역임. □

4) 연립일차방정식이 해를 가지는지 판별하기

Theorem 3.11 연립일차방정식 $Ax = b$에 모순이 없기 위한 필요충분조건은 $rank(A) = rank(A|b)$인 것임.

Proof. 모순이 없음 $\Leftrightarrow$ $Ax = b$가 해를 가짐 $\Leftrightarrow$ $b \in R(L_A)$임. $\Leftrightarrow$ $R(L_A) = span(\{a_1, \cdots, a_n\})$($a_i$는 A의 i번째 열)이므로, $b \in span(\{a_1, \cdots, a_n\}) \Leftrightarrow span(\{a_1, \cdots, a_n\}) = span(\{a_1, \cdots, a_n, b\})$. 즉, b가 a_i의 일차결합으로 표현됨. $\Leftrightarrow dim(span(\{a_1, \cdots, a_n\})) = dim(span(\{a_1, \cdots, a_n, b\})) \Leftrightarrow rank(A) = rank(A|b)$임.

□

5. 연립일차방정식 : 계산적 측면

기본행연산을 사용하여 연립방정식의 모든 해를 찾을 수 있음.

1. 동치(equivalent)

1) 정의

Definition 38. *두 연립일차방정식의 해집합이 서로 같을 때, 두 연립일차방정식은 동치(equivalent)라 함.*

어떤 연립일차방정식의 해집합을 구할 때, 동치인 더 쉬운 연립일차방정식으로 바꾸어 구하는 것이 더 쉬움.

2) 동치인 연립일차방정식으로 전환하기

Theorem 3.13의 Corollary에 의하면, 연립일차방정식 $Ax = b$에 대해 $(A|b)$에 기본행연산을 적용한 $(A'|b')$의 $A'x = b'$가 $Ax = b$와 동치임.

즉, $(A|b)$에 기본행연산을 적용하여 더 쉬운 연립일차방정식으로 바꾸어 해를 구할 수 있음.

2. 행간소사다리꼴(기약행사다리꼴, RREF, reduced row echelon form)

1) 정의

Definition 39. *아래의 세 조건을 만족하는 행렬을 행간소사다리꼴 또는 기약행사다리꼴(reduced row echelon form)이라고 함.*

1. *0이 아닌 성분을 가지는 행은 모든 성분이 0인 행보다 위에 위치함.*
2. *각 행의 처음으로 0이 아닌 성분은 그 성분을 포함하는 열에서 유일하게 0이 아닌 성분임.*
3. *각 행에서 처음으로 0이 아닌 성분은 1이고, 이전 행의 처음으로 0이 아닌 성분보다 오른쪽에 위치함.*

$(A|b)$꼴 연립일차방정식을 기본행연산을 통해 행간소사다리꼴로 바꾸면 계산이 굉장히 간단해짐. 또한, 행간소사다리꼴을 이용하여 해집합과 해의 존재 유무를 알아낼 수 있음. 즉, 그 연립일차방정식의 모든 것을 알게됨.

Theorem 3.16의 Corollary에 의하면 어떤 행렬에 대해서 그 행렬의 행간소사다리꼴은 유일함.

2) 행간소사다리꼴로 해집합 구하기

행간소사다리꼴로 원래 연립일차방정식의 해집합을 구하는 방법은 아래와 같음.

1. 모든 성분이 0인 행은 무시함.

2. 각 행에 대응하는 일차방정식에서, 가장 왼쪽에 있는 변수들에 매개변수를 부여함. 또한, 연립방정식에 영향을 주지 못하는 변수에도 매개변수를 부여함.[29]

3. 나머지 변수들을 매개변수를 부여한 변수들로 나타냄.

4. 매개변수로 표현한 변수들로 해를 나타내고, 매개변수에 대해서 정리함. 이것이 원래 연립일차방정식의 임의의 해임. 이때, 매개변수에 곱해져 있는 행렬들의 집합은 동차 연립일차방정식 해집합의 기저임. 또한 매개변수가 곱해져 있지 않은 행렬은 비동차 연립일차방정식의 한 해임. 이를 특수해(particular solution) 라고 함.

이때, 동차 연립일차방정식의 차원을 계산하여 정말 기저인지 확인해야 함.

아래는 이에 대한 예시임.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2t_1 + 2t_2 + 3 \\ t_1 - t_2 + 1 \\ t_1 \\ 2t_2 + 2 \\ t_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

정리하면, 연립일차방정식의 첨가행렬을 행간소사다리꼴로 바꾸면 아래의 두 가지를 알아낼 수 있음.
1. 처음 연립일차방정식($Ax = b$)의 특수해(한 해).
2. 처음 연립일차방정식에 대응하는 동차 연립일차방정식의 해집합의 기저.

즉, 처음 연립일차방정식의 해집합을 알아낼 수 있음.

3) 행간소사다리꼴로 해의 존재 유무 판정하기

Theorem 연립일차방정식의 해가 존재하기 위한 필요충분조건은 첨가행렬을 정리하여 행간소사다리꼴을 만들 때, 0이 아닌 유일한 성분이 마지막 열에 있는 행이 존재하지 않는 것임.

증명은 프리드버그 p.220 3번 문제 참고.

행간소사다리꼴로 만들기만 해도 해의 존재 유무를 판정할 수 있음.

3. 가우스 소거법(Gaussian elimination)[30]

1) 정의

> **Definition 40.** *1. 위에서 아래로*
> *위에서 아래로 내려가면서 기본행연산으로 첨가행렬을 변형하여 각 행의 최초로 0이 아닌 성분은 1 이고, 이 성분이 이전 행의 최초로 0이 아닌 성분보다 오른쪽 열에 위치하는 행렬(상삼각행렬)로 만듦. (정의 1,3 만족)*
>
> *2. 아래에서 위로*
> *아래에서 위로 올라가면서 행렬(상삼각행렬)을 기본행연산으로 변형하여 각 행의 최초로 0이 아닌 성분이 이 성분을 포함하는 열에서 유일하게 0이 아닌 성분인 행간소사다리꼴로 만듦. (정의 2 만족)*

Theorem 3.14에 의하면, 가우스 소거법은 첨가행렬을 행간소사다리꼴로 만듦.

행렬을 행간소사다리꼴로 변환할 때, 산술 연산을 가장 적게 하는 방법이 가우스 소거법임.

모든 행렬은 가우스 소거법을 사용하여 행간소사다리꼴로 바꿀 수 있음.

[29] $x_1 = t_1, x_3 = t_2$등으로 설정함.
[30] 가우스-조던 소거법에 대한 내용이지만, 프리드버그에서는 가우스 소거법으로 명명하므로 동일하게 정리함.

4. 일반해(general solution)

1) 정의

Definition 41. *n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식 $Ax = b$에 대해서, $(A|b)$를 기본행연산을 사용하여 행간소사다리꼴 $(A'|b')$으로 바꾸고 0이 아닌 행의 변수들에 대해 매개변수를 설정하면 임의의 해의 꼴을 얻을 수 있음. 이때, r은 A'의 0이 아닌 행의 개수임.*

$$s = s_0 + t_1 u_1 + t_1 u_1 + \cdots + t_{n-r} u_{n-r}$$

이 식을 연립일차방정식 $Ax = b$에 대한 일반해(general solution)라고 함.

2) 의미

Theorem 3.15에 따르면, r은 A의 랭크이고, $n-r$은 해집합의 차원이고, s_0은 원래 비동차 연립일차방정식의 특수해이고, $u_1, u_2, \cdots, u_{n_r}$은 원래 연립일차방정식에 대응하는 동차 연립일차방정식의 기저임.

5. 행간소사다리꼴에 대한 해석

1) 행간소사다리꼴에 대한 해석

Theorem 3.16에 의하면, 행간소사다리꼴을 확인하여 원래 행렬에서 서로 일차독립인 열들을 확인할 수 있음. 또한 행간소사다리꼴의 열로 원래 행렬의 열을 알아낼 수도 있음.

2) 일차독립 판정하기

Theorem 3.16을 이용하면, 행렬의 행간소사다리꼴을 확인하여 일차독립 여부를 판정할 수 있음.

F^n의 원소들의 집합에 대해서, 각 원소를 열로 하는 행렬을 생각했을 때 이 행렬의 행간소사다리꼴로 일차독립인 부분집합을 알아낼 수 있음.

유한차원 벡터공간 V와 그 기저 β에 대해서, V의 일차독립인 부분집합 S를 확장하여 V의 기저를 얻을 수 있음. $S \cup \beta$의 각 원소를 열로 하는 행렬을 생각했을 때 이 행렬의 행간소사다리꼴로 일차독립인 부분집합을 알아낼 수 있음.

6. 연립일차방정식의 풀이 총정리

1) 연립일차방정식의 풀이 총정리

이론적/계산적 측면을 모두 반영한 연립일차방정식의 풀이에 대한 요약은 아래와 같음.

0. $Ax = b$를 행간소사다리꼴(RREF)로 만듦.

해의 존재 여부 판단
1. $rank(A) = rank(A|b)$인지 확인
2. A가 가역인지 확인($A^{-1}b$)
3. 마지막 열에 0이 아닌 유일한 성분을 가지는 행이 존재하는지 확인

해집합 구하기 방법1
1. 동차 연립일차방정식 해집합의 차원을 계산하고, 기저를 '대충 때려넣어서' 만듦.
2. 비동차 연립일차방정식의 특수해를 '대충 때려넣어서' 만듦.
3. 일반해를 구함.

해집합 구하기 방법2
1. 맨 앞에 있거나 영향을 주지 못하는 변수들에 매개변수를 부여함. 자유변수에 매개변수를 부여하여 선행변수를 자유변수에 대해 정리할 수도 있는데, 이 방법이 더 간단함.
2. 정리하여 일반해를 구함. 이때, 동차 연립일차방정식의 차원을 계산하여 정말 기저인지 확인해야 함.

7. 관련 정리

1) 동치인 두 연립일차방정식

Theorem 3.13 n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식 $Ax = b$와 $m \times m$ 가역행렬 C에 대해서, 아래의 두 연립일차방정식은 동치임.

$$Ax = b,\ (CA)x = Cb$$

Proof. $Ax = b$의 해집합을 K, $(CA)x = Cb$의 해집합을 K'이라 하자.

1. $K \subseteq K'$
$w \in K$가 있음. $Aw = b$이므로, $(CA)w = Cb$, $w \in K'$임. 즉, $K \subseteq K'$임.

2. $K' \subseteq K$
$w \in K'$가 있음. $(CA)w = Cb, C^{-1}(CA)w = C^{-1}Cb, Aw = b, w \in K$임. 즉, $K' \subseteq K$임.

즉, $K = K^{-1}$임. □

Corollary n개의 미지수와 m개의 일차방정식으로 이루어진 연립일차방정식 $Ax = b$가 있음. $(A|b)$에 기본행연산을 유한 번 적용하여 얻은 $(A'|b')$은 처음 주어진 연립일차방정식과 동치임.

2) 가우스 소거법

Theorem 3.14 가우스 소거법은 임의의 행렬을 행간소사다리꼴로 바꾸어 줌.

3) 행간소사다리꼴로 얻은 일반해의 의미

Theorem 3.15 $Ax = b$를 n개의 미지수와 r개의 영이 아닌 방정식[31]으로 이루어진 연립일차방정식이라 하자. $rank(A) = rank(A|b)$이고 $(A|b)$가 행간소사다리꼴이면 아래가 성립함.

(1) $rank(A) = r$

(2) 앞선 과정을 거쳐 얻은 일반해가 아래와 같은 꼴이라고 하자.

$$s = s_0 + t_1u_1 + t_1u_1 + \cdots + t_{n-r}u_{n-r}$$

이때, $\{u_1, u_2, \cdots, u_{n-r}\}$은 대응하는 동차 연립일차방정식의 해집합의 기저이고, s_0은 처음 연립일차방정식의 해임.

4) 행간소사다리꼴에 대한 해석

Theorem 3.16 랭크가 r인 $m \times n$ 행렬 A에 대해서(단, $r > 0$) 행간소사다리꼴을 B라 하면 아래가 성립함.

1. B의 영이 아닌 행의 개수는 r임.
2. 각 $i = 1, 2, \cdots, r$에 대해서 $b_{j_i} = e_i$인 B의 열 b_{j_i}가 존재함.(j_i는 적절한 인덱스 값을 가짐.)
3. A의 $j_1, j_2, \cdots, j_r$열은 일차독립임.
4. 각 $k = 1, 2, \cdots, n$에 대해서 B의 k열이 $d_1e_1 + d_1e_2 + \cdots + d_re_r$이면 A의 k열은 $d_1a_{j_1} + d_1a_{j_2} + \cdots + d_ra_{j_r}$

증명은 프리드버그 p.214를 참고하자.

Corollary 행렬의 행간소사다리꼴은 유일함.

[31] 행간소사다리꼴에서 모두 0인 행을 제외한 것.

Part IV

행렬식

4장에서는 가역성을 결정하는 행렬식(determinant)에 대한 이야기를 함.

1. 행렬식의 엄밀한 정의

교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$ 중 $\delta(I) = 1$인 δ는 행렬식임.

1. n-선형함수(n-linear function, multi-linear)

1) 정의

Definition 42. *$n \times n$ 행렬의 다른 행이 고정되어 있을 때, 행렬의 각 행에 대해서 선형인 함수 $\delta :$ $M_{n\times n}(F) \to F$를 n-선형함수(n-linear function, multi-linear)라고 함.*

즉, 모든 $r = 1, 2, \cdots, n$에 대해서, F^n에 속하는 임의의 세 벡터 u, v, a_i와 스칼라 k에 대해서 아래의 관계식을 만족하는 δ는 n-선형임.

$$\delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u+kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + k\delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

n-선형함수중 가장 주요한 것은 행렬식임.

n-선형함수는 선형은 아니지만, 일종의 선형성을 가짐.

2. 교대(alternating)

1) 정의

Definition 43. *이웃한 두 행이 서로 같은 행렬 $A \in M_{n\times n}(F)$에 대해서, $\delta(A) = 0$인 n-선형함수 $\delta : M_{n\times n}(F) \to F$를 교대(alternating)라고 함.*

즉, 이웃한 두 행이 서로 같은 행렬을 넣으면 0이 되는 n-선형함수를 교대라고 함.

Theorem 4.10에 의하면, 임의의 두 행이 같기만 해도 0이 됨.

2) 교대와 기본연산

Theorem 4.10의 Corollary 3과 Theorem 4.11에 의해, 교대인 n-선형함수에 넣는 행렬에 대해서 기본연산을 적용하는 것은 기본연산의 종류에 따라 상이한 효과가 발생함.

1형 기본연산은 값의 부호를 바꾸고, 2형 기본연산은 값에 k가 곱해지고, 3형 기본연산은 값을 보존함.

3. 행렬식의 엄밀한 정의

1) 행렬식을 정의하는 세 번째 성질

어떤 교대 n-선형함수가 특정 조건에서 행렬식임을 보이려고 할 때, 임의의 교대 n-선형함수의 스칼라 곱은 여전히 교대 n-선형함수임.[32] 즉, 어떤 스칼라 곱이 행렬식에 해당하는지 정의해야 함.

Theorem 4.12에 의하면, 이를 위한 행렬식을 정의하는 세 번째 성질은 $n \times n$ 항등행렬 I_n의 값이 1인 것임.

2) 행렬식의 엄밀한 정의

> **Definition 44.** *어떤 함수 $\delta : M_{n\times n}(F) \to F$가 행렬식이기 위한 조건은 아래와 같음.*
>
> 1. *δ가 n-선형함수(multi-linear)임.*
> 2. *δ가 교대(alternating)임.*
> 3. *$\delta(I) = 1$임.*

즉, 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$ 중 $\delta(I) = 1$인 δ는 행렬식임.

당연히 n-선형함수(multi-linear)와 교대(alternating)의 성질을 행렬식에 적용할 수 있음.

이렇게 정의된 행렬식은 정사각행렬 집합을 정의역으로 하고 스칼라를 함숫값으로 하는 특별한 함수임.

4. 관련 정리

1) 행렬식을 정의하는 세 번째 성질

Theorem 4.12 $\delta(I) = 1$인 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$를 생각하자. 모든 $A \in M_{n\times n}(F)$에 대해서 $\delta(A) = det(A)$임.

Proof. $\delta(I) = 1$인 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$와 $A \in M_{n\times n}(F)$를 생각하자.

A의 랭크가 n 미만이면(즉, A가 가역이 아님.) Theorem 4.10 Corollary 2, Theorem 4.7 Corollary에 의해 $\delta(A) = det(A) = 0$임.

A의 랭크가 n이면(즉, A가 가역임.) A는 기본행렬의 곱으로 나타낼 수 있음. $A = E_1E_2\cdots E_m$라 하자. Theorem 4.3, Theorem 4.5, Theorem 4.6[33], Theorem 4.10 Corollary 3에 의하면, 기본행렬 E에 대해 $\delta(E) = det(E)$가 성립함. 따라서 $\delta(A)$를 정리하면 아래와 같음.

$$\delta(A) = \delta(E_1E_2\cdots E_m) = det(E_1)\delta(E_2\cdots E_m) = det(E_1E_2\cdots E_m) = det(A)$$

모든 A에 대해서 $\delta(A) = det(A)$임. 즉, $\delta(I) = 1$인 교대 n-선형함수 δ는 행렬식임. □

[32] 이러면 행렬식 값의 개수가 무한 개가 됨.

[33] Theorem 4.3, Theorem 4.5, Theorem 4.6은 이 필기에 따로 정리하지 않았음.

2) 교대(alternating)의 성질

Theorem 4.10 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$에 대해서 아래가 성립함.

1. $A \in M_{n\times n}(F)$의 임의의 두 행을 교환하여 얻은 행렬을 B라고 하면, $\delta(B) = -\delta(A)$임.
2. $A \in M_{n\times n}(F)$의 임의의 두 행이 같으면 $\delta(A) = 0$임.

Proof. (1) 이웃한 두 행을 교환하는 경우는 아래의 식에서 확인할 수 있음.

$$0 = \delta\begin{pmatrix} a_1 \\ \vdots \\ a_r + a_{r+1} \\ a_r + a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \delta\begin{pmatrix} a_1 \\ \vdots \\ a_r \\ a_r \\ \vdots \\ a_n \end{pmatrix} + \delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r+1} \\ a_r \\ \vdots \\ a_n \end{pmatrix} + \delta\begin{pmatrix} a_1 \\ \vdots \\ a_r \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + \delta\begin{pmatrix} a_1 \\ \vdots \\ a_{r+1} \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = 0 + \delta(A) + \delta(B) + 0$$

이므로 $\delta(A) = -\delta(B)$임.

이웃하지 않은 두 행을 교환하는 경우에는 이웃한 두 행끼리의 반복 교환을 통해 확인할 수 있음.

(2) 두 행이 이웃한 경우, 교대의 정의에 의해 성립함.

두 행이 이웃하지 않은 경우, 동일한 행이 이웃하도록 행을 교환하면 교대의 정의에 의해 성립함. □

Corollary 1 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$가 있음. 행렬 $A \in M_{n\times n}(F)$의 어느 행의 스칼라 배를 다른 행에 더하여 얻은 행렬을 B라 하면 $\delta(B) = \delta(A)$임.

즉, 3형 기본행연산은 행렬식 값을 보존함.

