통계공부와 관련된 글들

Seoncheol Park

2019-08-14

차 례

일러두기

John W. Tukey는 이렇게 말했다.

“통계학은 과학이라는 것이 내 견해다. 통계학은 더이상 수학이나 물리학, 화학 또는 경제학의 한 부류가 아니다.”

각 장은 독립된 구성으로 되어 있으며, 한 권 이상의 책들을 참고문헌으로 하여 그들의 정의 및 표현을 따라가는 방식으로 구성되어 있다. 따라서 각 장마다 표현 및 한국어 용어 번역이 상이할 수 있다.

install.packages(c(“wavethresh”, “splines2”, “astsa”, “xts”, “quantspec”, “TSA”, “kza”, “itsmr”, “dismo”, “waveslim”, “geoR”, “SpatialEpi”, “spBayes”, “fields”, “RandomFields”, “SpatialExtremes”, “spatstat”, “ggplot2”, “splancs”, “evd”, “plotrix”, “geoRglm”, “bookdown”, “svglite”, “ape”))

편 I

Basic Concepts

1 기본적인 수학 개념들

이 장에서는 앞으로 다룰 내용을 이해하기 위해 필요한 기본적인 수학 개념을 정리하였다.

1.1 집합론(set theory)

1.1.1 카디널리티(cardinality)

카디널리티(cardinality)는 집합의 원소의 갯수를 세기 위해 도입되었다. 유한집합에서는 원소의 갯수를 세는 것이 어렵지 않지만, 무한집합의 경우는 갯수를 세는 것이 문제가 될 수 있다. 집합 A가 주어졌을 때, 그것의 카디널리티를 $|A|$로 쓰도록 하자. 자연수 집합의 카디널리티는 특별히 $|\mathbb{N}| = \aleph_0$로 쓰며 **알레프-널(aleph-null)**로 부른다.

Definition 1.1.1 (카디널리티가 같다)**.** 두 집합 A, B 사에 전단사(bijection, 일대일 대응) 관계가 성립할 때, 두 집합의 카디널리티가 같다고 정의하고, $|A| = |B|$로 표기한다. □

1.2 극한(limit)

이 부분은 (**?**)의 1장을 참고하였다.

Definition 1.2.1 (수열의 수렴)**.** $\{x_n\}_{n=1}^{\infty}$를 실수들의 수열이라고 하고 $x \in \mathbb{R}$을 실수라고 하자. 그러면 모든 $\epsilon > 0$에 대해 $n_\epsilon \in \mathbb{N}$이 존재해 모든 $n \geq n_\epsilon$에 대해 $|x_n - x| < \epsilon$일 때 x_n은 $n \to \infty$함에 따라 x에 **수렴(converge)**한다고 말하며, 다음과 같이 쓴다.

$$\lim_{n \to \infty} x_n = x.$$

□

Definition 1.2.2 (상계와 하계). $\{x_n\}_{n=1}^{\infty}$를 실수들의 수열이라고 하자. 만약 $x_n \le u, \forall n \in \mathbb{N}$을 만족하는 실수 $u \in \mathbb{R}$이 존재할 경우 u를 수열 $\{x_n\}_{n=1}^{\infty}$의 **상계(upper bound)**라고 한다. 마찬가지로 $x_n \ge l, \forall n \in \mathbb{N}$을 만족하는 실수 $l \in \mathbb{R}$을 수열 $\{x_n\}_{n=1}^{\infty}$의 **하계(lower bound)**라고 한다. □

