2장 부록

보충설명: 가산집합과 비가산집합 (Countable and Uncountable Sets)

가산집합 (Countable Set)

> **정의2A-1.1**

집합 S가 **가산집합(countable set)**이라는 것은 S의 원소들이 자연수 집합 $\mathbb{N}$과 **일대일 대응(one-to-one correspondence)**을 가질 수 있다는 의미이다. 즉, S의 모든 원소를

$$s_1, s_2, s_3, s_4, \dots \tag{33}$$

와 같이 순서대로 배열할 수 있을 때, S는 가산집합이다.

- 집합이 **유한**하면 가산집합이다.
- 집합이 **무한**하지만 자연수 집합과 같은 크기(즉, 가부번집합, countably infinite)라면 가산집합이다.

예시

1. 자연수 집합 $\mathbb{N} = \{1, 2, 3, 4, \dots\}$ 은 자기 자신과 일대일 대응이 가능하므로 가산집합이다.
2. 정수 집합 $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ 도 가산집합이다. 그 이유는 다음과 같이 자연수와 일대일 대응이 가능하기 때문이다.

$$0, 1, -1, 2, -2, 3, -3, \dots \tag{34}$$

3. 유리수 집합 $\mathbb{Q}$도 가산집합이다. 유리수는 두 정수의 비율로 표현되므로, 분자와 분모를 자연수의 순서쌍으로 정리하면 자연수와 일대일 대응이 가능하다.

비가산집합 (Uncountable Set)

> **정의2A-1.2**

집합 S가 **비가산집합(uncountable set)**이라는 것은 **자연수 집합과 일대일 대응을 만들 수 없는 무한집합**이라는 의미이다. 즉, S의 원소를 $\mathbb{N}$의 원소와 순서대로 정리할 수 없다.

예시

1. 실수 집합 $\mathbb{R}$은 비가산집합이다.
2. 닫힌구간 $[0, 1] \subset \mathbb{R}$ 도 비가산집합이다.
3. **칸토르 집합(Cantor Set)**과 같은 일부 집합도 비가산집합이 될 수 있다.

칸토르의 대각선 논법 (Cantor's Diagonal Argument)**

게오르크 칸토르(Georg Cantor)는 **대각선 논법(diagonal argument)**을 사용하여 실수 집합 $\mathbb{R}$이 가산집합이 아님을 증명했다.

대각선 논법의 개요: 이진수 표현을 이용한 대각선 논법

- 가정: $(0, 1)$ 사이의 실수를 이진수로 표현하고, 이를 가산집합으로 나열할 수 있다고 가정하자.

$$
\begin{aligned}
&0.\mathbf{0}1100101010\ldots \\
&0.1\mathbf{0}101010011\ldots \\
&0.11\mathbf{0}00101100\ldots \\
&0.110\mathbf{1}1011010\ldots \\
&0.1101\mathbf{0}011011\ldots \\
&0.11010\mathbf{1}01101\ldots
\end{aligned}
\tag{35}
$$

- **대각선 숫자**: 위에서 **대각선 방향**으로 숫자를 읽으면 $000101\ldots$ 가 된다.
- **새로운 숫자**: 이 숫자의 각 자릿수를 바꿔서 $111010\ldots$ 를 만들면, 이 숫자는 원래 나열된 리스트에 포함되지 않는다.
- 따라서, 이 목록이 실수 전체를 포함할 수 없으며, 실수 집합은 비가산임이 증명된다.

칸토르의 대각선 논법은 수학에서 **비가산집합이 존재함을 보이는 강력한 도구**로, 이후 **괴델의 불완전성 정리(Gödel's incompleteness theorem)**나 **튜링의 결정문제(Entscheidungsproblem)** 증명에도 사용되었다.