Proof. B를 i행의 k배를 j행에 더한 행렬, 행렬 C를 A의 j행을 i행으로 바꾼 행렬이라고 하자. 교대 n-선형함수의 정의에 의해 $\delta(B) = \delta(A) + k\delta(C) = \delta(A)$임. □

Corollary 2 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$가 있음. $M \in M_{n\times n}(F)$의 랭크가 n 미만이면 $\delta(M) = 0$임.

Proof. 독립인 행의 개수가 랭크이므로, 랭크가 n 미만이라는 것은 어떤 행이 다른 행들의 일차결합으로 표현된다는 것임. 즉, 3형 기본연산을 적용하면 동일한 행이 2개 이상 존재하도록 만들 수 있음. 이럴 경우 Theorem 4.10에 의해 $\delta(M) = 0$임. □

Corollary 3 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$와 $M_{n\times n}(F)$에 속하는 1형 기본행렬 E_1, 2형 기본행렬 E_2, 3형 기본행렬 E_3를 생각하자. 특히, E_2는 I의 어떤 행에 영이 아닌 스칼라 k를 곱해서 얻은 행렬임. 이때, 아래가 성립함.

$$\delta(E_1) = -\delta(I),\ \delta(E_2) = k\delta(I),\ \delta(E_3) = \delta(I)$$

즉, 1형 기본연산은 행렬식 값의 부호를 바꾸고, 2형 기본연산은 행렬식 값에 k가 곱해지고, 3형 기본연산은 행렬식 값을 보존함.

Proof. Theorem 4.10의 (1), n-선형함수(multi-linear)의 정의, Theorem 4.6 참고. □

2. n차 정사각행렬의 행렬식

1. n차 정사각행렬의 행렬식

1) 정의[34]

Definition 45. *체 F의 원소를 성분으로 가지는 $n \times n$ 행렬 A의 행렬식(determinant)은 $det(A)$ 또는 $|A|$로 표기하고, 아래와 같이 계산할 수 있음.*

1. A가 1×1 행렬인 경우, $det(A) = A_{11}$임.

2. A가 2×2 행렬인 경우, $det(A) = A_{11}A_{22} - A_{12}A_{21}$임. $(ad - bc)$

3. A가 $n > 2$인 n에 대해서 $n \times n$ 행렬인 경우, 모든 i에 대해서 i행에 대한 여인수 전개를 하면 행렬식은 아래와 같음.

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} det(\tilde{A}_{ij})$$

또는 모든 j에 대해서 j열에 대한 여인수 전개를 하면 행렬식은 아래와 같음.

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} det(\tilde{A}_{ij})$$

n차 정사각행렬의 행렬식은 그 정의[35]에 의하면 첫번째 행에 대한 여인수 전개로 구하는 것임. 하지만 Theorem 4.4에 의해, n차 정사각행렬의 행렬식은 임의의 행에 대한 여인수 전개로 구할 수 있음. 또한 Theorem 4.8에 의하면, 행렬식은 임의의 열에 대한 여인수 전개로도 구할 수 있음.

2) 여인수(cofactor)

Definition 46. *스칼라 $(-1)^{i+j} A_{ij} det(\tilde{A}_{ij})$는 A의 i행 j열 성분에 대한 여인수(cofactor)라 함.*

여인수를 $c_{ij} = (-1)^{i+j} A_{ij} det(\tilde{A}_{ij})$로 표기했을 때, 행렬식을 아래와 같이 여인수들의 일차결합으로 나타낼 수 있음. 이를 i행에 대한 여인수 전개(cofactor expansion) 또는 라플라스 전개(Laplace expansion)라고 함.

$$det(A) = A_{i1}c_{i1} + A_{i2}c_{i2} + \cdots + A_{in}c_{in}$$

3) 소행렬

Definition 47. *A의 i행과 j열을 지워서 얻은 $(n-1) \times (n-1)$ 행렬을 A의 (i,j) 소행렬이라고 하고, $\tilde{A}_{ij}$로 표기함.*

[34] 이 정의에 대한 유도를 생각해 볼 수 있음. $det(A)$를 스칼라와 기본행렬의 곱으로 정리해 보자. 자세한 건 Theorem 4.4 참고.
[35] 프리드버그 p.234.

2. 상삼각행렬로 행렬식 구하기

1) 상삼각행렬로 행렬식 구하기

Theorem 상삼각행렬의 행렬식은 대각성분의 곱과 같음.

임의의 정사각행렬은 1형과 3형 기본행연산만으로 상삼각행렬로 바꿀 수 있으므로, 상삼각행렬로 전환하여 행렬식을 구할 수 있음.

3. 관련 정리

1) 임의의 행에 대한 여인수 전개

Theorem 4.4 정사각행렬의 행렬식은 임의의 행에 대해서 여인수 전개하여 구할 수 있음. 즉 $A \in M_{n\times n}(F)$와 임의의 정수 $i(1 \le i \le n)$에 대해서 아래가 성립함.

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} det(\tilde{A}_{ij})$$

증명은 프리드버그 p.238 참고.

3. 행렬식의 성질

1. 행렬식의 엄밀한 정의에 의한 성질

1) 기본 성질

행렬식은 $\delta(I) = 1$을 만족시키는 교대 n-선형함수임. 이에 따른 (매우 당연한) 성질을 정리하면 아래와 같음.

1. 행렬식은 선형이 아니지만 각 행에 대해서 선형성을 가짐.[36] (multi-linear)
2. 행렬의 두 행이나 열이 서로 같으면 그 행렬식은 0임.[37] (alternating)
3. $det(I) = 1$임.

2) 기본연산과 행렬식

Theorem 4.8, Theorem 4.10의 Corollary 3에 의하면, 기본연산이 행렬식에 미치는 영향은 아래와 같음.

1. $n \times n$ 행렬 A의 두 행 또는 두 열을 교환하여 얻은 행렬을 B라 하면 $det(B) = -det(A)$임. (1형)
2. $n \times n$ 행렬 A의 한 행 또는 열에 스칼라 k를 곱하여 얻은 행렬을 B라 하면 $det(B) = kdet(A)$임. (2형)
3. $n \times n$ 행렬 A의 한 행 또는 열에 다른 행에 스칼라 배를 더하여 얻은 행렬을 B라 하면 $det(B) = det(A)$임. (3형)

2. 행렬식의 성질

1) 행렬식과 가역성

Theorem 4.7의 Corollary에 의하면, 행렬이 가역이기 위한 필요충분조건은 그 행렬식이 0이 아닌 것임.

이때 Theorem 4.2에 의하면, 2×2 정사각행렬인 경우 그 역행렬을 간단히 구할 수 있음.

추가로 Theorem 4.3 Corollary에 의하면, 모든 원소가 0인 행이나 열이 존재할 경우, 그 행렬의 행렬식은 0임.

2) 행렬식과 행렬 곱

Theorem 4.7에 의하면, $det(AB) = det(A)det(B)$임. 즉, 행렬식은 행렬의 곱을 보존함.[38].

3) 전치행렬의 행렬식

Theorem 4.8에 의하면 $det(A^t) = det(A)$임.

이를 응용하면 행에 대한 개념들을 열에 대한 것으로까지 확장할 수 있음. 행렬식은 열에 대한 여인수 전개로도 구할 수 있고, 기본행연산 대신 기본열연산으로 행렬을 변형하여 행렬식을 구할 수도 있음. 이때, 기본열연산에 따른 행렬식의 변화는 기본행연산의 그것과 같음.

[36] Theorem 4.1, Theorem 4.3, Theorem 4.5은 행렬식이 multi-linear의 성질을 보이고 있기 때문에 이 필기에 정리하지 않음.

[37] Theorem 4.4 Corollary, Theorem 4.5, Theorem 4.6는 행렬식이 alternating의 성질을 보이고 있기 때문에 이 필기에 정리하지 않음.

[38] 즉, 곱하고 행렬식을 구하나 행렬식을 각각 구하고 곱하나 똑같음.

3. 행렬식의 기하학적 해석

1) 2차 정사각행렬

2차 정사각행렬과 그 행렬식은 평행사변형의 넓이로서 기하학적으로 해석해 볼 수 있음. 행렬의 각 행을 묶어 각각을 좌표계에서 원점을 시점으로 하는 화살표로 생각하면 평행사변형이 유일하게 결정됨. 이때, 행렬식의 절댓값이 평행사변형의 넓이라는 것.

프리드버그 p.226 참고.

2) n차 정사각행렬

행렬의 각 행을 묶어 좌표계에서 원점을 시점으로 하는 화살표로 생각하면 n차원 입체도형이 유일하게 결정됨. 이때, 행렬식의 절댓값이 해당 입체도형의 n차원 부피라는 것.

프리드버그 p.251 참고. 미적분에 대한 이해가 필요하다고 함.

4. 관련 정리

1) 행렬식과 행렬 곱

Theorem 4.7 임의의 $A, B \in M_{n\times n}(F)$에 대해서, $det(AB) = det(A)det(B)$임.

프리드버그 p.267 Theorem 4.11에 의하면, 임의의 교대 n-선형함수 $\delta : M_{n\times n}(F) \to F$에 대해 $\delta(AB) = \delta(A)\delta(B)$가 성립함.

Proof. A가 기본행렬인 경우, $det(AB) = det(A)det(B)$임이 성립함.

A의 랭크가 n보다 작은 경우, $rank(AB) \leq rank(A) < n$이므로 $det(AB) = det(A)det(B) = 0$으로 성립함.

A의 랭크가 n인 경우, A는 기본행렬의 곱으로 나타낼 수 있음. $A = E_m \cdots E_2E_1$이라고 하자. A가 기본행렬인 경우 $det(AB) = det(A)det(B)$이므로, $det(AB) = det(E_m \cdots E_2E_1B) = det(E_m)\cdots det(E_1)det(E_B) = det(E_m \cdots E_2E_1)det(B) = det(A)det(B)$로 성립함. □

Corollary 행렬 $A \in M_{n\times n}(F)$가 가역이기 위한 필요충분조건은 $det(A) \neq 0$임. 특히, A가 가역이면 $det(A^{-1}) = \frac{1}{det(A)}$임.

Proof. 1. 행렬 $A \in M_{n\times n}(F)$가 가역 $\to det(A) \neq 0$
$det(A)det(A^{-1}) = det(AA^{-1}) = det(I) = 1$이므로 $det(A) \neq 0$이고, $det(A) = \frac{1}{det(A^{-1})}$임.

2. $det(A) \neq 0 \to$ 행렬 $A \in M_{n\times n}(F)$가 가역[39]
행렬 A가 가역이 아니면 $det(A) = 0$임을 보이자. A가 가역이 아니면 어떤 행을 다른 행들의 일차결합으로 표현할 수 있음. A의 각 행을 $a_1, \cdots, a_n$, 일차결합으로 표현될 수 있는 행을 a_r이라 하자. a_r은 스칼라 $c_1, \cdots, c_n$에 대해 $a_r = c_1a_1 + \cdots + c_{r-1}a_{r-1} + c_{r+1}a_{r+1} + \cdots + c_na_n$으로 표현할 수 있음. 행렬 B를 A의 a_r을 제외한 각 행 $a_1, \cdots, a_n$에 각각 스칼라 $-c_1, \cdots, -c_n$를 곱해 r행에 더한 행렬이라고 하자. 즉, B는 A에 3형 기본연산을 반복해 얻은 행렬임. B의 r번째 행은 모든 원소가 0이므로, $det(B) = det(A) = 0$임. □

[39] 프리드버그 기준 Theorem 4.6 Corollary의 내용임. Theorem 4.10 Corollary 2에서 확인할 수도 있음.

2) 2차 정사각행렬의 행렬식과 가역성

Theorem 4.2 행렬 $A \in M_{2\times 2}(F)$에 대해서 A의 행렬식이 0이 아니기 위한 필요충분조건은 A가 가역행렬인 것임. 특히, A가 가역행렬이면 역행렬은 아래와 같음.

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

Proof. 1. $det(A) \neq 0 \to A$가 가역
$det(A) \neq 0$인 경우, 위의 역행렬 식을 A에 직접 곱해서 가역임을 확인할 수 있음.

2. A가 가역 $\to det(A) \neq 0$
행렬 A가 아래와 같다고 하자.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

A가 가역이므로 A_{11}, A_{21}이 모두 0일 수 없음. $A_{11} \neq 0$일 때, 3형 기본연산으로 A를 아래와 같이 바꿀 수 있음.

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} - \frac{A_{12}A_{21}}{A_{11}} \end{pmatrix}$$

이때, A가 가역이므로 $A_{22} - \frac{A_{12}A_{21}}{A_{11}} \neq 0$이어야 함. 즉, 정리하면 $det(A) = A_{11}A_{22} - A_{12}A_{21} \neq 0$임. □

3) 어떤 행의 성분이 모두 0인 경우의 행렬식

Theorem 4.3 Corollary 행렬 $A \in M_{n\times n}(F)$의 어느 행의 모든 성분이 0이면 $det(A) = 0$임.

당연히 열에 대해서도 적용 가능함.

Proof. $n \times n$ 행렬의 어느 행의 모든 성분이 0이면 랭크가 n 미만이므로 가역이 아님. 즉, $det(A) = 0$임. □

4) 전치행렬의 행렬식

Theorem 4.8 임의의 $A \in M_{n\times n}(F)$에 대해서 $det(A^t) = det(A)$임.

Proof. A가 가역이 아니면 $rank(A) = rank(A^t) \leq n$이므로 $det(A) = det(A^t) = 0$임.

A가 가역이면 기본연산으로 표현할 수 있으므로, $A = E_1E_2\cdots E_m$, $A^t = E_m^t \cdots E_2^tE_1^t$임. $det(E) = det(E^t)$이므로[40], $det(A^t)$를 Theorem 4.7을 사용하여 정리하면 $det(A^t) = det(A)$임. □

[40] 프리드버그 4.2절 연습문제 29 참고.

Part V

대각화

행렬을 대각화하면 행렬 곱 등을 굉장히 간단하게 계산할 수 있음.

1. 교윳값과 고유벡터

1. 대각화(diagonalization)

1) 대각화(diagonalization))

대각화는 대각행렬을 만드는 것 또는 그 과정을 의미함.

한 선형변환에 대해서 다수의 행렬표현과 기저가 있을 수 있음. 이 중 최고의 행렬표현은 대각행렬이고, 최고의 기저는 그런 대각행렬을 만드는 기저임.

어떤 선형연산자가 언제 대각화가능한지, 그리고 대각행렬이 되도록 하는 기저는 무엇인지를 알아내야 함.

2) 대각화가능(diagonalizable)

Definition 48. *유한차원 벡터공간 V에서 정의된 선형연산자 T에 대해서,*

$[T]_\beta$가 대각행렬이 되도록 하는 V의 순서기저 β가 존재할 때, 선형연산자 T는 대각화가능(diagonalizable)하다고 함.

L_A가 대각화가능할 때, 정사각행렬 A는 대각화가능(diagonalizable)하다고 함.

Theorem 5.1과 그 Corollary에 의하면, 선형변환과 행렬이 대각화가능하기 위한 필요충분조건은 고유벡터로 이루어진 순서기저가 존재하는 것임. 또한, 행렬이 대각화가능하기 위한 필요충분조건은 해당 행렬과 닮음인 대각행렬이 존재하는 것임.

2. 교윳값(eigenvalue)과 고유벡터(eigenvector)

1) 정의

Definition 49. *1. 선형변환*
벡터공간 V의 선형연산자 T에 대해서,

영벡터가 아닌 벡터 $v \in V$와 어떤 스칼라 λ가 존재하여 $T(v) = \lambda v$를 만족할 때, 벡터 v를 T의 고유벡터(eigenvector)라 함.

스칼라 λ를 고유벡터 v에 대응하는 고윳값(eigenvalue)이라 함.

2. 행렬
$M_{n\times n}(F)$에 속하는 행렬 A에 대해서,

L_A의 고유벡터, 즉 $Av = \lambda v$인 스칼라 λ가 존재하게 만드는 영벡터가 아닌 벡터 $v \in F^n$을 A의 고유벡터(eigenvector)라 함.

스칼라 λ를 고유벡터 v에 대응하는 행렬 A의 고윳값(eigenvalue)이라 함.

고윳값과 고유벡터는 어떤 선형변환을 대각행렬로 표현하기 위한 개념임.

1. $[T]_\beta$가 대각행렬 $\rightarrow T(v) = \lambda v$
$D = [T]_\beta$가 대각행렬이면, 각 $v_j \in \beta$에 대해서 $T(v_j) = \sum_{i=1}^{n} D_{ij}v_i = D_{jj}v_j = \lambda_j v_j$ (단, $\lambda_j = D_{jj}$)가 성립함.

2. $T(v) = \lambda v \rightarrow [T]_\beta$가 대각행렬
$v \in \beta$이면 적절한 스칼라 $\lambda_1, \cdots, \lambda_n$에 대해서 $[T]_\beta$를 아래와 같이 나타낼 수 있음.

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

즉, 선형변환 행렬표현이 대각행렬이 되는 기저의 각 원소 v에 대해, $T(v) = \lambda v$가 성립함. 역으로 $T(v) = \lambda v$인 v들의 집합에서 기저를 찾아 선형변환을 행렬로 표현하면 대각행렬이 됨.

고유벡터 대신 특성벡터(characteristic vector, proper vector), 고윳값 대신 고유치 또는 특성값(characteristic value, proper value)라고 하기도 함.

2) 대각행렬의 성질

$n \times n$ 행렬 A의 고유벡터로 이루어진 F^n의 순서기저를 β라 하고, 표준 순서기저를 α라 하자. 이때 $Q = [I]_\beta^\alpha$에 대해서 $A' = Q^{-1}AQ$인 A'은 대각행렬임.

Theorem A'의 대각성분은 A의 고윳값이고, 대응하는 고유벡터는 Q의 각 열임.

3. 특성다항식(characteristic polynomial)

1) 정의

> **Definition 50.** *행렬 $A \in M_{n\times n}(F)$에 대해서 다항식 $f(t) = det(A - tI_n)$을 A의 특성다항식(characteristic polynomial)이라 함.*

Theorem 5.2에 의하면, t가 행렬 A의 고윳값이기 위한 필요충분조건은 $det(A - \lambda I_n) = 0$인 것임. 이를 다항식으로 나타낸 것이 특성다항식임. 즉, 특성다항식을 t에 대해 해를 구하면 그게 고윳값임.

Theorem 5.3은 특성다항식의 성질에 대한 것임.

2) 선형변환의 행렬식과 특성다항식[41]

> **Definition 51.** *유한차원 벡터공간 V에서 정의된 선형연산자 T를 생각하자. V의 임의의 순서기저 β에 대해서 $A = [T]_\beta$의 행렬식을 T의 행렬식(determinant)이라고 하고 $det(T)$라고 표기함. 또한 행렬 A의 특성다항식 $f(t) = det(A - tI_n)$을 T의 특성다항식(characteristic polynomial)이라고 함.*

선형변환의 행렬식은 행렬의 행렬식이 가지는 성질들을 동일하게 가짐.[42]

[41]지금까지는 행렬의 행렬식과 특성다항식에 대해서 이야기했는데, 이제 선형변환(어차피 같음.)에 대해서도 이야기할 수 있음.
[42]가역성 확인, 역행렬의 행렬식, 행렬식과 행렬 곱 등. 프리드버그 p.284 8번 문제 참고.