Definition 1.2.3 (최소상계와 최대하계). $\{x_n\}_{n=1}^{\infty}$를 실수들의 수열이라고 하자. 수열 $\{x_n\}_{n=1}^{\infty}$의 **최소상계(supremum)** u_l은 상계들 중 가장 작은 상계, 즉 모든 다른 상계들 u에 대해 $u_l \le u$인 상계이며,

$$u_l = \sup_{n \in \mathbb{N}} x_n$$

으로 쓴다. 마찬가지로 수열 $\{x_n\}_{n=1}^{\infty}$의 **최대하계(infimum)** l_u는 하계들 중 가장 큰 하계, 즉 모든 다른 하계들 l에 대해 $l_u \ge l$인 하계이며,

$$l_u = \inf_{n \in \mathbb{N}} x_n$$

으로 쓴다. □

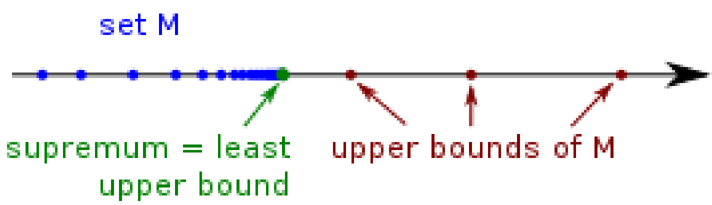


그림 1.1: 최소상계.

우리는 최소상계와 최대하계의 극한 또한 생각해 볼 수 있다.

Definition 1.2.4 (상극한과 하극한 그리고 극한). $\{x_n\}_{n=1}^{\infty}$를 실수들의 수열이라고 하자. 수열 $\{x_n\}_{n=1}^{\infty}$의 **상극한(limit supremum)**은

$$\lim_{n \to \infty} \sup x_n = \inf_{n \in \mathbb{N}} \sup_{k \ge n} x_k$$

이다. 마찬가지로 수열 $\{x_n\}_{n=1}^{\infty}$의 **하극한(limit infimum)**은

$$\lim_{n \to \infty} \inf x_n = \sup_{n \in \mathbb{N}} \inf_{k \ge n} x_k$$

로 정의한다. 만약

$$\lim_{n \to \infty} \sup x_n = \lim_{n \to \infty} \inf x_n = c \in \mathbb{R}$$