보조정리 2-2.2. 증명

함수 $f : \mathbb{R}^n \to \mathbb{R}$ 가 **볼록(convex)** 하다고 하는 것은 임의의 두 점 x_1, x_2 및 $0 \leq \lambda \leq 1$ 에 대해 다음의 **젠슨의 부등식(Jensen's inequality)** 을 만족할 때를 의미한다:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{36}$$

확률론의 맥락에서, 젠슨의 부등식은 일반적으로 다음과 같이 주어진다. 즉, X 가 확률 변수이고, φ 가 볼록 함수일 때,

$$\varphi\left(\mathbb{E}[X]\right) \leq \mathbb{E}\left[\varphi(X)\right]. \tag{37}$$

보조정리 2A.1 [영의 곱셈 부등식(Young's Inequality for Products)]

a, b 를 0 이상의 실수라 하고, p, q 를 1보다 큰 실수라고 하자. 만약 다음 관계식이 성립한다면:

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{38}$$

다음의 부등식이 성립한다:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{39}$$

이 부등식에서 등호가 성립하는 경우는 $a^p = b^q$ 일 때이며, 그 경우에만 등호가 성립한다.

증명 (Proof)

이 주장은 $a = 0$ 또는 $b = 0$ 일 때 자명하게 성립한다. 따라서, 이제 $a > 0$ 그리고 $b > 0$ 인 경우를 고려하자.

$t = \frac{1}{p}$, 그리고 $1 - t = \frac{1}{q}$ 라 두자. 그러면 로그 함수가 오목(concave)하다는 사실에 의해

$$\ln(ta^p + (1-t)b^q) \geq t\ln(a^p) + (1-t)\ln(b^q) = \ln(a) + \ln(b) = \ln(ab) \tag{40}$$

이 성립하며, 등호는 $a^p = b^q$ 일 때만 성립한다. Young의 부등식은 양변에 지수를 취하여 얻어진다.

Young의 곱셈 부등식을 사용하면 다음과 같은 식을 얻을 수 있다:

$$\left|\frac{x_i y_i}{\|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q}\right| \leq \frac{|x_i|^p}{p\|\boldsymbol{x}\|_p^p} + \frac{|y_i|^q}{q\|\boldsymbol{y}\|_q^q}. \tag{41}$$

양변을 합하면

$$\left\|\frac{\boldsymbol{x}^T \boldsymbol{y}}{\|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q}\right\|_1 = \sum_i \left|\frac{x_i y_i}{\|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q}\right| \leq \frac{\sum_i |x_i|^p}{p\|\boldsymbol{x}\|_p^p} + \frac{\sum_i |y_i|^q}{q\|\boldsymbol{y}\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1. \tag{42}$$

이를 통해 주장을 증명하였다. ¶

보충 설명: 2.3절

m개의 방정식과 n개의 미지수를 가질때의 해(The Solution of m Equations in n Unknowns)

m개의 방정식과 n개의 미지수를 가지는 선형 연립방정식 $\boldsymbol{Ax} = \boldsymbol{b}$ (여기서 $A \in \Re^{m\times n}$)에 대해 $m < n$인 경우 거의 대부분(almost all) 무한히 많은 해를 가진다. 좀 더 자세한 예로, 3개의 방정식과 4개의 미지수를 가지는 선형연립방정식을 고려하자.

$$Ax = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 4 & 3 & 6 & -5 \\ -1 & -2 & -4 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 1 \end{bmatrix} = b \tag{43}$$

가우시안 소거법을 활용하면,

$$\begin{aligned} [A;b] &= \begin{bmatrix} 1 & 1 & 2 & -1 & -1 \\ 4 & 3 & 6 & -5 & -4 \\ -1 & -2 & -4 & 0 & 1 \end{bmatrix} \\ \xrightarrow[R_2-(4)R_1]{} \ E_{21}(4)[A;b] &= \begin{bmatrix} 1 & 1 & 2 & -1 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & -4 & 0 & 1 \end{bmatrix} \\ \xrightarrow[R_3-(-1)R_1]{} \ E_{31}(-1)E_{21}(4)[A;b] &= \begin{bmatrix} 1 & 1 & 2 & -1 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{bmatrix} \\ \xrightarrow[R_3-(1)R_2]{} \ E_{32}(1)E_{31}(-1)E_{21}(4)[A;b] &= \begin{bmatrix} 1 & 1 & 2 & -1 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{44}$$