3) 선형변환과 행렬의 관계[43]

Theorem 닮은 행렬끼리는 서로 같은 행렬식과 특성다항식을 가짐.

즉, 각 선형변환에 대해 행렬식과 특성다항식은 유일하므로 선형변환의 행렬식과 특성다항식에 대해 정의할 수 있음. 이때, 유한차원 벡터공간 V에서 정의된 선형연산자 T와 V의 순서기저 β에 대해서 λ가 T의 고윳값이기 위한 필요충분조건은 λ가 $[T]_\beta$의 고윳값인 것임.

정리하면, 선형변환과 그 모든 행렬표현은 행렬식, 특성다항식, 고윳값, 고유벡터[44]를 공유함.

4. 대각화 총정리

1) 대각화

고윳값/고유벡터/대각행렬를 구하는 과정을 대각화라고 함. 행렬의 고윳값/고유벡터는 아래와 같은 방법으로 구할 수 있음.

1. 특성다항식으로 고윳값을 구함.
$f(t) = det(A - tI_n)$의 근을 찾으면 됨. (Theorem 5.2)

2. 고윳값으로 고유벡터를 구함.
$(A - \lambda I_n)v = 0$인 영이 아닌 v를 찾으면 됨. 이때 이 방정식은 $(A - \lambda I_n)$가 계수행렬인 동차 연립일차방정식이므로, 그것의 해집합을 찾으면 됨.

3. 고유벡터 집합이 순서기저인지 확인함.
순서기저인 경우 이 집합은 해당 행렬/선형변환을 대각화하는 기저임.

4. 닮음인 대각행렬 구하기
순서기저인 고유벡터의 집합으로 기저변환행렬을 구성하여 $Q^{-1}AQ$를 계산하거나, 직접 선형변환에 대해 (L_A 또는 T) 해당 순서기저로 행렬을 구함.

이때의 순서기저인 고유벡터의 집합은 고유기저(Eigen Basis)라고 부름.

대각화가능한 $n \times n$ 행렬 A의 고유벡터를 열로 갖는 행렬을 Q, A를 대각화한 행렬을 B라 하자. A는 $A = QBQ^{-1}$로 분해가 가능한데 이를 고윳값 분해(eigen decomposition)라 함.

선형변환과 그 행렬표현은 고윳값을 공유하고, 선형변환 고유벡터의 좌표벡터는 행렬의 고유벡터이므로[45], 선형변환을 임의의 행렬표현으로 나타내 선형변환에 대해서도 구할 수 있음.

[43]증명은 프리드버그 p.276, p.278 참고.
[44]단, 행렬의 고유벡터는 선형변환 고유벡터와 완전히 같은 것은 아니고, 선형변환 고유벡터의 좌표벡터임.
[45]증명은 프리드버그 p.278 참고.

5. 고윳값의 기하학적 의미[46]

1) 고윳값의 기하학적 의미

고윳값의 종류에 따라 선형변환이 벡터(고유벡터, 고유벡터의 span 위에 벡터)에 다르게 작용함.

T의 고유벡터는 v, 대응하는 고윳값을 λ라 하자. v에 의해 생성된 V의 일차원 부분공간 $W = span(\{v\})$를 0과 v를 지나는 직선으로 생각할 수 있음.

경우 1. $\lambda > 1$일 때, T는 W의 벡터를 λ에 비례하여 0에서 밀어냄.
경우 2. $\lambda = 1$일 때, T는 W에서 항등연산자임.
경우 3. $0 < \lambda < 1$일 때, T는 W의 벡터를 λ에 비례하여 0방향으로 당김.
경우 4. $\lambda = 0$일 때, T는 W에서 영 연산자임.
경우 5. $\lambda < 0$일 때, T는 W의 방향을 바꿈. 즉, 0의 반대 방향으로 옮김.

좌표평면에서 고유벡터가 생성하는 일차원(직선) 부분공간 위에 있는 벡터에 대해서, 위에 작성한 고윳값의 종류에 따라 선형변환이 다르게 작용함.

영이 아닌 임의의 벡터 v에 대해서 $0, v, T(v)$가 한 직선 위에 있지 않다면, T가 고윳값과 고유벡터를 가지지 않는다는 것을 알 수 있음.

6. 관련 정리

1) 대각화가능한 경우(당연한 말)

Theorem 5.1 유한차원 벡터공간 V의 선형연산자 T가 대각화가능하기 위한 필요충분조건은 T의 고유벡터로 이루어진 V의 순서기저 β가 존재하는 것임. 또한 T가 대각화가능하고 $\beta = \{v_1, \cdots, v_n\}$이 T의 고유벡터로 이루어진 순서기저이면 $D = [T]_\beta$는 대각행렬임. 이때, D_{jj}는 v_j에 해당하는 고윳값임.

고유벡터와 고윳값의 정의가 이 정리의 증명임.

Corollary 행렬 $A \in M_{n\times n}(F)$가 대각화가능하기 위한 필요충분조건은 A의 고유벡터로 이루어진 F^n의 순서기저가 존재하는 것임. 또한 $\{v_1, \cdots, v_n\}$은 A의 고유벡터로 이루어진 F^n의 순서기저이고 j열이 벡터 v_j인 $n \times n$ 행렬 Q에 대해서 $D = Q^{-1}AQ$는 D_{jj}가 v_j에 대응하는 A의 고윳값인 대각행렬임. 즉, 행렬이 대각화가능하기 위한 필요충분조건은 해당 행렬과 닮음인 대각행렬이 존재하는 것임.

2) 행렬의 고윳값 구하기

Theorem 5.2 행렬 $A \in M_{n\times n}(F)$에 대해서 스칼라 λ가 A의 고윳값이기 위한 필요충분조건은 $det(A - \lambda I_n) = 0$임.

Proof. 스칼라 λ가 A의 고윳값이기 위한 필요충분조건은 $Av = \lambda v$를 만족하는 영이 아닌 벡터 $v \in F^n$이 존재하는 것임. 즉, $(A - \lambda I_n)(v) = 0$임. 이 식이 성립하기 위한 필요충분조건은 $(A - \lambda I_n)$이 가역이 아닌 것임(nullity가 0보다 커야 함.). 즉, $det(A - \lambda I_n) = 0$인 것. □

3) 특성다항식의 특성

Theorem 5.3 행렬 $A \in M_{n\times n}(F)$에 대해서 아래가 성립함.

1. A의 특성다항식은 n차 다항식이고, 최고차항의 계수는 $(-1)^n$임.
2. A에는 최대 n개의 서로 다른 고윳값이 있음.

4) 행렬의 고유벡터 구하기

Theorem 5.4 행렬 $A \in M_{n\times n}(F)$와 고윳값 λ에 대해서 벡터 $v \in F^n$이 λ에 대응하는 A의 고유벡터이기 위한 필요충분조건은 $v \neq 0$이고 $(A - \lambda I)v = 0$인 것임.

고유벡터의 정의가 곧 이 정리의 증명임.

[46] 프리드버그 p.281 참고.

2. 대각화 가능성

1. 다항식의 인수분해

1) 정의

Definition 52. *아래와 같이 표현되는 다항식 $f(t) \in P(F)$를 F 위에서 완전히 인수분해된다(split over F)고 함. 이때 스칼라 $a_1, a_2, \cdots, a_n \in F$ 중 같은 값이 있을 수 있음.*

$$f(t) = c(t - a_1)(t - a_2)\cdots(t - a_n)$$

2) 체에 따른 인수분해

$f(t)$가 F 위 선형연산자 또는 특성다항식일 때, $f(t)$가 완전히 인수분해된다는 것은 체 F 위에서 완전히 인수분해된다는 것임.

다항식의 근이 체 위에 있어야 하므로 체에 따라서 인수분해가 가능한 영역이 달라짐.

예를 들어, R 위에서는 실수근만 존재할 수 있어 t^2+1를 더이상 인수분해할 수 없지만, C 위에서는 허수근도 존재할 수 있어 $t^2 + 1 = (t + i)(t - i)$로 인수분해할 수 있음.

2. 대수적/기하적 중복도

대수적/기하적 중복도는 완전히 인수분해되는 특성다항식을 가지는 연산자가 대각화가능한지를 살필 수 있게 함.

1) 대수적 중복도

Definition 53. *특성다항식이 $f(t)$인 선형연산자(또는 행렬)의 고윳값 λ에 대해서, $(t - \lambda)^k$가 $f(t)$의 인수가 되도록 하는 가장 큰 자연수 k를 λ의 중복도(multiplicity) 또는 대수적 중복도(algebraic multiplicity)라고 함.*

즉, 특성다항식에서 λ에 해당되는 차수.

Theorem 5.1에 의하면 대각화로 만든 대각행렬의 대각성분은 고윳값이고, T의 고윳값은 대각성분에 각 대수적 중복도만큼 나타남. 또한, 대각행렬을 만드는 순서기저에는 동일한 고윳값에 대응되는 고유벡터가 해당 고윳값의 대수적 중복도만큼 나타남.

2) 고유공간(eigenspace)

Definition 54. *벡터공간 V의 선형연산자 T의 고윳값 λ에 대해서, 아래의 집합 E_λ를 고윳값 λ에 대응하는 T의 고유공간(eigenspace)이라고 함.*

$$E_\lambda = \{x \in V | T(x) = \lambda x\} = N(T - \lambda I_V)$$

이와 비슷하게 λ에 대응하는 L_A의 고유공간을 λ에 대응하는 정사각행렬 A의 고유공간이라고 함.

이때, $dim(E_\lambda)$를 λ의 기하적 중복도라고 함.

즉, λ의 고유공간은 대응하는 고유벡터와 영벡터로 이루어진 V의 부분공간임. 또한 기하적 중복도는 고유벡터의 차원이므로 일차독립인 고유벡터들의 개수임.

$(A - \lambda I)(v) = 0$을 만족시키는 v가 하나뿐인 경우, 즉 $nullity(A - \lambda I) = 0$인 경우는 존재하지 않음. λ는 특성다항식 $f(t) = det(A - tI) = 0$를 만족시키므로 $A - \lambda I$는 가역이 아님.

3. 대각화가능 판정법

1) 대각화가능하기 위한 필요충분조건

Definition 55. *n차원 벡터공간 V의 선형연산자 T가 대각화가능하기 위한 필요충분조건은 아래의 두 조건이 모두 성립하는 것임.*

1. *T의 특성다항식이 체 위에서 완전히 인수분해됨.*
2. *T의 각 고윳값에 대해서 대수적 중복도와 기하적 중복도가 같음. 즉, λ의 중복도가 $nullity(T - \lambda I) = n - rank(T - \lambda I)$임.*

이는 Theorem 5.8에 의해 성립함.

행렬 A의 대각화 가능성과 L_A의 대각화 가능성은 동치이므로, 행렬에 대해서도 위 판정법을 적용할 수 있음.

2) 대각화가능 판정법

방법1. 특성다항식 사용.

1. 특성다항식을 인수분해.
2. 중근을 가지는 고윳값에 대해 대수적 중복도와 기하적 중복도가 같은지 확인.

Theorem 5.5에 의해 대수적 중복도가 1인 고윳값에 대해서는 확인하지 않아도 됨.
기하적 중복도는 $nullity(A - \lambda I) = dim(V) - rank(A - \lambda I)$이므로 직접 기저를 구하지 않아도 알 수 있음.
행렬식과 특성다항식은 기본적으로 행렬에 대해 정의되어 있으므로, 선형연산자를 행렬로 바꾸어 계산하는 것이 편함.

방법2. T의 고유벡터로 이루어진 V의 순서기저가 존재하는지 확인.

방법3. V가 T의 고유공간의 직합인지 확인.[47]

[47] 원리는 방법2와 동일함.

4. 관련 정리

1) 고유벡터 집합들의 독립성

Theorem 5.5 벡터공간의 선형연산자 T와 그 고윳값 $\lambda_1, \lambda_2, \cdots, \lambda_k$를 생각하자. 각 $i = 1, \cdots, k$에 대해서 λ_i에 대응하는 고유벡터의 집합을 S_i라 하자. 각 S_i가 일차독립이면 $S_1 \cup S_2 \cup \cdots \cup S_k$도 일차독립임.

즉, 각 고윳값에 대응하는 고유벡터들의 집합이 일차독립이면, 서로 합집합해도 일차독립임. 대수적 중복도가 1인 고윳값에 대한 기하적 중복도는 항상 1임.

Proof. 고윳값의 개수에 대해서 수학적 귀납법으로 증명함. □

Corollary n차원 벡터공간의 선형연산자 T가 서로 다른 n개의 고윳값을 가지면 T는 대각화가능함.

2) 대각화가능성과 특성다항식의 인수분해

Theorem 5.6 F-벡터공간 V의 대각화가능한 선형연산자의 특성다항식은 F 위에서 완전히 인수분해됨.

역은 성립하지 않음.

Proof. 선형변환의 행렬표현들에 대해서 특성다항식은 모두 같으므로, 대각화하여 만든 대각행렬에 대해 특성다항식을 작성하면 완전히 인수분해된다는 것을 알 수 있음. □

3) 대수적 중복도와 기하적 중복도 사이의 관계

Theorem 5.7 유한차원 벡터공간 V의 선형연산자 T와 중복도가 m인 T의 고윳값 λ에 대해서, $1 \leq dim(E_\lambda) \leq m$ 임.

즉, 임의의 고윳값에 대해 기하적 중복도는 대수적 중복도보다 작음.

Proof. E_λ의 순서기저를 $\{v_1, \cdots, v_p\}$라 하고 이를 확장하여 만든 V의 순서기저 $\beta = \{v_1, \cdots, v_p, v_{p+1}, \cdots, v_n\}$에 대해 생각하자. $[T]_\beta$를 아래와 같이 나타낼 수 있음.

$$[T]_\beta = \begin{pmatrix} \lambda I_p & B \\ O & C \end{pmatrix}$$

T의 특성다항식은 아래와 같이 나타낼 수 있음.

$$f(t) = det\begin{pmatrix} (\lambda - t)I_p & B \\ O & C - tI_{n-p} \end{pmatrix} = (\lambda - t)^p det(C - tI_{n-p})$$

λ의 대수적 중복도는 p(기하적 중복도) 이상임. □

4) 대각화가능하기 위한 필요충분조건

Theorem 5.8 유한차원 벡터공간 V의 선형연산자 T에 대해서 T의 특성다항식이 완전히 인수분해되고 $\lambda_1, \cdots, \lambda_k$가 T의 서로다른 고윳값일 때 아래가 성립함.

1. T가 대각화가능하기 위한 필요충분조건은 모든 i에 대해서 λ_i의 대수적 중복도가 $dim(E_\lambda)$(기하적 중복도)와 같은 것임.
2. T가 대각화가능하고 각각의 i에 대해서 β_i가 E_{λ_i}의 순서기저일 때, $\beta = \beta_1 \cup \cdots \cup \beta_k$는 T의 고유벡터로 이루어진 V의 순서기저임.

즉, 모든 고윳값에 대해서 고윳값이 최대 개수만큼 나와야 함. 최대 개수만큼 나온 고윳값들을 모으면 T를 대각행렬로 만드는 V의 순서기저가 됨.

3. 직합

1. 직합(direct sum)

벡터공간을 부분공간으로 분해하는 방법. 특히, 벡터공간을 선형연산자의 고유공간으로 분해할 수 있음.

1) 부분공간의 합(sum)

Definition 56. *벡터공간 V의 부분공간 $W_1, W_2, \cdots, W_k$에 대해서 아래의 집합을 부분공간의 합(sum)이라 함.*

$$\{v_1, v_2, \cdots, v_k \,|\, 1 \leq i \leq k,\ v_i \in W_i\}$$

이를 $W_1 + \cdots + W_k$ 또는 $\sum_{i=1}^{k} W_i$라 표기함.

즉, 각 집합의 원소들을 더한 것을 원소로 가지는 집합을 합이라고 하는 것.

부분공간의 합은 부분공간임.

2) 직합(direct sum)

Definition 57. *벡터공간 V의 부분공간 $W, W_1, \cdots, W_k$를 생각하자. 모든 $i = 1, 2, \cdots, k$에 대해서 $W_i \subseteq W$이고 아래를 만족하는 W를 부분공간 $W_1, W_2, \cdots, W_k$의 직합(direct sum)이라 하며, $W = W_1 \oplus W_2 \oplus \cdots \oplus W_k$라 표기함.*

$$W = \sum_{i=1}^{k} W_i \text{이고 각 } 1 \leq j \leq k \text{에 대해서 } W_j \cap \sum_{i \neq j} W_i = \{0\}$$

즉, 부분공간들의 합 중에서 합쳐진 모든 부분공간들의 각 교집합이 $\{0\}$인 것.

3) 직합과 대각화 가능성

Theorem 5.10에 의하면 유한차원 벡터공간 V의 선형연산자 T가 대각화가능하기 위한 필요충분조건은 V가 T의 고유공간의 직합인 것임.

2. 관련 정리

1) 직합과 동치인 조건들

Theorem 5.9 유한차원 벡터공간 V와 그 부분공간 $W_1, \cdots, W_k$에 대해서 아래 조건들은 서로 동치임.

1. $V = W_1 \oplus \cdots \oplus W_k$
2. $V = \sum_{i=1}^k W_i$이고 $1 \le i \le k$에 대해서 $v_i \in W_i$인 임의의 벡터 $v_1, \cdots, v_k$가 있을 때, $v_1 + \cdots + v_k = 0$이면 모든 i에 대해서 $v_i = 0$임.
3. 모든 $v \in V$마다 $v = v_1 + \cdots + v_k$ 꼴로 표현하는 방법이 유일함.
4. W_i의 순서기저 γ_i에 대해서 $\gamma_1 \cup \cdots \cup \gamma_k$는 V의 순서기저임.
5. $\gamma_1 \cup \cdots \cup \gamma_k$가 V의 순서기저가 되도록 하는 γ_i가 존재함.

Proof. 1. 3에 의하면 $V = \sum_{i=1}^k W_i$임. 어떤 벡터 $v \in V$에 대해서 $v \in W_j \cap \sum_{i \ne j} W_i$가 성립한다고 가정하자. $v \in span(\gamma_i)$, $v \in span(\cup_{i \ne j} \gamma_i)$임. 즉, v를 V의 순서기저인 $\gamma_1 \cup \cdots \cup \gamma_k$의 일차결합으로 표현하는 방법이 2가지인 것이므로 모순임. 따라서 $W_j \cap \sum_{i \ne j} W_i = \{0\}$임.

4. 3에 의하면 $V = \sum_{i=1}^k W_i$이므로, $\gamma_1 \cup \cdots \cup \gamma_k$는 V를 생성함. $\gamma_1 \cup \cdots \cup \gamma_k$는 일차독립이므로 V의 기저임. □

2) 직합과 대각화 가능성

Theorem 5.10 유한차원 벡터공간 V의 선형연산자 T가 대각화가능하기 위한 필요충분조건은 V가 T의 고유공간의 직합인 것임.

Proof. 1. 대각화가능 $\rightarrow$ V가 고유공간의 직합
i에 대해서 고유공간 E_{λ_i}의 순서기저를 γ_i라 하면, $\gamma_1 \cup \cdots \cup \gamma_k$가 V의 순서기저임. Theorem 5.9에 의해 V는 고유공간의 직합임.