일 경우 c를 수열 $\{x_n\}_{n=1}^{\infty}$의 **극한(limit)**이라고 한다. □

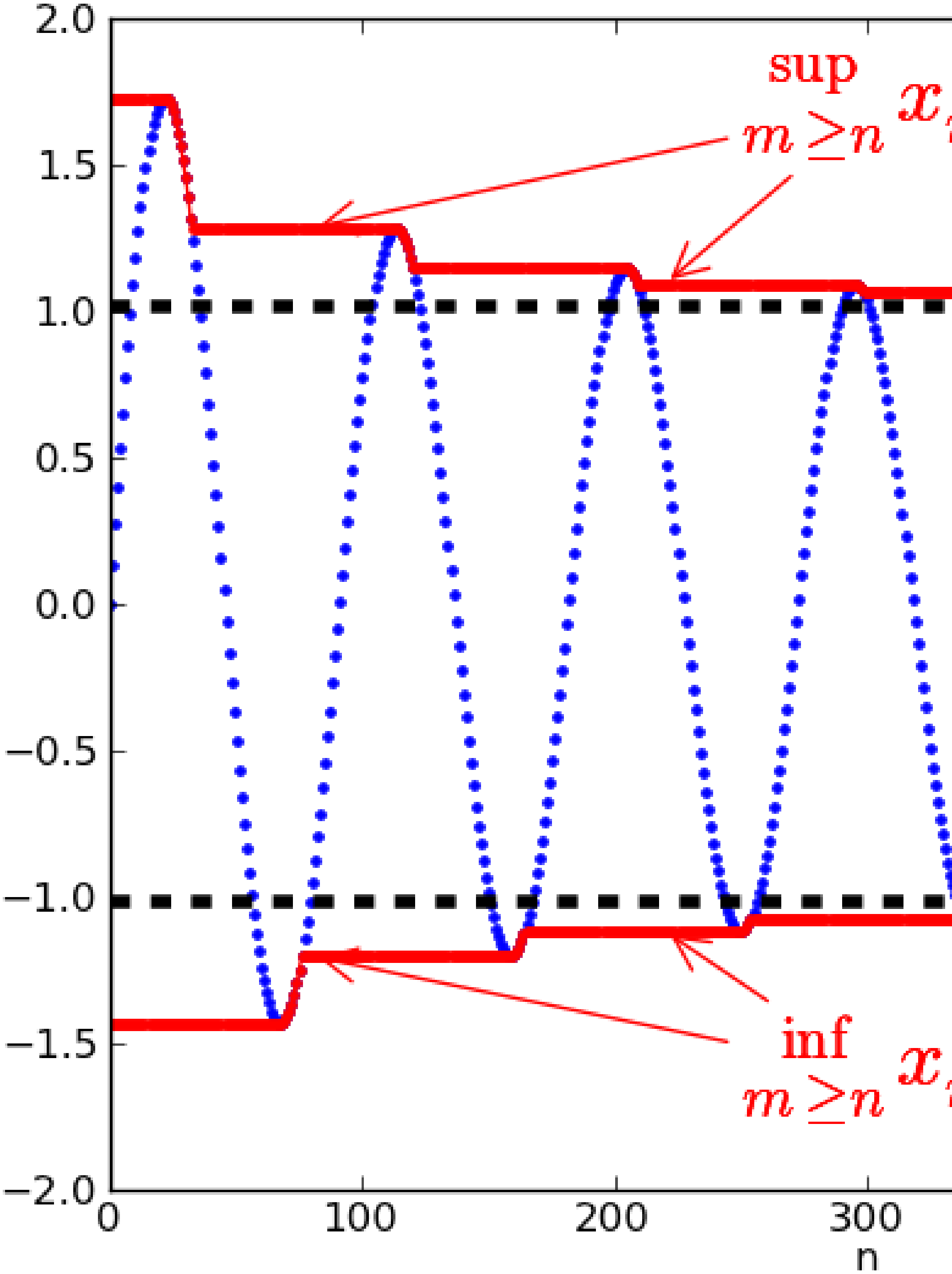


그림 1.2: 상극한과 하극한.

1.3 연산자들과 노름(operators and norms)

1.3.1 직합(direct sum)

Definition 1.3.1 (직합)**.** 크기 $m \times n$인 행렬 $\mathbf{A}$와 $p \times q$인 행렬 $\mathbf{B}$가 있을 때 이들의 **직합(direct sum)**은

$$\mathbf{A} \oplus \mathbf{B} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{p1} & \cdots & b_{pq} \end{bmatrix}$$

□

Example 1.3.1 (직합의 예)**.**

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{bmatrix} \oplus \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 & 0 & 0 \\ 2 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

□

1.3.2 크로네커 곱(Kronecker product)

Definition 1.3.2 (크로네커 곱)**.** 크기 $m \times n$인 행렬 $\mathbf{A}$와 $p \times q$인 행렬 $\mathbf{B}$가 있을 때 이들의 **크로네커 곱(Kronecker product)**은

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

□

Example 1.3.2 (크로네커 곱의 예)**.** 다음은 크로네커 곱의 한 예이다.

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} \otimes \begin{bmatrix}0 & 5\\ 6 & 7\end{bmatrix} = \begin{bmatrix}1\cdot 0 & 1\cdot 5 & 2\cdot 0 & 2\cdot 5\\ 1\cdot 6 & 1\cdot 7 & 2\cdot 6 & 2\cdot 7\\ 3\cdot 0 & 3\cdot 5 & 4\cdot 0 & 4\cdot 5\\ 3\cdot 6 & 3\cdot 7 & 4\cdot 6 & 4\cdot 7\end{bmatrix} = \begin{bmatrix}0 & 5 & 0 & 10\\ 6 & 7 & 12 & 14\\ 0 & 5 & 0 & 20\\ 18 & 21 & 24 & 28\end{bmatrix}$$

□

1.3.3 크로네커 거듭곱(Kroneckerian power)