다음과 같이 동등하지만 풀기 쉬운 선형시스템인 **행사다리꼴(Row echelon form)** 행렬 U를 얻게 된다.

$$
\begin{aligned}
Ax &= \begin{bmatrix} 1 & 1 & 2 & -1 \\ 4 & 3 & 6 & -5 \\ -1 & -2 & -4 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 1 \end{bmatrix} = b \\
\Longrightarrow Ux &= \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & -1 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = c
\end{aligned} \tag{45}
$$

여기서 U를 A로 되돌리는 행렬은 다음과 같이 정방 하삼각(lower traingular) 행렬이 된다.

$$
L = E_{21}(-4)E_{31}(1)E_{32}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \tag{46}
$$

이 소거 과정은 행렬이 다음과 같이 더 간단한 **기약행사다리꼴(reduced row echelon form,rref)**가 될 때까지 계속 할 수 있다.

$$
R = \begin{bmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{47}
$$

전형적인 행사다리꼴 행렬의 0이 아닌 성분과 전형적인 **기약행사다리꼴 (reduced row echelon form)**의 모양은 다음과 같다.

$$
U = \begin{bmatrix} \otimes & * & * & * & * & * & * & * & * \\ 0 & \otimes & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \otimes & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \otimes \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & * & 0 & * & * & * & * & 0 \\ 0 & 1 & * & 0 & * & * & * & * & 0 \\ 0 & 0 & 0 & 1 & * & * & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{48}
$$

행끼리의 교환이나 행을 나누는 것을 포함한 기본 연산(elementary operation)의 선택에 상관없이, 언제나 마지막 A의 rref는 언제나 동일하다. 기본 변수(위의 예에서는 u, w)는 피벗이 있는 열에 대응되며, 자유 변수들은 피벗이 없는 행들에 대응된다. 여기서 U 또는 R 의 형태를 통해 선형독립인 열벡터와 선형독립인 행벡터의 갯수는 모두 피벗의 갯수와 같게 되고 이는 행렬의 **랭크(rank)**와 같다.

일반적으로 소거법이 $A\boldsymbol{x} = \boldsymbol{b}$ 를 $U\boldsymbol{x} = \boldsymbol{c}$로 줄인다고 하자. r개의 피벗이 존재할 때, U의 마지막 $m - r$ 개의 행은 0이다. c의 마지막 $m - r$ 개 요소가 0일 때만 해가 존재한다. 만약 $r = m$이면, 언제나 해가 존재한다. 여기서 r은 행렬A의 **랭크(rank)**라고 한다. 이는 행렬에서 실제로 독립된 행의 수를 의미한다.

동차시스템(A homogeneous system) $A\boldsymbol{x} = 0$ 이 방정식보다 많은 미지수를 갖는다면 ($n > m$), 거의 대부분(almost all) 자명하지 않은(nontrivial)한 해를 갖는다: 즉, 자명한(trivial) 해 $\boldsymbol{x} = 0$ 이외에 다른 해가 존재하고 모든 동차해의 집합을 A의 영공간(null space)이라 한다. 예를들면 위 예에서 동차해는 다음과 같이 주어진다.

$$\boldsymbol{x} = v\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \tag{49}$$

만약 $r = n$이고 자유 변수가 없으면, 동차해는 유일하게 $\boldsymbol{x} = \mathbf{0}$ 만을 갖는다.

따라서 앞절 예에서 일반해는 다음과 같다.

$$x = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \in \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \mathcal{N}(A) \tag{50}$$

각주