2. V가 고유공간의 직합 $\rightarrow$ 대각화가능
i에 대해서 고유공간 E_{λ_i}의 순서기저를 γ_i라 하면, $\gamma_1 \cup \cdots \cup \gamma_k$가 V의 순서기저임. 이 기저는 T의 고유벡터로 이루어진 순서기저이므로 T는 대각화가능함. □

4. 행렬의 극한과 마르코프 연쇄

1. 행렬의 극한

1) 복소수열의 극한

Definition 58. *복소수로 이루어진 수열의 극한은 실수부와 허수부 각각의 실수열의 극한으로 정의함. 즉, r_m과 s_m이 각각 수렴하는 실수열, i는 $i^2 = -1$인 허수라고 할 때, $z_m = r_m + is_m$이면 아래가 성립함.*

$$\lim_{m\to\infty} z_m = \lim_{m\to\infty} r_m + i \lim_{m\to\infty} s_m$$

실수열의 극한은 안다고 가정함.

2) 행렬열의 극한

Definition 59. *복소수 성분을 가지는 $n \times p$ 행렬 $L, A_1, A_2 \cdots$를 생각하자. 모든 $1 \le i \le n$, $1 \le j \le p$에 대해서 $\lim_{m\to\infty} (A_m)_{ij} = L_{ij}$일 때, 행렬열 $A_1, A_2, \cdots$는 $n \times p$ 행렬 L로 수렴(converge)한다고 함. 이때 행렬 L을 이 행렬열의 극한(limit)이라 함. 행렬열의 극한을 L로 표기하면 간단히 $\lim_{m\to\infty} A_m = L$이라 나타내기도 함.*

즉, 각각의 성분에 대해 극한을 취한 것.

Theorem 5.11에 의하면 실수열의 극한에서처럼 상수(행렬)를 극한 밖으로 뺄 수 있음.

3) 행렬열의 극한 쉽게 계산하기

행렬 A가 대각화가능한 경우, $Q^{-1}AQ = D$가 대각행렬이 되도록 하는 가역행렬 Q가 존재함. A^m의 극한을 계산하는 대신 아래와 같이 D^m의 극한을 이용하여 계산하는 것.

$$\lim_{m\to\infty} A^m = \lim_{m\to\infty} (QDQ^{-1})^m = QLQ^{-1}$$

2. 극한의 존재성

복소수/행렬의 거듭제곱 또한 복소수열/행렬열의 하나로 생각할 수 있음.

1) 복소수 극한의 존재성

Theorem 집합 $S = \{\lambda \in C \,|\, |\lambda| < 1 \, or \, \lambda = 1\}$을 생각하자. 복소수 x에 대해서 $\lim_{m\to\infty} x^m$이 존재하기 위한 필요충분조건은 $x \in S$인 것임.

집합 S를 복소평면에서 생각해 보면 복소수 1과 단위원의 내부로 이루어져 있음.

2) 행렬 극한의 존재성

Theorem 5.12에 의하면 행렬 $\lim_{m\to\infty} A^m$이 존재하려면 A의 모든 고윳값이 S의 원소여야 하고, 고윳값 중 1에 해당하는 기하적 중복도와 대수적 중복도가 같아야(고유벡터가 꽉 차있어야) 함(또는 A가 대각화가능해야 함.).

2. 마르코프 연쇄(Markov process)

1) 추이행렬(transition matrix)

Definition 60. *음이 아닌 성분만을 가지고, 각 열의 합이 1인 행렬을 추이행렬(transition matrix) 또는 확률행렬(stochastic matrix)이라 함.*

임의의 $n \times n$ 추이행렬 M에 대해서 각 행과 열은 n가지 상태(state)에 대응함. M_{ij} 성분은 한 단계(stage)에 의해 상태 j에서 상태 i로 이동할 확률을 나타냄.

또한 $(M^m)_{ij}$ 성분은 m 단계를 거쳐 상태 j에서 상태 i로 이동할 확률임.

2) 확률벡터(probability vector)

Definition 61. *음이 아닌 성분만을 가지고, 각 성분의 합이 1인 열벡터를 확률벡터(probability vector) 라고 함.*

추이행렬의 각 열은 확률벡터임.

각 성분을 확률, 비율, 백분율 등으로 생각할 수 있음.

3) 추이행렬과 확률벡터의 곱

3.

1)

4. 관련 정리

1) 행렬열 극한에서의 상수(행렬)

Theorem 5.11 복소수 성분을 가지는 $n \times p$ 행렬열 $A_1, A_2, \cdots$이 행렬 L로 수렴한다고 하자. 임의의 $P \in M_{r\times n}(C)$, $Q \in M_{p\times s}(C)$에 대해서 아래가 성립함.

$$\lim_{m\to\infty} PA_m = PL, \quad \lim_{m\to\infty} A_m Q = LQ$$

Corollary $\lim_{m\to\infty} A^m = L$인 행렬 $A \in M_{n\times n}(C)$와 임의의 가역행렬 $Q \in M_{n\times n}(C)$에 대해서 아래가 성립함.

$$\lim_{m\to\infty} (QAQ^{-1})^m = QLQ^{-1}$$

2) 행렬열 극한의 존재성

Theorem 5.12 복소수 성분을 가지는 정사각행렬 A를 생각하자. $\lim_{m\to\infty} A^m$이 존재하기 위한 필요충분조건은 아래의 두 조건을 만족하는 것임.

1. A의 모든 고윳값은 S의 원소임. (S는 복소수 극한의 존재성에서의 집합임.)
2. 1이 A의 고윳값이면 1에 대응하는 기하적 중복도는 고윳값 1의 대수적 중복도와 같음.

2번 조건은 'A는 대각화가능함.'으로 바꾸어 쓸 수 있음.

Proof. 1. $\lim_{m\to\infty} A^m v = \lim_{m\to\infty} (Av)^m = \lim_{m\to\infty} (\lambda v)^m = \lim_{m\to\infty} \lambda^m v$이므로 λ가 S의 원소여야 $\lim_{m\to\infty} A^m$이 존재함.

2. 조르당 표준형을 알아야 설명할 수 있음. □

Part VI

내적공간

내적공간이라는 개념을 통해 벡터공간에 길이와 거리의 개념을 도입함.

벡터공간을 학습한 순서와 유사하게 내적공간(벡터공간), 정규직교기저(기저), 연산자(선형변환)의 순서로 내용이 구성됨.

1. 내적과 노름

1. 내적(inner product)

1) 정의

Definition 62. *F-벡터공간 V를 생각하자. V에 정의된 내적(inner product) $\langle x, y\rangle$는 V의 임의의 벡터 x와 y의 순서쌍을 F에 속한 스칼라에 대응시키는 함수로, 아래의 조건을 만족함.*

임의의 $x, y, z \in V$에 대해서

1. $\langle cx + z, y\rangle = c\langle x, y\rangle + \langle z, y\rangle$ *(left-linearity)*
2. $\overline{\langle x, y\rangle} = \langle y, x\rangle$ *(뒤집기)*
3. $\langle x, x\rangle \geq 0$*이고,* $x = 0$*은* $\langle x, x\rangle = 0$*이기 위한 필요충분조건임. (positive-definite)*

이를 응용하면 아래 또한 성립함을 알 수 있음.

1. $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$, $\langle x, cy\rangle = \bar{c}\langle x, y\rangle$
2. $\langle x, 0\rangle = \langle 0, x\rangle = 0$
3. *모든* $x \in V$*에 대해서* $\langle x, y\rangle = \langle x, z\rangle$*이면* $y = z$

아래쪽 1번에서 내적은 두 번째 성분에 대해서 켤레를 적용한 선형성을 띄는데, 이를 켤레선형(conjugate linear)이라 함.

위쪽 1번 조건에 의하면 내적은 왼쪽 성분에 대해 선형임.

2) 켤레 전치행렬(conjugate transpose)

Definition 63. *$A \in M_{n\times n}(F)$를 생각하자. 모든 i, j에 대해서 $(A^*)_{ij} = \overline{A}_{ji}$인 $n \times m$ 행렬 A^*를 A의 켤레 전치행렬(conjugate transpose) 또는 수반행렬(adjoint)이라 함.*

A의 모든 성분이 실수이면 A^*는 단순 전치행렬임.

3) 여러 가지 내적

1. 표준 내적(standard inner product)

> **Definition 64.** *F^n의 두 벡터 $x = (a_1, \cdots, a_n)$, $y = (b_1, \cdots, b_n)$에 대해서 $\langle x, y\rangle = \sum_{i=1}^{n} a_i \bar{b}_i$인 내적을 F^n의 표준 내적(standard inner product)이라 함.*
>
> *$F = R$인 경우 x, y 원소들의 단순 곱인데, 이 경우의 내적을 dot product라고 함. 이때 $\langle x, y\rangle$ 대신 $x \cdot y$로 표기하기도 함.*

F^n을 표준 내적이 주어진 내적공간으로 지칭하기도 함.

어떤 내적을 사용하더라도 내적공간들끼리는 위상이 같기 때문에 유사한 결과를 가짐. 그래서 대부분의 경우에는 내적들 중에서도 간단한 표준 내적(복소수체) 또는 dot product(실수체)를 사용함. 별 말이 없으면 표준 내적(복소수체) 또는 dot product(실수체)를 사용하는 것임.

2. 프로베니우스 내적(Frobenius inner product)

> **Definition 65.** *벡터공간 $V = M_{n\times n}(F)$와 두 행렬 $A, B \in V$에 대해서 $\langle A, B\rangle = tr(B^*A)$라 할 때, 이 내적을 프로베니우스 내적(Frobenius inner product)이라 함.*

$M_{n\times n}(F)$를 프로베니우스 내적이 주어진 내적공간으로 지칭하기도 함.

2. 내적공간(inner product space)

1) 정의

> **Definition 66.** *내적이 주어진 F-벡터공간 V를 내적공간(inner product space)이라 함.*
>
> *$F = C$이면 V를 복소내적공간(complex inner product space)이라 하고, $F = R$이면 V를 실내적공간(real inner product space)이라 함.*

이번 장에서 등장하는 모든 벡터공간은 R-벡터공간(실수체) 또는 C-벡터공간(복소수체)임.

같은 벡터공간에 서로 다른 내적을 주면 두 공간은 서로 다른 내적공간임.

내적공간의 부분공간은 동일한 내적이 정의된 내적공간임.

이제부터 벡터공간은 별 말이 없으면 내적공간으로 가정함.

3. 노름(norm)

1) 정의

> **Definition 67.** *내적공간 V와 벡터 $x \in V$에 대해서 x의 노름(norm) 또는 길이(length)를 $||x|| = \sqrt{\langle x, x\rangle}$로 정의함.*

하나의 벡터에 대한 값임.

R^2, R^3에서의 길이의 개념을 임의의 내적공간으로 일반화한 것. 벡터 $x = (a, b, c) \in R^3$의 길이가 $\sqrt{a^2 + b^2 + c^2} = \sqrt{\langle x, x\rangle}$인 것을 생각해 보면 유사하다는 것을 알 수 있음.

기하적으로 생각했을 때도 노름(길이)는 R^2, R^3에서의 '길이'와 같은 의미를 가짐.

2) 성질

Definition 68. *F-내적공간 V와 임의의 벡터 $x, y \in V$, 스칼라 $c \in F$에 대해서 아래가 성립함.*

1. $||cx|| = |c| \cdot ||x||$
2. $||x|| = 0 \Leftrightarrow x = 0$*임. 또한 모든* x*에 대해서* $||x|| \geq 0$*임.*
3. *코시-슈바르츠 부등식(Cauchy-Schwarz inequality)* $|\langle x, y\rangle| \leq ||x|| \cdot ||y||$
4. *삼각 부등식(triangle inequality)*[a] $||x + y|| \leq ||x|| + ||y||$

[a]삼각형의 세 변의 길이에 대한 공식.

R^2, R^3 유클리드 길이에서 성립하는 대표 성질은 일반적인 내적공간에서도 성립함.

Proof. 3. $y = 0$이면 성립함. $y \neq 0$인 경우에 대해 생각하자. 임의의 $c \in F$에 대해서 아래가 성립함.

$$0 \geq \langle x - cy, x - cy\rangle = \langle x, x\rangle - \bar{c}\langle x, y\rangle - c\langle y, x\rangle + c\bar{c}\langle y, y\rangle$$

특히 $c = \frac{\langle x,y\rangle}{\langle y,y\rangle}$라 하면, $\frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle} = \frac{|\langle x,y\rangle|^2}{||y||^2}$ 이므로[48] 위 부등식을 아래와 같이 정리할 수 있음.

$$0 \geq \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{||y||^2} = ||x||^2 - \frac{|\langle x, y\rangle|^2}{||y||^2}$$

정리하면 코시-슈바르츠 부등식을 얻을 수 있음.

방정식과 근의 공식을 이용한 방법도 있음. 여종헌 필기 참고.

4. 코시-슈바르츠 부등식을 이용함. 프리드버그 p.357 참고. □

3) 사이각

Definition 69. *F-내적공간 V와 임의의 벡터 x, y에 대해서 코시-슈바르츠 부등식을 정리하면 아래와 같음.*

$$-1 \leq \frac{|\langle x, y\rangle|^2}{||x||^2||w||^2} \leq 1$$

이를 이용하여 아래와 같이 $\cos\theta$를 정의할 수 있는데, 이때의 θ를 x와 y의 사이각이라고 함.

$$\cos\theta = \frac{|\langle x, y\rangle|}{||x|| \cdot ||w||}$$

4) 여러 가지 노름

필요에 따라 노름을 다르게 정의해 사용하기도 함. 내적을 이용하지 않는 노름도 존재함.

L^2-norm : $v = (a_1, \cdots, a_n)$, $||v||_2 = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2}$
내적(dot product)을 이용한 노름임. 이 장에서 다루는 노름은 모두 이 유형임.

L^p-norm : $v = (a_1, \cdots, a_n)$, $||v||_p = (|v_1|^p + |v_2|^p + \cdots + |v_n|^p)^{\frac{1}{p}}$
p에는 정수, inf 등이 들어갈 수 있음. p가 2가 아닌 경우는 내적을 사용하는 것이 아님. p가 커질수록 $v_1, \cdots, v_n$ 중 가장 큰 값의 영향이 커짐.

[48]복소수의 성질에 의하면 $z\bar{z} = |z|^2$임.

4. 직교(orthogonal)

1) 정의

Definition 70. *내적공간 V를 생각하자. V의 벡터 x, y에 대해서 $\langle x, y\rangle = 0$이면 두 벡터는 직교(othogonal) 또는 수직(perpendicular)이라고 정의함. 또한,*

1. *V의 부분집합 S에 대해서 S에 속하는 서로 다른 임의의 두 벡터가 직교할 때, 집합 S를 직교(orthogonal) 집합이라고 함. (모든 원소가 서로 직교.)*
2. *$||x|| = 1$인 벡터 $x \in V$를 단위벡터(unit vector)라고 함.*
3. *V의 부분집합 S가 직교집합이고 단위벡터로만 이루어져 있을 때, 집합 S를 정규직교(orthonormal) 집합이라고 함. (모든 원소가 서로 직교이고 단위벡터.)*

즉, 집합 $S = \{v_1, v_2, \cdots\}$가 정규직교집합이기 위한 필요충분조건은 $\langle v_i, v_j\rangle = \delta_{ij}$인 것임.

기하적으로 생각했을 때 직교는 R^2, R^3에서의 '직교'와 같은 의미를 가짐. 다시 말해, 두 벡터가 수직이라는 것.

2) 정규화

Definition 71. *영이 아닌 벡터에 길이의 역수만큼의 스칼라를 곱해서 단위벡터로 만드는 것을 정규화(normalizing)라고 함.*

벡터에 영이 아닌 스칼라를 곱해도 직교성에 영향이 가지 않고, 영이 아닌 임의의 벡터 x에 대해서 $\frac{x}{||x||}$는 단위벡터임. 즉, 직교집합의 각 원소에 적절한 스칼라를 곱해서 정규직교집합으로 만들 수 있음.

2. 그람-슈미트 직교화와 직교여공간

1. 정규직교기저(orthonormal basis)

1) 정의

> **Definition 72.** *내적공간 V의 순서기저 중 정규직교집합인 것을 정규직교기저(orthonormal basis)라고 함.*

즉, 서로 직교이면서 각각이 단위벡터인 순서기저.

표준 순서기저는 정규직교기저임.

벡터공간을 구성하는 것이 기저였다면, 내적공간을 구성하는 것은 정규직교기저임.

Theorem 6.7에 의하면 정규직교집합을 확장하여 정규직교기저를 만들 수 있음.

2) 쓸모

Theorem 6.3에 의하면 정규직교집합(직교집합)의 일차결합으로 표현된 벡터는 각각의 계수를 간단히 계산할 수 있음.

즉, 임의의 벡터를 정규직교기저의 일차결합으로 나타낼 수 있음. 점공간이 아닌 유한차원 내적공간 V가 정규직교기저 β를 가지고, $\beta = \{v_1, \cdots, v_n\}$이라 하면 $x \in V$를 아래와 같이 나타낼 수 있음.

$$x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$$

각각의 계수를 안다면, 정규직교기저에 대한 행렬표현의 성분을 간단히 구할 수 있음. V의 선형연산자 T에 대해서 $A = [T]_\beta$일 때 임의의 i, j에 대해 아래가 성립함.

$$A_{ij} = \langle T(v_j), v_i \rangle$$

3) 푸리에 계수(Fourier coefficient)

> **Definition 73.** *내적공간 V의 정규직교 부분집합 β와 $x \in V$를 생각하자. $y \in \beta$일 때 $\langle x, y \rangle$를 β에 대한 x의 푸리에 계수(Fourier coefficient)라고 함.*

즉, 어떤 벡터에 대해서, 정규직교 부분집합의 원소와의 내적으로 얻은 계수를 의미함.

2. 그람-슈미트 직교화(Gram-Schmidt process)

1) 정의

Definition 74. *내적공간 V와 일차독립인 부분집합 $S = \{w_1, \cdots, W_n\}$, 스칼라 $a_1, \cdots, a_{k-1} \in F$에 대해서 집합 $S' = \{v_1, \cdots, v_n\}$을 아래와 같이 정의하자.*

$$v_1 = w_1,\ v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{||v_j||^2} v_j\ \ (2 \le k \le n)$$

이때, S'은 $span(S') = span(S)$이고 영이 아닌 벡터로 이루어진 직교집합임.

그람-슈미트 직교화로 임의로 주어진 일차독립인 집합과 동일한 부분공간을 생성하는 직교집합을 얻을 수 있음.

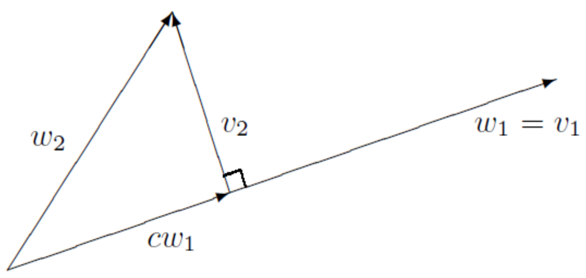
이는 점공간이 아닌 모든 유한차원 내적공간이 정규직교기저를 가짐을 보증함.

그람-슈미트 직교화를 통해 얻은 직교집합을 정규화하면 정규직교집합을 얻을 수 있음. Theorem 6.3의 Corollary에 의해 이 정규직교집합은 일차독립이므로, 원소의 개수가 차원과 같아질 때까지 정규직교집합을 구성하면 정규직교기저를 얻을 수 있음.

계산에는 위 수식을 활용하고, 이해에 기하적인 해석을 사용하자.

Proof. 1. 수학적 귀납법 사용
프리드버그 p.367 참고.

2. 기하적 해석
v_1은 w_1과 동일하게 설정함. k번째 원소는 1부터 $k-1$까지의 원소들과 직교해야 하므로, 해당 원소들로 이루어진 부분공간(S라 하자.)에 직교해야 함. S의 k번째 원소에서 이 부분공간에 수선의 발(사영)을 내린 선분이 k번째 원소가 됨. 다시 말해, w_k의 $S^{\perp}$에 대한 정사영이 v_k가 되는 것. □



이를 수식적으로 나타내려면 w_k를 $v_1, \cdots v_{k-1}$의 일차결합과 v_k의 합으로 나타내고, $1 \le i \le k-1$에 대해서 $\langle v_k, v_i \rangle = 0$임을 이용해 식을 정리하면 됨.

2) OEB(Orthonormal Eigen Basis)

고유벡터로 이루어진 정규직교기저를 OEB(Orthonormal Eigen Basis)라고 함. 고유기저와 정규직교기저의 특성을 모두 가지고 있는 기저이므로 지금까지 등장한 것들 중에 가장 강력한 기저임.

고유기저에 그람-슈미트 직교화와 정규화를 적용하면 OEB를 얻을 수 있음.

3) 르장드르 다항식(Legendre polynomial)

Definition 75. *$P(R)$의 기저 $\{1, x, x^2, \cdots\}$에 그람-슈미트 직교화를 반복하면 직교기저 $\{v_1, v_2, \cdots\}$을 얻을 수 있음. 각 n에 대해서 다항식 $\frac{v_k}{v_k(1)}$를 k차 르장드르 다항식이라고 함. 르장드르 다항식의 집합은 $P(R)$의 직교기저임.*

3. 직교여공간(orthogonal complement)

1) 정의

Definition 76. *내적공간 V의 공집합이 아닌 부분집합 S에 대해서, S의 모든 벡터와 수직인 V의 벡터의 집합을 $S^{\perp}$라 하자. 이 집합을 S의 직교여공간(orthogonal complement)이라 함.*

즉, $S^{\perp} = \{x \in V | y \in S, \langle x, y \rangle = 0\}$임.

2) 성질

임의의 내적공간 V와 그 부분집합 S에 대해서 아래가 성립함.

1. $S^{\perp}$은 V의 부분공간임.
2. $\{0\}^{\perp} = V$
3. $V^{\perp} = \{0\}$

3) 직교여공간과 직합

Theorem 내적공간 V와 그 부분공간 W에 대해서, $V = W \oplus W^{\perp}$가 성립함.

Proof. 영벡터가 아닌 임의의 벡터 $w \in W$에 대해서 $w \in W^{\perp}$이면 $\langle w, w \rangle \neq 0$이므로 모순임. 즉, $W \cap W^{\perp} = \{0\}$임.

Theorem 6.7에 의해, W의 정규직교기저를 확장하여 V의 정규직교기저를 만들면 W의 정규직교기저가 아닌 부분은 $W^{\perp}$의 정규직교기저임. 즉, 임의의 벡터 $x \in V$를 W와 $W^{\perp}$로 표현할 수 있으므로 $W + W^{\perp} = V$임. □

4. 정사영(orthogonal projection)

1) 정의

Definition 77. *Theorem 6.6과 그 Corollary에서 보인 u를 y의 W에 대한 정사영(orthogonal projection)이라고 함.*

즉, W에 대해서 직교인 성분을 제거하여 W에 비추어 나타낸(사영) 것.

5. 관련 정리

1) 정규직교기저의 쓸모

Theorem 6.3 내적공간 V와 영이 아닌 벡터로 이루어진 V의 직교 부분집합 $S = \{v_1, \cdots, v_k\}$를 생각하자. $y \in span(S)$, $a_1, \cdots, a_k \in F$에 대해서 아래가 성립함.

$$y = \sum_{i=1}^{k} a_i v_i = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{||v_i||^2} v_i$$

S가 정규직교집합이면 $y \in span(S)$일 때 아래가 성립함.

$$y = \sum_{i=1}^{k} a_i v_i = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$$

즉, 어떤 벡터를 S의 원소들의 일차결합으로 표현했을 때의 각 스칼라 값을 알 수 있음.

Corollary 내적공간 V와 영이 아닌 벡터로 이루어진 V의 직교 부분집합 S를 생각하자. 집합 S는 일차독립임.

기하적으로 생각해 보면 직교하는 벡터들끼리는 당연히 일차독립일 수밖에 없음.

Proof. S가 직교 부분집합이고 $\sum_{i=1}^{k} a_i v_i = y = 0$이라고 하자. $y = 0$이면 모든 j에 대해서 $a_j = \frac{\langle 0, v_j \rangle}{||v_j||^2} = 0$임. 즉, $v_1, \cdots, v_k$의 일차결합으로 0을 만드는 방법은 자명한 방법밖에 없음. □

2) 정사영

Theorem 6.6 내적공간 V의 유한차원 부분공간 W와 벡터 $y \in V$에 대해서 $y = u + z$인 유일한 벡터 $u \in W$와 $z \in W^{\perp}$가 존재함. 또한 $\{v_1, \cdots, v_k\}$가 W의 정규직교기저일 때, 아래가 성립함.

$$u = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$$

y는 V의 기저로 표현할 수 있고, $\langle y, v_i \rangle$에서 v_i를 제외한 기저의 계수는 제거되므로, $\langle y, v_i \rangle$이 v_i의 계수가 됨.

Proof. W와 $W^{\perp}$은 직합이므로, Theorem 5.9에 의해 성립함. □

Corollary Theorem 6.6의 표기법을 따를 때 벡터 u는 W의 벡터 중 y에 가장 가까운 유일한 벡터임[49]. 수식으로 표현하면 임의의 $x \in W$에 대해서 $||y - x|| \geq ||y - u||$임. 등호는 $x = u$일 때 성립함.

여기서 '어떤 벡터와 가장 가까운 벡터'란 어떤 벡터와 뺐을 때의 길이(노름)이 가장 작은 것을 말함.

3) 정규직교집합의 확장과 직교여공간과의 직합

Theorem 6.7 n차원 내적공간 V의 정규직교집합 $S = \{v_1, \cdots, v_k\}$에 대해서 아래가 성립함.

1. S를 확장하여 V의 정규직교기저 $\{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$을 얻을 수 있음.
2. $W = span(S)$일 때, $S_1 = \{v_{k+1}, v_{k+2}, \cdots, v_n\}$은 $W^{\perp}$의 정규직교기저임.
3. V의 임의의 부분공간 W에 대해서 $dim(V) = dim(W) + dim(W^{\perp})$임.

즉, 정규직교집합을 확장하여 정규직교기저를 만들 수 있음. 또한 부분집합 W의 기저와 $W^{\perp}$을 합치면 전체 내적공간의 기저가 됨. 더 나아가 생각해 보면 $V = W \oplus W^{\perp}$인 것.

Proof. 1. 대체정리의 Corollary 2와 그람-슈미트 직교화에 의해 성립함.

2. S_1은 기저의 부분집합이므로 일차독립임. 즉, $span(S_1) = W^{\perp}$임을 보이면 됨.

임의의 벡터 $x \in V$는 $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$로 나타낼 수 있음. 이때 $x \in W^{\perp}$이라면 $x = \sum_{i=k+1}^{n} v_i$인데, 이 경우 $x \in span(S_1)$이므로 성립함. □

[49] 여기서 가까운 벡터라는 것은 W위의 벡터들 중 y와 뺐을 때 가장 작은 크기를 가지는 벡터를 의미함.

3. 수반연산자

1. 수반연산자(adjoint)

1) 정의

Definition 78. *Theorem 6.9의 선형연산자 T^*를 T의 수반연산자(adjoint)라 함.*

즉, T^*는 모든 $x, y \in V$에 대해서 $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$를 만족시키는 유일한 선형연산자임.

Theorem $\langle x, T(y)\rangle = \langle T^*(x), y\rangle$가 성립함.

Proof. $\langle x, (T(y))\rangle = \overline{\langle T(y), x\rangle} = \overline{\langle y, T^*(x)\rangle} = \langle T^*(x), y\rangle$ □

내적기호 안에서 T의 위치를 바꿀 때 $*$가 붙는다고 생각하면 편함.

2) 확장된 정의

Definition 79. *유한차원 내적공간 V, W에 내적이 각각 $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$로 주어져 있음. 선형변환 $T : V \to W$와 모든 $x \in V$와 모든 $y \in W$에 대해서 $\langle T(x), y\rangle_1 = \langle x, T^*(y)\rangle_2$인 함수 $T^* : W \to V$를 T의 수반연산자(adjoint)라 함.*

선형연산자에 대한 수반연산자의 정의를 임의의 선형변환 $T : V \to W$로 확장한 것.

3) 수반연산자의 계산

Theorem 6.10에 의해, T^*의 행렬표현은 T 행렬표현의 수반행렬(켤레전치)와 같음. 즉, T^*의 행렬표현은 T 행렬표현의 수반행렬을 구함으로써 알아낼 수 있음.

4) 수반연산자와 켤레복소수의 유사성

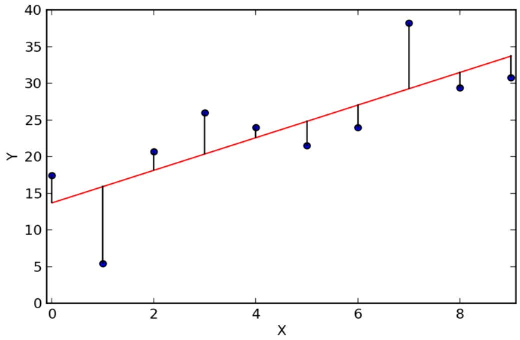
Theorem 6.11과 그 Corollary에 의하면 수반연산자와 켤레복소수는 서로 유사한 성질을 가짐.

2. 최소제곱법(least squares approximation)

1) 최소제곱직선(least squares line)

어떤 데이터를 좌표평면 위의 점들로 나타냈다고 해 보자. 이 점들의 경향성을 나타내는 직선을 구하려고 함. 데이터의 각 점과 이 직선 사이에 x축에 수직인 직선으로 주어지는 거리의 제곱의 합을 오차 E라고 하면, 이 오차가 최소가 되는 직선이 이 데이터들을 가장 잘 대표한다고 할 수 있음. 이 직선을 최소제곱직선(least squares line)이라고 함.

최소제곱법은 좌표평면 위 점들에 대한 경향성을 나타내는 최적의 다항식을 근사적으로 구하는 방법임. 최소제곱직선은 최소제곱법으로 구할 수 있는 다항식 중 하나임.



2) 과정

n차 다항식에 대해서 m개의 데이터 $(t_1, y_1), (t_2, y_2), \cdots, (t_m, y_m)$이 있다고 하자. 이 데이터들의 경향성을 나타내는 다항식 $y = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$을 구하려고 함. 이때의 $a_n, \cdots a_0$은 x축에 수직인 직선으로 주어지는 거리의 제곱의 합인 오차 E를 최소화함.

A, x, y를 각각 아래과 같이 정의하자.

$$A = \begin{pmatrix} t_1^n & t_1^{n-1} & \cdots & 1 \\ t_2^n & t_2^{n-1} & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ t_m^n & t_m^{n-1} & \cdots & 1 \end{pmatrix}, \ x = \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \end{pmatrix}, \ y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

오차 E는 아래와 같음.

$$E = \sum_{i=1}^{n} (y_i - a_n t_n - \cdots - a_0)^2 = ||y - Ax||^2$$

E를 최소화하는 벡터 x_0을 구하는 것이 목표임. $m \times n$ 행렬 A와 $y \in F^m$에 대해서 $W = \{Ax | x \in F^n\}$이라 하자. $W = R(L_A)$임. Theorem 6.6(정사영)의 Corollary에 의해 y에 가장 가까운 벡터가 유일하게 존재함. 이를 수식적으로 표현하면, 이 벡터를 x_0이라고 했을 때 모든 x에 대해서 $||y - Ax_0|| \le ||y = Ax||$가 성립함. 이 x_0은 지금 구하려고 하는 벡터임.

이제 x_0을 구하는 구체적인 방법에 대해 생각해 보자. Theorem 6.6과 그 Corollary에 의하면, $Ax_0 - y \in W^{\perp}$이므로 모든 $x \in F^n$에 대해서 $\langle Ax, Ax_0 - y \rangle_m = 0$임. Lemma 1에 의해 $\langle x, A^*(Ax_0 - y) \rangle = 0$으로 정리하면 $A^*Ax_0 - A^*y = 0$이고 $A^*Ax_0 = A^*y$임. 즉, 방정식 $A^*Ax_0 = A^*y$의 해를 구하면 목표했던 다항식을 알게 됨.

$rank(A) = n$인 경우 Lemma 2의 Corollary에 의해 A^*A가 가역이므로 $A^*Ax_0 = A^*y$이 유일한 해 $x_0 = (A^*A)^{-1}A^*y$를 가짐. 이 경우 더 쉽게 x_0을 구할 수 있음.

3) 결론

'2) 과정'의 표기법을 그대로 따름. 목표했던 다항식에 대한 A, x, y을 가지고 $A^*Ax_0 = A^*y$의 해를 구하면 해당 다항식을 알 수 있음.

$rank(A) = n$인 경우 $x_0 = (A^*A)^{-1}A^*y$로 더 쉽게 알아낼 수 있음.[50]

오차 E는 $E = ||Ax_0 - y||^2$으로 구함.

[50] 최소제곱직선을 구하는 경우 A는 $m \times 2$ 행렬인데 이 경우 $rank(A) = 2$인 것이 일반적임. A의 두 번째 열은 모든 성분이 1이므로 랭크가 1이라는 것은 실험자가 단 한 번 측정을 했다는 것이므로 현실성이 떨어짐.

3. 연립일차방정식의 최소해

1) 정의

> **Definition 80.** *방정식 $Ax = b$의 임의의 해 u에 대해서 $||s|| \leq ||u||$를 만족하는 해 s를 최소해(minimal solution)라 함.*

즉, 연립일차방정식의 해 중 노름이 가장 작은 것이 최소해임.

2) 구하는 법

Theorem 6.13에 의하면 모순이 없는 연립일차방정식은 유일한 최소해를 가짐.

$(AA^*)x = b$의 해 중 하나를 u라 하면, 최소해는 $s = A^*u$임.

4. 관련 정리

1) 선형범함수(linear functional)의 유일한 내적 표현

Theorem 6.8 유한차원 F-내적공간 V와 선형변환 $g : V \to F$를 생각하자. 모든 $x \in V$에 대해서 $g(x) = \langle x, y \rangle$를 만족하는 벡터 $y \in V$가 유일하게 존재함. 이때, $y = \sum_{i=1}^{n} \overline{g(v_i)} v_i$로 y를 구할 수 있음.

즉, V에서 F로 가는 선형범함수를 유일한 내적 표현으로 나타낼 수 있음.

Proof. V의 정규직교기저를 $\beta = \{v_1, \cdots, v_n\}$이라 하고, $y = \sum_{i=1}^{n} \overline{g(v_i)} v_i$라 하자. 함수 $h : V \to F$를 $h(x) = \langle x, y \rangle$로 정의하면 $1 \le j \le n$에 대해서 아래가 성립함.

$$h(v_j) = \langle v_j, y \rangle = \sum_{i=1}^{n} g(v_i) \langle v_j, v_i \rangle = \sum_{i=1}^{n} g(v_i) \delta_{ji} = g(v_j)$$

즉, $g = h$이므로 $y = \sum_{i=1}^{n} \overline{g(v_i)} v_i$를 값으로 가지는 선형변환은 g가 유일함. 다시 말해, $g(x) = \langle x, y \rangle$로 표현할 수 있음.

$g(x) = \langle x, y' \rangle$인 y'이 있다고 해 보자. $\langle x, y \rangle = \langle x, y' \rangle$이고 Theorem 6.1에 의해 $y = y'$임. 즉, g에 대해서 y가 유일함. □

2) 수반연산자

Theorem 6.9 유한차원 내적공간 V와 선형연산자 $T : V \to V$를 생각하자. 모든 $x, y \in V$에 대해서 $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$인 함수 $T^* : V \to V$가 유일하게 존재함. 특히 T^*는 선형변환임.

Proof. $y \in V$를 고정하고, 모든 $x \in V$에 대해서 함수 $g : V \to F$를 $g(x) = \langle f(x), y \rangle$라 정의하자.

1. g는 선형인가?
T가 선형이므로 g는 선형인 것을 알 수 있음.

g가 선형이므로, Theorem 6.8에 의해 $g(x) = \langle T(x), y \rangle = \langle x, y' \rangle$으로 나타낼 수 있음. $y' = T^*(y)$라 정의하면 $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$임.

2. T^*는 선형인가?
$\langle x, T^*(cy_1 + y_2) \rangle = \langle T(x), cy_1 + y_2 \rangle$를 전개해 보면 T^*가 선형인 것을 알 수 있음.

3. T^*은 유일한가? $\langle T(x), y \rangle = \langle x, U(y) \rangle$인 $U : V \to V$가 있다고 해 보자. $\langle x, T^*(y) \rangle = \langle x, U(y) \rangle$이고 Theorem 6.1에 의해 $T^* = U$임. 즉, T^*가 유일함. □

3) 수반연산자의 행렬표현과 행렬표현의 수반행렬

Theorem 6.10 유한차원 벡터공간 V와 그 정규직교기저 β, V의 선형연산자 T에 대해서 아래가 성립함.

$$[T^*]_\beta = ([T]_\beta)^*$$

즉, T^*의 행렬표현과 T 행렬표현의 수반행렬(켤레전치)이 같음.

Proof. $A = [T]_\beta$, $B = [T^*]_\beta$, $\beta = \{v_1, \cdots, v_n\}$이라 하자. 아래의 식이 성립함.

$$B_{ij} = \langle T^*(v_j), v_i \rangle = \langle v_j, T(v_i) \rangle = \overline{\langle T(v_i), v_j \rangle} = \overline{A_{ji}} = (A^*)_{ij}$$

□

Corollary $n \times n$ 행렬 A에 대해서 $L_{A^*} = (L_A)^*$임.
즉, 좌측 곱 연산의 수반연산자는 해당 행렬의 수반행렬을 곱하는 좌측 곱 연산임.

Proof. β가 표준기저일 때 $[L_A]_\beta = A$이므로 아래가 성립함.

$$[(L_A)^*]_\beta = ([(L_A)]_\beta)^* = A^* = [L_{A^*}]_\beta$$

즉, $L_{A^*} = (L_A)^*$임. □

4) 켤레복소수와 수반연산자의 유사성

Theorem 6.11 내적공간 V와 그 수반연산자가 존재하는 선형연산자 T, U에 대해서 아래가 성립함.

1. $T + U$의 수반연산자가 존재하고 $(T + U)^* = T^* + U^*$
2. 임의의 $c \in F$에 대해서 cT의 수반연산자가 존재하고 $(cT)^* = \bar{c}T^*$임.
3. TU의 수반연산자가 존재하고 $(TU)^* = U^*T^*$임.
4. T^*의 수반연산자가 존재하고 $T^{**} = T$임.
5. I의 수반연산자가 존재하고 $I^* = I$임.

Corollary $n \times n$ 행렬 A와 B에 대해서 아래가 성립함.

1. $(A + B)^* = A^* + B^*$
2. $c \in F$에 대해서 $(cA)^* = \bar{c}A^*$
3. $(AB)^* = B^*A^*$
4. $A^{**} = A$
5. $I^* = I$

즉, 행렬에 대해서도 성립함.

5) 최소제곱법

두 벡터 $x, y \in F^n$의 표준 내적을 $\langle x, y\rangle_n$이라 표기함. x, y를 열벡터 취급하면 $\langle x, y\rangle_n = y^*x$로 표기할 수 있음.

Lemma 1 $A \in M_{m\times n}(F), x \in F^n, y \in F^m$에 대해서 아래가 성립함.

$$\langle Ax, y\rangle_m = \langle x, A^*y\rangle_n$$

Proof. $\langle Ax, y\rangle_m = y^*(Ax) = (y^*A)x = (A^*y)^*x = \langle x, A^*y\rangle_n$ □

Lemma 2 $A \in M_{m\times n}(F)$에 대해서 $rank(A^*A) = rank(A)$임.

Proof. 차원정리에 의하면, $rank(A^*A) = n - nullity(A^*A)$, $rank(A) = n - nullity(A)$이므로 $A^*Ax = 0$이기 위한 필요충분조건이 $Ax = 0$임을 보이면 됨.

1. $A^*Ax = 0 \to Ax = 0$
$0 = \langle A^*Ax, x\rangle = \langle Ax, A^{**}x\rangle = \langle Ax, Ax\rangle$이므로 $Ax = 0$임.

2. $Ax = 0 \to A^*Ax = 0$
당연함. □

Corollary $rank(A) = n$인 $m \times n$ 행렬 A에 대해서 A^*A는 가역임.

Proof. A^*A는 $n \times n$ 행렬이므로 당연함. □

Theorem 6.12 $A \in M_{m\times n}(F), y \in F^m$이 주어지면, 모든 $x \in F^n$에 대해서 $||AX_0 - y|| \le ||Ax - y||$이고, $(A^*A)x_0 = A^*y$인 벡터 $x_0 \in F^n$이 존재함. 특히 $rank(A) = n$이면 $x_0 = (A^*A)^{-1}A^*y$임.

이는 추후에 등장할 유사역행렬과 연립일차방정식의 관계와도 연관이 있음.

6) 연립일차방정식의 최소해

Corollary 6.13 $A \in M_{m\times n}(F), b \in F^m$에 대해서 $Ax = b$가 모순이 없는 연립일차방정식이라 가정하자. 아래가 성립함.

1. $Ax = b$의 유일한 최소해 s가 존재하고 $s \in R(L_{A^*})$임.
2. 벡터 u가 $(AA^*)u = b$를 만족하면 $s = A^*u$임.

Proof. 프리드버그 p.387 참고. □

4. 정규연산자와 자기수반연산자

1. 정규연산자(normal operator)

1) 정의

Definition 81. *내적공간 V와 그 선형연산자 T를 생각하자.*

$TT^* = T^*T$인 선형연산자 T를 정규연산자(normal operator)라고 함.

$AA^* = A^*A$인 $n \times n$ 실행렬/복소행렬 A를 정규행렬(normal matrix)이라 함.

Theorem 6.10에 의해, 정규직교기저 β에 대해서 T가 정규연산자이기 위한 필요충분조건은 $[T]_\beta$가 정규행렬인 것임.

Theorem 6.15를 보면 정규연산자가 여러 성질들을 가지는 것을 확인할 수 있음.

2) T-불변

Definition 82. *V의 부분공간 W에 대해서 $T(W)$가 W의 부분집합일 때, W를 T-불변(T-invariant)이라 함. W가 T-불변이면 제한 $T_W : W \to W$를 아래와 같이 정의함.*

$$\text{모든 } x \in W\text{에 대해서 } T_W(x) = T(x)$$

즉, W에 대해서 T가 닫혀 있는 것으로 이해할 수 있음. T_W는 정의역을 W로 제한한 것.

3) 정규연산자와 OEB

Theorem 6.16에 의하면, 복소내적공간 V의 선형연산자 T가 정규연산자이기 위한 필요충분조건은 T에 대한 정규직교 고유기저(Orthonormal Eigen Basis, OEB)가 존재하는 것임.

즉, 정규연산자이면 복소내적공간에 대해 OEB가 존재한다는 것.

V가 실내적공간이면 T의 특성다항식이 완전히 인수분해되지 않으므로 성립하지 않음.

2. 자기수반연산자(self-adjoint)

1) 정의

Definition 83. *내적공간 V의 선형연산자 T를 생각하자.*

$T = T^*$인 T를 자기수반연산자(self-adjoint) 또는 에르미트 연산자(Hermitian)라 함.

$A = A^*$인 $n \times n$ 실행렬 또는 복소행렬 A를 자기수반행렬(self-adjoint matrix) 또는 에르미트 행렬(Hermitian matrix)이라 함.

실내적공간에서의 자기수반행렬은 실대칭행렬(Symmetric matrix)이라 함.

정규직교기저 β에 대해서 T가 자기수반연산자이기 위한 필요충분조건은 $[T]_\beta$가 자기수반행렬인 것임.

실행렬에서는 자기수반행렬인 조건을 해당 행렬이 대칭인 것($A = A^t$)으로 생각할 수 있음.

Theorem 6.17의 증명 과정에 의하면, 자기수반연산자의 고윳값은 모두 실수임.

2) 자기수반연산자와 OEB

Theorem 6.17에 의하면, 실내적공간 V의 선형연산자 T가 자기수반연산자이기 위한 필요충분조건은 T에 대한 정규직교 고유기저(Orthonormal Eigen Basis, OEB)가 존재하는 것임.

자기수반연산자는 정규연산자 중 더 강한 조건을 가진 연산자임.

3) 총정리

정규연산자는 복소내적공간에서 OEB의 존재를 보증함. 자기수반연산자는 실내적공간에서 OEB의 존재를 보증함. 복소수체에서 실수체로 범위를 줄이면서 OEB의 존재를 보증하는 연산자의 조건을 더 엄격하게 잡은 것임.

실대칭행렬은 실 자기수반행렬이고, 자기수반행렬은 정규행렬임.(역은 성립하지 않음.)

정규연산자와 자기수반연산자는 실내적공간이든 복소내적공간이든 존재할 수 있음. 하지만 OEB의 존재성에 대해서는 체에 따른 차이가 존재함.

1. $F = C$ (복소내적공간)
정규연산자 $\iff$ OEB 존재.
자기수반연산자 $\Longrightarrow$ 정규연산자[51] $\Longrightarrow$ OEB 존재.

2. $F = R$ (실내적공간)
자기수반연산자 $\iff$ OEB 존재.
정규연산자 $\Longrightarrow$ X. (특성다항식을 완전하게 인수분해할 수 없음.)

3. 관련 정리

1) 정규연산자의 성질

Theorem 6.15 내적공간 V와 그 정규연산자 T에 대해서 아래가 성립함.

1. 모든 $x \in V$에 대해서, $||T(x)|| = ||T^*(x)||$임.
2. 임의의 $c \in F$에 대해서, $T - cI$는 정규연산자임.
3. 고윳값 λ에 대응하는 T의 고유벡터 x는 고윳값 $\bar{\lambda}$에 대응하는 T^*의 고유벡터이기도 함. 즉, $T(x) = \lambda x$이면 $T^*(x) = \bar{\lambda} x$임.
4. 고유벡터 x_1, x_2에 대응하는 T의 고윳값을 각각 λ_1, λ_2라 하자. $\lambda_1 \neq \lambda_2$이면 x_1와 x_2는 직교함.

자기수반연산자는 정규연산자이므로 자기수반연산자에 대해서도 위 정리가 성립함.

Proof. 2. $U = T - cI$에 대해서 $UU^* = U^*U$가 성립함.

3. $\langle v_i, Tv_j \rangle = \bar{\lambda} \langle v_i, v_j \rangle = \langle \bar{\lambda} v_i, v_j \rangle = \langle T^* v_i, v_j \rangle$

4. $\lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \langle T(x_1), x_2 \rangle = \langle x_1, T^*(x_2) \rangle = \langle x_1, \bar{\lambda_2} x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$인데, $\lambda_1 \neq \lambda_2$이므로 $\langle x_1, x_2 \rangle = 0$임. □

[51] 역은 성립하지 않음.

2) 정규연산자와 OEB

Theorem 6.16 유한차원 복소내적공간 V의 선형연산자를 T를 생각하자. T가 정규연산자이기 위한 필요충분조건은 T의 고유벡터로 이루어진 V의 정규직교기저(Orthonormal Eigen Basis, OEB)가 존재하는 것임.

Proof. 1. OEB 존재 $\rightarrow$ T가 정규연산자
OEB를 β라 하자. $[T]_\beta$와 $[T^*]_\beta = ([T]_\beta)^*$는 대각행렬이므로 T와 T^*는 가환적임. 즉, $TT^* = T^*T$임.

2. T가 정규연산자 $\rightarrow$ OEB 존재
방법1. 슈어의 정리(Schur's Theorem)로 증명.
프리드버그 p.394 참고.

방법2. 귀납법으로 증명.
대수학의 기본 정리에 의해 복소수체에서 T의 특성다항식은 완전히 인수분해됨. 이 경우 T에 대해서 고유벡터 v와 고윳값 λ가 적어도 하나씩은 존재함. v에 대해서 V를 $V = span(v_1) \oplus W$로 분해하자. 이때 $W = (span(v_1))^\perp$임. $dim(W) = n - 1$이고 W의 벡터들은 v와 직교하므로, W에 고유벡터가 존재하는지만 확인하면 귀납적으로 OEB가 존재함을 알 수 있음. 귀납적 과정을 위해선 아래의 3가지를 확인해야 함.

1. W가 T-불변인가?
W가 T-불변이 아니면 함수값이 공역 밖에 있는 것이므로 귀납적인 증명이 불가능함. 임의의 $x \in W$에 대해서 $\langle v, Tx\rangle = \langle T^*v, x\rangle = \bar{\lambda}\langle v, x\rangle = 0$이므로 W는 T-불변임.

2. W가 T^*-불변인가?
임의의 $x \in W$에 대해서 $\langle v, T^*x\rangle = \langle Tv, x\rangle = \lambda\langle v, x\rangle = 0$이므로 W는 T^*-불변임.

3. T_W도 정규연산자인가?
$TT^* = T^*T$가 W에 대해 닫혀 있는지만 보이면 됨. 1, 2번에 의해, 임의의 $x \in W$에 대해서 $TT^*x, T^*Tx \in W$임. $\square$

3) 자기수반연산자와 OEB

Theorem 6.17 유한차원 실내적공간 V의 선형연산자 T에 대해서 T가 자기수반연산자이기 위한 필요충분조건은 T의 고유벡터로 이루어진 V의 정규직교기저(Orthonormal Eigen Basis)가 존재하는 것임.

Proof. 1. OEB 존재 $\rightarrow$ T가 자기수반연산자
β가 T의 OEB라 하자. $[T]_\beta$는 대각행렬이므로, $[T]_\beta = ([T]_\beta)^*$가 성립함.

2. T가 자기수반연산자 $\rightarrow$ OEB 존재
$TT^* = T^*T$이므로 T는 정규연산자임. Theorem 6.16에 의해, T의 모든 고윳값이 실수인지를 확인하여 T의 특성다항식이 실수체에서 완전히 인수분해된다는 것을 보이면 됨. $x \neq 0$에 대해서 $T(x) = \lambda x$라 하자. $T = T^*$이고 $\lambda x = T(x) = T^*(x) = \bar{\lambda}x$이므로 $\lambda = \bar{\lambda}$임. 즉, λ는 실수임. $\square$

5. 유니타리 연산자와 직교연산자

유니타리/직교 연산자는 길이(내적)를 보존하는 함수임.[52]

1. 유니타리(unitary)/직교(orthogonal) 연산자(operator)

1) 정의

Definition 84. *유한차원 F-내적공간 V의 선형연산자 T를 생각하자.*

$F = C$일 때 모든 $x \in V$에 대해서 $||T(x)|| = ||x||$인 T를 유니타리 연산자(unitary operator)라 함.

$F = R$일 때 모든 $x \in V$에 대해서 $||T(x)|| = ||x||$인 T를 직교연산자(orthogonal operator)라 함.

즉, 원래의 길이와 보낸 후의 길이가 동일함.

2) 성질

Theorem 6.18과 그 Corollary들에 의하면 유니타리/직교 연산자는 나름의 성질들을 가짐.

3) 대칭변환(reflection)

Definition 85. *R^2의 1차원 부분공간 L을 생각하자. L은 평면 위 원점을 지나는 직선임. 모든 $x \in L$에 대해서 $T(x) = x$이고, 모든 $x \in L^{\perp}$에 대해서 $T(x) = -x$로 정의한 R^2의 선형연산자 T를 L에 대한 R^2의 대칭변환(reflection)이라 함.*

즉, 벡터를 직선 L에 대해서 대칭시키는 함수임.

Theorem 대칭변환은 직교연산자임.

Proof. 원점을 지나는 직선 L에 대한 R^2의 대칭변환 T와 $||v_1|| = ||v_2|| = 1$을 만족하는 $v_1 \in L, v_2 \in L^{\perp}$을 생각하자. $T(v_1) = v_1, T(v_2) = -v_2$이므로 v_1, v_2는 고유벡터이고 대응하는 고윳값은 각각 $1, -1$임. $\{v_1, v_2\}$는 고유기저이면서 정규직교기저임. 즉, OEB임. Theorem 6.18의 Corollary 1에 의해 T는 직교연산자임. □

2. 유니타리(unitary)/직교(orthogonal) 행렬(matrix)

1) 정의

Definition 86. *$A^*A = AA^* = I$인 정사각행렬 A를 유니타리 행렬이라 하고, $A^tA = AA^t = I$인 정사각행렬 A를 직교행렬(orthogonal matrix)이라고 함.*

내적공간 V의 선형연산자 T가 유니타리/직교 연산자이기 위한 필요충분조건은 V의 적절한 정규직교기저 β에 대해서 $[T]_\beta$가 유니타리/직교 행렬인 것임.

실행렬 A에 대해서 $A^* = A^t$이므로 실 유니타리 행렬은 직교행렬임. 이 경우 A는 그냥 직교행렬이라고 하는 것이 일반적임.

2) 정규행렬과의 관계

유니타리/직교 행렬은 정규행렬임. 역은 성립하지 않음.

유니타리/직교연산자는 정규연산자임. 물론 역은 성립하지 않음.

[52]선형변환은 벡터 합과 스칼라 곱을, 동형사상은 벡터공간의 모든 구조를, 유니타리/직교 연산자는 길이(내적)를 보존하는 함수임.

2) 유니타리/직교 행렬의 가역성

Theorem 유니타리/직교 행렬은 가역임.

Proof. 유니타리 행렬 A를 생각하자. 6.3 연습문제 18에 의해 $det(A^*) = \overline{det(A)}$이므로, $det(A^*A) = \overline{det(A)}det(A) = det(I) = 1$임. $det(A) \neq 0$이므로 A는 가역임. □

유니타리/직교 행렬은 가역이므로 유니타리/직교 행렬에 대해 아래와 같이 표현할 수 있음.

유니타리 행렬 A에 대해, $A^* = A^{-1}, A = (A^*)^{-1}$, 직교 행렬 A에 대해, $A^t = A^{-1}, A = (A^t)^{-1}$.

3) 행/열과 정규직교기저

Theorem $AA^* = I/A^*A = I$는 A의 각 행/열이 F^n의 정규직교기저인 것과 동치임.

Proof. $\delta_{ij} = I_{ij} = (AA^*)_{ij} = \sum_{k=1}^{n} A_{ik}\bar{A}_{jk}$임. 즉, 각 행과 열의 곱이 내적인 것을 확인할 수 있음. □

3. 유니타리(unitary)/직교(orthogonal) 동치(equivalent)

1) 정의

Definition 87. *복소정규[실대칭] 행렬 A에 대해서 F^n의 OEB β가 존재하므로 A는 특정 대각행렬 D와 닮음임. β를 열로 가지며 $D = Q^{-1}AQ$를 만족시키는 Q가 존재하는데, Q의 열벡터가 F^n의 정규직교기저이므로 Q는 유니타리[직교]행렬임.*

이 경우 A를 D와 유니타리(unitary)/직교(orthogonal) 동치(equivalent)라고 함.

즉, 유니타리/직교행렬에 의해 동치(닮음)이 되는 관계를 유니타리/직교 동치라고 하는 것.

유니타리/직교행렬은 가역이므로 유니타리/직교 동치인 두 행렬은 닮음임.

Theorem 두 행렬 A, B가 유니타리/직교 동치이기 위한 필요충분조건은 $A = P^*BP$인 유니타리/직교행렬 P가 존재하는 것임.

2) 유니타리/직교 동치와 정규성

Theorem 6.19, Theorem 6.20에 의하면 유니타리/직교 동치는 실/복소내적공간에서 정규성을 보증함.

대각행렬 A와 행렬 B가 유니타리/직교 동치인 경우를 생각하자. $A = P^*BP$의 P가 유니타리/직교행렬이고 P의 각 행, 열이 정규직교기저임. A가 대각행렬이므로 이 정규직교기저는 OEB임. OEB가 존재하므로 B는 정규행렬/실 자기수반행렬인 것. 즉, 대각행렬과의 유니타리/직교 동치는 OEB의 존재성을 보여줌.

3) 총정리

유니타리/직교 연산자는 정규연산자이고, 길이(내적, 노름)를 보존함.

대각행렬과 유니타리/직교 동치인 행렬은 정규행렬/실 자기수반행렬임.

체에 따른 적용은 아래와 같음.

1. $F = C$ (복소)
유니타리 연산자 $\Longrightarrow$ 정규연산자 $\Longleftrightarrow$ OEB 존재.

대각행렬과 유티나리 동치 $\Longleftrightarrow$ 정규행렬 $\Longleftrightarrow$ OEB 존재.

OEB로 가는 기저변환행렬(P)은 유니타리행렬.

2. $F = R$ (실)
직교연산자 $\Longrightarrow$ X. (실 자기수반연산자라는 보증이 없음.)

대각행렬과 직교 동치 $\Longleftrightarrow$ 실 자기수반행렬 $\Longleftrightarrow$ OEB 존재.

OEB로 가는 기저변환행렬(P)은 직교행렬.

4. 관련 정리

1) 유니타리/직교 연산자의 성질

Theorem 6.18 유한차원 내적공간 V의 선형연산자 T에 대해서 아래의 명제들은 서로 동치임.

1. $T^*T = I$
2. $TT^* = I$
3. 모든 $x, y \in V$에 대해서 $\langle T(x), T(y)\rangle = \langle x, y\rangle$임. (내적 보존)
4. V의 정규직교기저 β에 대해서 $T(\beta)$도 V의 정규직교기저임.
5. $T(\beta)$가 V의 정규직교기저가 되도록 하는 V의 정규직교기저 β가 존재함.
6. 모든 $x \in V$에 대해서 $||T(x)|| = ||x||$임. (노름 보존, 유니타리/직교 연산자의 정의)

1,2번 성질에 의하면 유니타리/직교 연산자는 정규연산자임.

3번 성질을 만족하는 T는 내적을 보존한다고 함. 6번 성질을 만족하는 T는 노름을 보존한다고 함.

Corollary 1 유한차원 실내적공간 V의 선형연산자 T를 생각하자. V가 T의 절댓값이 1인 고윳값에 대응하는 고유벡터로 이루어진 정규직교기저를 포함하기 위한 필요충분조건은 T가 자기수반 직교연산자인 것임.

즉, 실내적공간에서 자기수반 직교연산자는 절댓값이 1인 고윳값만을 가지고, 해당 고윳값에 대한 고유벡터로 OEB를 만들 수 있음.

Corollary 2 유한차원 복소내적공간 V의 선형연산자 T를 생각하자. V가 T의 절댓값이 1인 고윳값에 대응하는 고유벡터로 이루어진 정규직교기저를 포함하기 위한 필요충분조건은 T가 유니타리 연산자인 것임.

즉, 복소내적공간에서 유니타리 연산자는 절댓값이 1인 고윳값만을 가지고, 해당 고윳값에 대한 고유벡터로 OEB를 만들 수 있음.

2) 유니타리/직교 동치와 정규성

Theorem 6.19 $n \times n$ 복소행렬 A를 생각하자. A가 정규행렬이기 위한 필요충분조건은 A가 대각행렬과 유니타리 동치인 것임.

Theorem 6.20 $n \times n$ 실행렬 A를 생각하자. A가 실 자기수반행렬이기 위한 필요충분조건은 A가 실(real) 대각행렬과 직교 동치인 것임.

6. 정사영과 스펙트럼 정리

1. 사영(projection)

1) 정의

Definition 88. *내적공간 V와 $V = W_1 \oplus W_2$인 부분공간 W_1, W_2에 대해서 아래와 같이 정의한 함수 $T : V \to V$를 W_2에 대한 W_1 위로의 V의 사영(projection)이라 함.*

$$x = x_1 + x_2 \text{일 때, } T(x) = x_1 \text{ (단, } x_1 \in W_1,\ x_2 \in W_2\text{)}$$

T는 'W_1로의 사영' 또는 그냥 '사영'이라고 부르기도 함.

즉, a에 대한 b 위로의 사영은, a 성분만 제거하여 x를 b 위로 '비추는' 것.

2) 사영의 영공간과 상공간

Theorem 내적공간 V와 그 부분공간 W_1, W_2에 대해서 $V = W_1 \oplus W_n$이라 하자. $T : V \to V$가 W_1 위로의 사영일 때, $R(T) = W_1$, $N(T) = W_2$이고 $V = R(T) \oplus N(T)$임.

즉, 모든 사영은 치역(상공간)과 영공간에 의해 유일하게 결정됨.

Proof. $R(T) = W_1$이고, $N(T) = W_2$이므로 $V = W_1 \oplus W_2 = R(T) \oplus N(T)$임. □

3) 사영의 합성

Theorem 내적공간 V에 대해서 $T : V \to V$가 $N(T)$에 대한 $R(T)$ 위로의 사영이기 위한 필요충분조건은 $T = T^2$인 것임.

Proof. 1. 사영 $\to T = T^2$
당연함.

2. $T = T^2 \to$ 사영
우선 $V = R(T) \oplus N(T)$인 것을 보이고, T가 $R(T)$로의 사영인 것을 보이면 됨.

임의의 $v \in V$에 대해서 $v = Tv + (v - Tv)$라 하자. $Tv \in R(T)$이고 $T(v - Tv) = Tv - T^2v = 0$, $v - Tv \in N(T)$이므로 $V = R(T) + N(T)$임. 또한 임의의 $y \in R(T) \cap N(T)$에 대해서 $T(y) = 0$이고 $T(x) = y$인 $x \in V$가 존재함. $y = T(x) = T(T(x)) = T(y) = 0$임. 즉, $R(T) \cap N(T) = \{0\}$임. 정리하면 $V = R(T) \oplus N(T)$임.

$u \in R(T), w \in N(T)$라 하자. $T(k) = u$인 $k \in V$가 존재하므로 $T(k) = T(T(k)) = T(u) = u$임. 즉, $T(u + w) = T(u) + T(w) = T(u) = u$이므로 T가 사영임. □

Corollary 사영 T는 0과 1만을 고윳값으로 가짐.

2. 정사영

1) 정의

Definition 89. *내적공간 V와 사영 $T : V \to V$를 생각하자. $R(T)^{\perp} = N(T)$, $N(T)^{\perp} = R(T)$를 만족하는 T를 정사영(orthogonal projection)이라 함.*

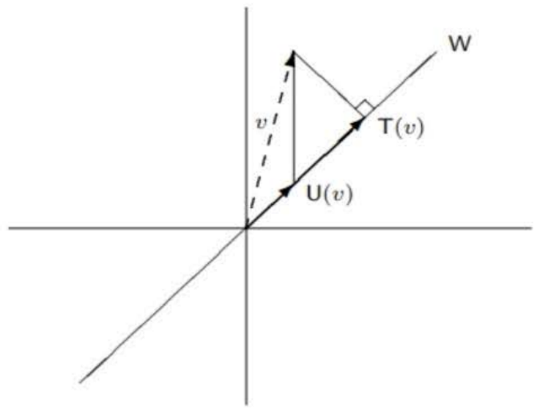
이는 전에 다룬 정사영과 동일한 개념임. 다만 여기서 더 깔끔하게 정의함.

Theorem 6.6에 의하면 어떤 부분공간에 대해 정사영은 유일하게 존재함.

정사영을 구체적으로 구할 때는 Theorem 6.6을 사용하자. Linear Extension Theorem에 의해, 기저에 대한 정사영을 구하는 방법으로 두 공간에 대해 정사영을 알아낼 수 있음.

2) 사영과의 차이

$V = W_1 \oplus W_2$, $V = W_1 \oplus W_3$라고 해서 $W_2 = W_3$인 것은 아니므로 치역(W_1)은 사영 T를 유일하게 결정하지 않음. 하지만 $R(T)^{\perp} = N(T)$, $N(T)^{\perp} = R(T)$라는 조건을 추가한 정사영은 치역에 의해 유일하게 결정됨.



3) 최적 근사성(approximation property)

내적공간 V와 그 부분공간 W에 대해서, W 위로의 정사영 $T : V \to V$를 생각하자. Theorem 6.6에 의하면, $v \in V$에 대해서 $T(v)$는 v의 W로의 최적근사임. 다시 말해, $T(v)$는 W위의 벡터들 중 v와 가장 가까운 벡터임.

4) 자기수반과 정사영의 판정

Theorem 6.24에 의하면, T가 정사영이기 위한 필요충분조건은 T가 자기수반인 사영인 것임.

3. 스펙트럼 정리(spectral Theorem)

1) 정의

Definition 90. *Theorem 6.25를 스펙트럼 정리(spectral Theorem)라 함.*

스펙트럼 정리에서 T의 고윳값으로 이루어진 집합 $\{\lambda_1, \lambda_2, \cdots, \lambda_k\}$를 T의 스펙트럼(spectrum)이라 하고, 4번의 합 $I = T_1 + T_2 + \cdots + T_k$를 T로 유도된 항등연산자 분해(resolution of the identity operator)이라 하며, 5번의 $T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k$를 T의 스펙트럼 분해(spectral decomposition)라 함.

고윳값의 배열 순서를 무시하면 T의 스펙트럼 분해는 유일함.

4. 관련 정리

1) 정사영과 자기수반 사영

Theorem 6.24 내적공간 V와 선형연산자 T를 생각하자. T가 정사영이기 위한 필요충분조건은 T의 수반연산자 T^*가 존재하고 $T^2 = T = T^*$가 성립하는 것임.

즉, T가 정사영이기 위한 필요충분조건은 T가 자기수반인 사영인 것임.

Proof. 1. T가 정사영 $\to$ T^* 존재, $T^2 = T = T^*$
T가 사영이므로 $T^2 = T$가 성립함. 즉, T^*의 존재성과 $T = T^*$만 보이면 됨.

$V = R(T) \oplus N(T)$, $R(T) = N(T)^\perp$, $x, y \in V$, $x_1, y_1 \in R(T)$, $x_2, y_2 \in N(T)$, $x = x_1 + x_2$, $y = y_1 + y_2$라 하자. 계산해보면 $\langle T(x), y\rangle = \langle x, T(y)\rangle$가 성립하므로 T^*가 존재하고 $T^* = T$임.

2. T^* 존재, $T^2 = T = T^*$ $\to$ T가 정사영
$T = T^*$이므로 T는 자기수반연산자이고, 복소/실내적공간에서 모두 OEB를 가짐. OEB를 β라고 하면 $[T]_\beta$와 $[T^2]_\beta$는 아래와 같음.

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \ [T^2]_\beta = ([T]_\beta)^2 = \begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix}$$

$[T]_\beta = [T^2]_\beta$이므로, $\lambda_1, \lambda_2, \cdots, \lambda_n$는 0 또는 1임. β의 원소들을 적절히 재배치한 것을 γ라 하면 $[T]_\gamma$는 아래와 같으므로 T는 정사영임.

$$[T]_\gamma = \begin{pmatrix} I & O \\ O & O \end{pmatrix}$$

□

2) 스펙트럼 정리

Theorem 6.25 유한차원 벡터공간 V의 선형연산자 T의 서로 다른 고윳값을 $\lambda_1, \lambda_2, \cdots, \lambda_k$라 하자. $F = C$이면 T가 정규연산자, $F = R$이면 T가 자기수반연산자라 가정함. (즉, T에 대한 OEB가 존재함.)

각 i $(1 \le i \le k)$에 대해서 고윳값 λ_i에 대응하는 고유공간을 W_i라 하자. W_i로의 V의 정사영을 T_i라 할 때 아래가 성립함.

1. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$
2. $j \neq i$에 대해서 부분공간 W_j의 직합을 W_i'라 표기하자. $W_i^\perp = W_i'$임.
3. $1 \le i, j \le k$에 대해서 $T_iT_j = \delta_{ij}T_i$임.
4. $I = T_1 + T_2 + \cdots + T_k$
5. $T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k$

Proof. 1. 직합에서 다뤘음. (Theorem 5.10)

2. W_i와 W_j의 각 벡터는 서로 직교하므로 당연함. (Theorem 6.15)

3. $i = j$인 경우, T_i가 정사영이므로 당연함. $i \neq j$인 경우 $W_i^\perp = W_j$이므로, $u \in W_i$, $w \in W_j$에 대해서 $T_i w = 0$, $T_j u = 0$임. 즉, $T_iT_j = 0$임.

4. $x_i \in W_i$, $x \in V$에 대해 $T_i(x) = x_i$임. 1번에 의해 $I(x) = x = x_1 + \cdots + x_k = T_1(x) + \cdots + T_k(x)$이므로 성립함.

5. $x \in V$를 $x_i \in W_i$에 대해 $x = x_1 + \cdots + x_k$라 하자. $T(x)$를 정리하면, $T(x) = T(x_1) + \cdots + T(x_k) = \lambda_1 x_1 + \cdots + \lambda_k x_k = \lambda_1 T_1(x) + \cdots + \lambda_k T_k(x) = (\lambda_1 T_1 + \cdots + \lambda_k T_k)(x)$임. □

아래 Corollary들에 대한 증명은 프리드버그 p.428 참고.

Corollary 1 $F = C$일 때 T가 정규연산자이기 위한 필요충분조건은 적절한 다항식 g에 대해서 $T^* = g(T)$인 것임.

Corollary 2 $F = C$일 때 T가 유니타리이기 위한 필요충분조건은 T가 정규연산자이고 T의 모든 고윳값 λ에 대해서 $|\lambda| = 1$인 것임.

Corollary 3 $F = C$일 때 T가 자기수바연산자이기 위한 필요충분조건은 T가 정규연산자이고 T의 모든 고윳값이 실수인 것임.

Corollary 4 T의 스펙트럼 분해가 $T = \lambda_1 T_1 + \cdots + \lambda_k T_k$이면 각 T_j는 T에 대한 다항식임.

7. 특잇값 분해와 유사역행렬

1. 특잇값 분해(SVD, singular value decomposition)

1) 정의(Theorem 6.26, 6.27)

Definition 91. *유한차원 내적공간 V, W, 랭크가 r인 선형변환 $T : V \to W$를 생각하자.*

-관찰-

선형변환 $T^*T : V \to V$, $TT^* : W \to W$를 생각하자.

1. T^*T와 TT^*는 자기수반연산자임.

2. T^*T와 TT^*는 양의 준정부호(positive semi-definite)임.
즉, T^*T와 TT^*는 랭크가 r이고, T^*T와 TT^*는 고윳값이 모두 0 이상의 실수임.

-정의-

(선형변환의 특잇값 정리. Theorem 6.26)

1. V의 정규직교기저 $\beta = \{v_1, \cdots, v_n\}$와 W, 정규직교기저 $\gamma = \{u_1, \cdots, u_m\}$, 양의 스칼라 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$이 아래를 만족하며 존재함.

$$T(v_i) = \begin{cases} \sigma_i u_i & (1 \leq i \leq r) \\ 0 & (i > r) \end{cases}$$

2. 1번이 성립할 때, $v_1, \cdots, v_n$은 T^*T의 고유벡터(OEB)가 되고, 각 고유벡터에 해당되는 고윳값은 아래와 같음. 이때, 특잇값은 T에 의해 유일하게 결정됨.

$$\begin{cases} \sigma_i^2 & (1 \leq i \leq r) \\ 0 & (r < i \leq n) \end{cases}$$

(행렬의 특잇값 분해 정리. Theorem 6.27)

3. $[T]_\beta^\gamma = A$($[T]_\beta^\gamma$가 아니라도 임의의 행렬에 대해 성립함), $\sigma_1, \cdots, \sigma_r$이 A의 고윳값이라고 할 때, $m \times n$ 행렬 Σ를 아래와 같이 정의하자.

$$\Sigma_{ij} = \begin{cases} \sigma_i & (i = j \leq r) \\ 0 & (else) \end{cases}$$

$A = U\Sigma V^*$를 만족하는 $m \times m$ 유니타리 행렬 U와 $n \times n$ 유니타리 행렬 V가 존재함. 이 분해를 A의 특잇값 분해(singular value decomposition)라 함.

모든 j에 대해서 j열이 u_j인 행렬을 U, 모든 j에 대해서 j열이 v_j인 행렬을 V라 하면 U와 V는 유니타리 행렬이고, $A = U\Sigma V^*$가 성립함.

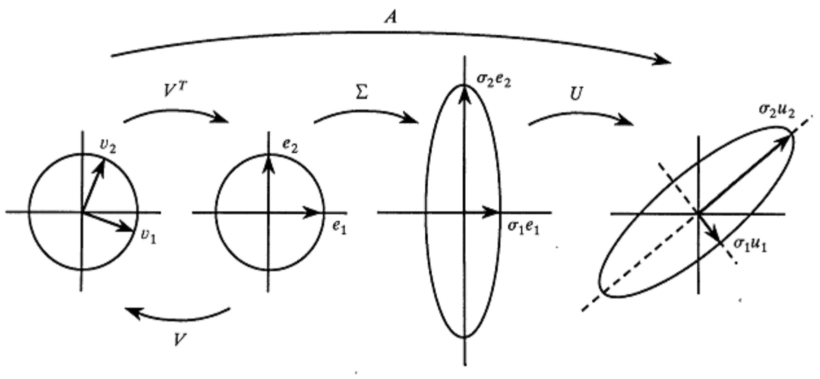
특잇값은 유일하게 결정되지만, 이때 등장하는 정규직교기저는 유일하게 결정되지 않음.

증명 과정에서도 알 수 있듯이, T와 T^*의 특잇값은 서로 같음. 또한, V와 W를 뒤집으면 T^*에 대한 정의로도 생각할 수 있음.

선형변환과 그 행렬표현의 특잇값은 같음.

정의의 1,2번은 선형변환에 대한 특잇값 정리임. 정의의 증명 과정을 보면, T^*T의 OEB와 그 고윳값으로 특잇값을 정의함. 그 특잇값은 V, W의 어떤 정규직교기저 쌍에 대해 1번 정의를 만족함. 1번에 의해 2번도 성립함. 이 증명 과정을 이해한다면 정의의 논리 관계를 알 수 있음.

$A = U\Sigma V^*$는 임의의 행렬(선형변환) A의 작용을 정사각행렬로 분해한 것으로 이해할 수 있음. U는 W의 기저 위치를 지정함. Σ는 특잇값(고윳값의 역할)으로서 크기를 지정(scaling)함. V는 V(내적공간)의 기저 위치를 지정함.



Proof. 1. '-관찰-'에 의해, T^*T는 고윳값 λ_i에 대응하는 고유벡터로 이루어진 OEB가 존재함. T^*T의 랭크가 r이므로 이때 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$이고(순서에 맞게 배치), $i > r$일 때 $\lambda_i = 0$임. $1 \leq i \leq r$에 대해서 아래와 같이 정의하자.

$$\sigma_i = \sqrt{\lambda_i},\ u_i = \frac{1}{\sigma_i}T(v_i)$$

위와 같이 정의했을 때 $\{u_1, \cdots, u_m\}$이 정의의 조건을 만족하는 W의 정규직교기저임을 보이면 됨. 우선 $\{u_1, \cdots, u_r\}$이 정규직교 부분집합임을 보이자. $1 \leq i, j \leq r$이라 하면 아래가 성립함.

$$\langle u_i, u_j\rangle = \langle \frac{1}{\sigma_i}T(v_i), \frac{1}{\sigma_i}T(v_j)\rangle = \frac{1}{\sigma_i\sigma_j}\langle T^*T(v_i), v_j\rangle = \frac{1}{\sigma_i\sigma_j}\langle \lambda_i v_i, v_j\rangle = \delta_{ij}$$

$\{u_1, \cdots, u_r\}$가 정규직교이므로 확장하여 정규직교기저 $\{u_1, \cdots, u_m\}$을 얻을 수 있음. $1 \leq i \leq r$일 때 $T(v_i) = \sigma_i u_i$임. $i > r$일 때 $T^*T(v_i) = 0$이므로 $T(v_i) = 0$임[53]. 즉, $\{u_1, \cdots, u_m\}$이 정의의 조건을 만족하는 W의 정규직교기저임.

2. 1번 정의가 성립한다고 가정하자. $1 \leq i \leq m$, $1 \leq j \leq n$에 대해서 $\langle T^*(u_i), v_j\rangle$가 아래와 같음. i, j가 r보다 커지면 $T^*(u_i) = 0$, $T(v_j) = 0$이 되므로 $i = j \leq r$이란 조건이 붙음.

$$\langle T^*(u_i), v_j\rangle = \langle u_i, T(v_j)\rangle = \langle u_i, \sigma_j u_j\rangle = \sigma_j\langle u_i, u_j\rangle = \begin{cases}\sigma_i & (i = j \leq r)\\ 0 & (else)\end{cases}$$

따라서 임의의 $1 \leq i \leq m$에 대해서 $T^*(u_i)$는 아래와 같이 나타낼 수 있음.

$$T^*(u_i) = \sum_{j=1}^{n}\langle T^*(u_i), v_j\rangle v_j = \begin{cases}\sigma_i v_i & (i = j \leq r)\\ 0 & (else)\end{cases}$$

$1 \leq i \leq r$일 때 $T * T(v_i) = T^*(\sigma_i u_i) = \sigma_i^2 v_i$이고, $r \leq i \leq n$일 때 $T * T(v_i) = T * (0) = 0$임. 즉, $1 \leq i \leq r$일 때 고윳값이 σ_i^2, $r \leq i \leq n$일 때 고윳값이 0임.

T에 의해 T^*T가 결정되고, σ_i^2는 T^*T의 고윳값이므로 T^*T에 의해 유일하게 결정됨. 즉, 특잇값은 T에 의해 유일하게 결정됨.

3. Theorem 2.13에 의해, AV의 j열은 $Av_j = \sigma_j u_j$임. Σ의 j열은 $\sigma_j e_j$이므로 $U\Sigma$의 j열은 $U(\sigma_j e_j) = \sigma_j U e_j = \sigma_j u_j$임. 즉, $AV = U\Sigma$이고 정리하면 $A = U\Sigma V^*$임. □

[53] 6.3절 연습문제 15(d) 참고.

2) 특잇값(singular value)

> **Definition 92.** *위에서 정의한 유일한 스칼라 $\sigma_1, \cdots, \sigma_r$를 선형변환 T의 특잇값(singular value)이라 함.*
>
> *$m \times n$ 행렬 A에 대해서 선형변환 L_A의 특잇값을 A의 특잇값(singular value)이라 함.*

정의에 의하면 특잇값은 T에 의해 유일하게 결정됨.

m과 n의 최솟값을 k라 하고, $r < m$, $r < n$이면 특잇값은 $\sigma_{r+1} = \cdots = \sigma_k = 0$으로까지 확장할 수 있음.

3) 선형변환의 특잇값/정규직교기저 찾기

내적공간 V, W에 대해서 $T : V \to W$의 특잇값과 그 정규직교기저를 찾는 방법은 아래와 같음.

행렬로 계산하는 것이 간단함.

1. V, W의 정규직교기저에 대해서 T의 행렬표현 A를 얻음.
2. A^*A의 고윳값을 구함.
3. 고윳값에 대한 고유벡터를 찾아서 OEB를 만듦. 이게 $\{v_1, \cdots, v_n\}$이고 그 고윳값으로 특잇값을 알 수 있음.
4. $v_1, \cdots, v_n$과 특잇값을 구했으므로 단순 대입해서 대응하는 u_i를 알 수 있음.[54]

4) 행렬에 특잇값 분해 적용하기

$m \times n$ 행렬 A에 특잇값 분해를 적용하는 방법은 아래와 같음.

1. A^*A의 고윳값과 그에 해당하는 고유벡터(OEB)를 찾음. $\{v_1, \cdots, v_n\}$과 특잇값을 알 수 있음.

2-1. T의 랭크와 W의 차원이 같으면, 단순 대입해서 $\{u_1, \cdots, u_m\}$을 알 수 있음.
2-2. T의 랭크와 W의 차원이 다르면, AA^*에 대한 OEB를 직접 찾아줘야 함.

3. U, Σ, V^*를 구성함.

A^*A에서 OEB를 얻어 V를 구성한 것과 같은 맥락으로, AA^*에서 OEB를 얻어 U를 구성할 수 있음. A^*에 대한 특잇값 분해를 생각하면 당연함.

v_i와 u_i의 관계는 T의 랭크(r)만큼만 존재하므로, $m = r$[55]이면 $\{v_1, \cdots, v_i\}$만으로 $\{u_1, \cdots, u_r\}$(전체)을 구할 수 있음. $m \neq r$이면 TT^*에 대해서 계산해줘야 함. n, m과 r을 비교해서 T에 대해서 구할지 T^*에 대해서 구할지 결정할 수 있음.

[54] W의 차원이 T의 랭크보다 클 수도 있으므로 W의 기저를 일반적으로 구하는 방법은 아님. W의 기저는 AA^*에 대해서 위의 과정을 적용해 얻을 수 있음.

[55] dimension theorem에 의해 $r > m$인 경우는 존재하지 않음.

2. 극분해(polar decomposition)

1) 정의

Definition 93. *임의의 정사각행렬 A에 대해서, $A = WP$를 만족하는 유니타리 행렬 W와 양의 준정부호 행렬 P가 존재함. 또한, A가 가역이면 이 표현을 유일함.*

A를 W와 P로 분해하는 것을 A의 극분해(polar decomposition)라 함.

복소수를 크기가 1인 복소수와 음이 아닌 실수, 즉 $z = re^{i\theta}$로 분해하는 것과 유사한 맥락의 분해임. 크기가 1인 복소수의 역할을 유니타리 행렬이, 음이 아닌 실수를 양의 준정부호 행렬을 사용한 것.

Proof. 행렬 A에 대해서 특잇값 분해를 하면 $A = U\Sigma V^*$임. $W = UV^*$, $P = V\Sigma V^*$라 하면 아래가 성립함.

$$A = U\Sigma V^* = UV^*V\Sigma V^* = WP$$

W는 유니타리 행렬의 곱이므로 유니타리 행렬임. P는 양의 준정부호 행렬인 Σ와 유니타리 동치이므로 양의 준정부호 행렬임.[56]

가역성에 대한 증명은 프리드버그 p.438 참고. □

2) 극분해 방법

그 증명에서도 알 수 있듯이, 임의의 정사각행렬을 극분해하는 방법은 특잇값 분해 이후 $W = UV^*$, $P = V\Sigma V^*$를 구하는 것임.

3. 유사역변환

1) 정의

Definition 94. *같은 체 위의 유한차원 내적공간 V, W, 선형변환 $T : V \to W$를 생각하자. 선형변환 $L : N(T)^\perp \to R(T)$를 모든 $x \in N(T)^\perp$에 대해서 $L(x) = T(x)$라 정의하자. 아래의 조건을 만족하는 W에서 V로 가는 유일한 선형변환을 T의 유사역변환(pseudo-inverse) 또는 무어-펜로즈 유사역변환(Moore-Penrose generalized inverse)라 하고, $T^\dagger$라 표기함.*

$$T^\dagger = \begin{cases} L^{-1}(y) & (y \in R(T)) \\ 0 & (y \in R(T)^\perp) \end{cases}$$

가역이 아닌 선형변환에 대해서도, 가역이 되는 부분으로 선형변환을 제한한다면 역변환은 아니지만 역변환의 장점은 가지고 있는 선형변환을 생각할 수 있음. $L : N(T)^\perp \to R(T)$[57]에 대해서 $L(x) = T(x)$라고 정의하면 L은 가역이므로 L의 역변환을 이용함.

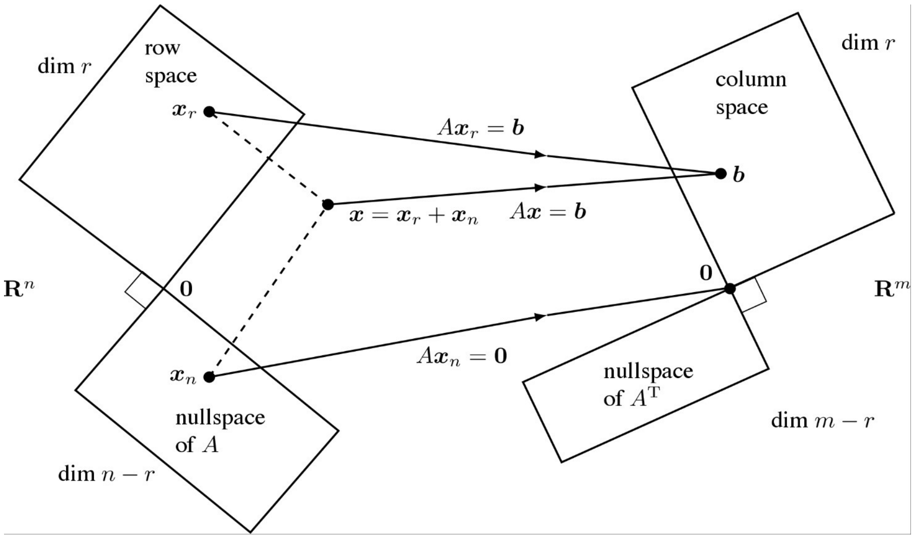
$L : N(T)^\perp \to W$에서 $N(T)^\perp$을 정의역으로 설정함으로서 공간(즉, 기저.)을 깔끔하게 분리할 수 있음.

T의 유사역변환은 T가 가역이 아닐 때에도 존재함. T가 가역이면 $T^\dagger = T^{-1}$임.

영변환의 유사역변환은 영변환임.

$T^\dagger$ 대신 T^+로 표기하는 경우도 많음.

[56] 6.5절 연습문제 14 참고.
[57] 직교집합은 일차독립이므로 $N(T)^\perp$으로 정의역을 지정한 것은 당연함.



2) 유사역변환 구하기

특잇값 정리를 이용하여 유사역변환을 구할 수 있음.

특잇값 정리의 정의, 유사역변환의 정의에서 나온 표기를 그대로 사용하자. $L^{-1}(u_i) = \frac{1}{\sigma_i}v_i$이므로 아래가 성립함은 당연함.

$$T^\dagger(u_i) = \begin{cases} \frac{1}{\sigma_i}v_i & (1 \le i \le r) \\ 0 & (r < i < m) \end{cases}$$

3. 유사역행렬

1) 정의

Definition 95. *$m \times n$ 행렬 A에 대해서 $(L_A)^\dagger : F^m \to F^n$이 좌측 곱 변환 L_B와 같도록 하는 $n \times m$ 행렬 B가 유일하게 존재함. 이 행렬 B를 A의 유사역행렬(pseudo-inverse)이라 하고, $B = A^\dagger$라 표기함. 즉, 아래가 성립함.*

$$(L_A)^\dagger = L_{A^\dagger}$$

2) 유사역행렬 구하기

행렬의 특잇값 분해를 안다면 그 행렬의 유사역행렬을 구하는 것은 굉장히 간단함.

유사역변환은 원래 선형변환에서 특잇값이 역수가 된 것으로 생각할 수 있음. 즉, 특잇값으로 원래 선형변환 특잇값의 역수를 적용하고, β, γ의 역할을 바꾸면 유사역행렬의 특잇값 분해를 쉽게 생각할 수 있음.

랭크가 r인 $m \times n$ 행렬 A의 특잇값 분해가 $A = U\Sigma V^*$이고, 영이 아닌 특잇값이 $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r$라 하자. $n \times m$ 행렬 $\Sigma^\dagger$를 아래와 같이 정의하자.

$$\Sigma^\dagger_{ij} = \begin{cases} \frac{1}{\sigma_i} & (i = j \le r) \\ 0 & (else) \end{cases}$$

$A^\dagger = V\Sigma^\dagger U^*$로 특잇값 분해를 할 수 있음. 또한, 정의한 $\Sigma^\dagger$는 실제로 Σ의 유사역행렬임.

3) 유사역행렬과 연립일차방정식

행렬로 표현된 연립일차방정식 $Ax = b$에 대해서 A가 가역인 경우 유일한 해 $A^{-1}b$가 존재함. A가 가역인 경우 $A^{-1}b = A^{\dagger}b$로 쓸 수 있는데, A가 가역이 아닌 경우에도 $A^{\dagger}b$는 존재함. Theorem 6.30에 의하면 $A^{\dagger}b$는 일반적으로 연립일차방정식 $Ax = b$와 관련이 있음.

Theorem 6.30에 의하면 $z = A^{\dagger}b$는 해가 존재하는 경우 노름이 가장 작은 유일한 해이고, 해가 존재하지 않는 경우 노름이 가장 작은 유일한 최적근사임. 즉, $z = A^{\dagger}b$는 해이거나 해에 가장 가까운 값임.

최소제곱법에서 등장한 Theorem 6.12에서 x_0이 $z = A^{\dagger}b$임.

4. 관련 정리

1) 유사역행렬과 연립일차방정식

Theorem 6.30 연립일차방정식 $Ax = b$를 생각하자. 이때 A는 $m \times n$ 행렬이고, $b \in F^m$임. $z = A^{\dagger}b$라 하면, z에 대해 아래의 성질이 성립함.

1. $Ax = b$에 모순이 없다면, z는 연립일차방정식의 노름이 가장 작은 유일한 해임.
2. $Ax = b$에 모순이 있다면, z는 노름이 가장 작은 유일한 최적근사임.

즉, 해가 존재하는 경우, 어떤 해를 y라 했을 때 $||z|| \leq ||y||$임(노름 최소). 이때 등호가 성립하기 위한 필요충분조건은 $z = y$인 것임(유일함.). 해가 존재하지 않는 경우, 임의의 $y \in F^n$에 대해서 $||Az - b|| \leq ||Ay - b||$임 (최적근사). 이때 등호가 성립하기 위한 필요충분조건은 $Az = Ay$임(유일함.). 또한 $Az = Ay$일 때 $||z|| < ||y||$가 성립함(노름 최소). 이때 등호가 성립하기 위한 필요충분조건은 $z = y$임(유일함.).[58]

Proof. 프리드버그 p.442 참고. □

[58] 최적근사로서도 유일하고, 노름이 가장 작은 것으로서도 유일하다는 것.

+ 기타

1. 순서쌍(tuple)

1) 정의

Definition 96. *$a_1, a_2, \cdots, a_n$이 체 F의 원소일 때, $(a_1, a_2, ..., a_n)$ 꼴의 수학적 대상을 F에서 성분을 가져온 n순서쌍(n-tuple)이라고 함.*

n순서쌍에서 $a_1, a_2, ..., a_n$을 n순서쌍의 성분(entry, component)이라고 함.

F에서 성분을 가져온 두 n순서쌍 $(a_1, a_2, ..., a_n)$과 $(b_1, b_2, ..., b_n)$은 $a_i = b_i$일 때 같다(equal)고 함.

2. 다항식(polynomial)

1) 정의

Definition 97. *계수가 체 F의 원소인 다항식은 $f(x) = a_n x_n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$로 정의함.*

이때 n은 음이 아닌 정수임.

각 a_k를 x_k의 계수(coefficient)라고 함.

$f(x) = 0$이면 이를 0 다항식(zero coefficient)라고 함.
0 다항식의 차수는 편의를 위해 -1로 정의함.

다항식의 차수(degree)는 계수가 0이 아닌 항의 x의 지수 중 가장 큰 값으로 정의함.

각 항이 전부 일치하는 두 다항식은 같다고 함.

F가 무한집합일 때, F에서 계수를 가져온 다항식을 F에서 F로 가는 함수로 볼 수 있음.

다항함수 $f(x) = a_n x_n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$는 간단히 $f(x)$또는 f로 씀.

3. 수열(sequence)

1) 정의

Definition 98. *체 F에서 정의된 수열은 자연수 집합을 정의역, F를 공역으로 하는 함수임.*

$\sigma(n) = a_n \ (n = 1, 2, 3, \cdots)$인 수열 σ는 (a_n)이라 표기하기도 함.

4. 정리

정리(Theorem) : 증명을 통해 참임이 밝혀진 명제.

보조정리(Lemma) : 증명된 명제로서 다른 결과를 증명하는 데 주로 사용되는 명제.

따름정리(Corollary) : 정리가 증명되었을 때, 그것으로부터 파생되어 나오는 명제.

공리(Axiom) : 증명할 필요 없이 자명한 진리이자 다른 명제들을 증명할 때 전제로 사용되는 기본적인 가정.

5. 행렬(matrix)

1) 정의

Definition 99. *F에서 성분을 가져온 $m \times n$ 행렬은 아래와 같은 직사각형 모양의 배열임.*

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

$a_{ij}\ (1 \le i \le m, 1 \le j \le n)$ $i = j$에 대해서 $i = j$인 성분을 대각성분(diagonal entry)라 함.

성분 $a_{i1}, a_{i2}, \cdots, a_{in}$은 이 행렬의 i행(row)이라 함.
행렬의 각 행은 F^n의 벡터로 나타낼 수 있음.

성분 $a_{1j}, a_{2j}, \cdots, a_{mj}$은 이 행렬의 j열(column)이라 함.
행렬의 각 열은 F^m의 벡터로 나타낼 수 있음.

두 행렬에 대해 대응하는 성분이 모두 일치할 때, 두 행렬을 같다고 함.

2) 영행렬(zero matrix)

Definition 100. *모든 성분이 0인 행렬을 영행렬(zero matrix)이라 하고 O로 표기함.*

3) 정사각행렬(정방행렬, square matrix)

Definition 101. *행의 개수와 열의 개수가 같은 행렬을 정사각행렬(정방행렬, square matrix)이라 함.*

4) 전치행렬(transpose matrix)

Definition 102. *$m \times n$ 행렬 A의 행과 열을 바꾸어 얻은 행렬을 A의 전치행렬(transpose matrix)라 하고, A^t로 표기함.*

$(A_t)_{ij} = A_{ji}$임.

임의의 두 행렬 A, B와 스칼라 a, b에 대해서, $(aA + bB)^t = aA^t + bB^t$임.
임의의 행렬 A와 스칼라 a에 대해서, $(aA)^t = aA^t$임.

임의의 두 행렬 A, B에 대해서, $(AB)^t = B^tA^t$임.

5) 대칭행렬(symmetric matrix)

Definition 103. *$A^t = A$인 행렬.*

대칭행렬이려면 정사각행렬이어야 함.

6) 상삼각행렬(위삼각행렬, upper triangular matrix)

Definition 104. *대각성분 아래의 모든 성분이 0인 행렬. 즉, $i > j$일 때 $A_{ij} = 0$인 행렬.*

7) 대각행렬(diagonal matrix)

Definition 105. *대각성분을 제외한 모든 성분이 0인 행렬.*

8) 크로네거 델타(Kronecker delta)와 항등행렬(identity matrix)

크로네커 델타(Kronecker delta)는 아래과 같이 정의함.

Definition 106. *$i = j$일 때 $\delta_{ij} = 1$이고, $i \neq j$일 때, $\delta_{ij} = 0$*

항등행렬(identity matrix)는 아래와 같이 정의함.

Definition 107. *$n \times n$ 항등행렬 I_n의 성분은 $(I_n)_{ij} = \delta_{ij}$임.*
즉, 항등행렬은 정사각행렬 중 대각성분은 전부 1이고 나머지는 0인 행렬임.

가리키는 것이 명확하면 n을 생략하여 I라 표기하기도 함.

항등행렬은 $M_{n \times n}(F)$에서 곱셈에 대한 항등원임.
즉, 행렬 A에 항등행렬을 곱한 결과는 A임.

6. 집합(set)

1) 집합의 합(sum)

Definition 108. *공집합이 아닌 S_1과 S_2는 벡터공간 V의 부분집합임. 두 집합의 합(sum) $S_1 + S_2$는 아래과 같이 정의함.*

$$x + y : x \in S_1, y \in S_2$$

2) 직합(direct sum)

Definition 109. *벡터공간 V와 부분공간 W_1, W_2에 대해서 $W_1 \cap W_2 = \{0\}$이고 $W_1 + W_2 = V$이면 V는 W_1과 W_2의 직합(direct sum)이라 하고 $V = W_1 \oplus W_2$라 표기함.*