We call p^k-vector $\mathbf{a}^{\otimes k}$, the k-th power of the p-vector $\mathbf{a}$, if $\mathbf{a}^{\otimes 0} = 1$ and

$$\mathbf{a}^{\otimes k} = \mathbf{a} \otimes \cdots \otimes \mathbf{a}(\text{ k times}).$$

In general, for any matrix $\mathbf{A}$, the Kroneckerian power is given by

$$\mathbf{A}^{\otimes k} = \mathbf{A} \otimes \cdots \otimes \mathbf{A}(\text{ k times}).$$

Furthermore, $\mathbf{A}^{\otimes k}\mathbf{B}^{\otimes k} = (\mathbf{AB})^{\otimes k}$. In particular, it is noted that

$$\mathbf{a}^{\otimes k} \otimes \mathbf{a}^{\otimes j} = \mathbf{a}^{\otimes (k+j)}, \qquad k, j \in \mathbb{N}.$$

The following statement makes it possible to identify where in a long vector of the Kroneckerian power a certain element of the product is situated.

1.3.4 텐서곱(tensor product)

이 부분은 (?)의 10.5절을 참고하였다. 우리가 covariance operator를 다룰 때 텐서곱과 텐서 공간을 생각하는 것이 편리할 때가 있다고 한다. 특히 무한차원이나 일반적인 힐버트 공간에서는 텐서곱의 활용이 절대적이다.

다음과 같은 행렬 $\mathbf{A} \in \mathbb{R}^{N\times M}$을 생각해보자. 이 행렬은 세 가지 방법으로 바라볼 수 있다.

1. 전통적인 행렬
2. $\mathbb{R}^N \to \mathbb{R}^M$으로 가는 선형 변환: 벡터의 왼쪽이나 오른쪽에 행렬을 곱하는 형태 (유클리드 공간에서의 모든 선형 변환을 떠올려보자)

3. $\mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$인 겹선형(bilinear) functional: 이차 형식에서 많이 다루는 형태이다. 다음과 같이 두 유클리드 공간 $\mathcal{H}_1 = \mathbb{R}^N, \mathcal{H}_2 = \mathbb{R}^M$가 있다고 하자. 두 개의 벡터 $x_1 \in \mathbb{R}^N$, $x_2 \in \mathbb{R}^M$에 **텐서곱(tensor product)**를 적용하면 다음과 같은 행렬을 유도한다.

$$x_1 \otimes x_2 := x_1 x_2^T.$$

또한 공간 $\mathbb{R}^N \otimes \mathbb{R}^M$은 위와 같은 원소들의 모든 유한한 선형 결합의 공간으로 정의할 수 있다. 다시 말하면 만약 $\mathbf{A} \in \mathbb{R}^N \otimes \mathbb{R}^M \to \mathbb{R}$이면 J개의 원소들 $x_{1j} \in \mathbb{R}^N$, $x_{2j} \in \mathbb{R}^M$이 존재해

$$\mathbf{A} = \sum_{j=1}^{J} x_{1j} \otimes x_{2j} = \sum_{j=1}^{J} x_{1j} x_{2j}^T$$

를 만족한다. 어떤 $N \times M$ 행렬도 이런 표현 방법으로 표현할 수 있으므로 $\mathbb{R}^N \otimes \mathbb{R}^M = \mathbb{R}^{N \times M}$으로 표현할 수 있다. 이제 이 세번째 관점을 다른 공간에 대해서도 특정화해보자. 어떤 $N \times M$ 행렬 $\mathbf{A}$는 다음과 같은

$$\mathbf{A}(y_1, y_2) := y_1^T \mathbf{A} y_2$$

인 겹선형 mapping을 이끌어낸다. 따라서 공간 $\mathbb{R}^N \otimes \mathbb{R}^M$은 모든 겹선형 map $\mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$들의 집합으로 identified될 수 있다. 오브젝트 $x_1 \otimes x_2$는 또한 다음과 같은 겹선형 mapping

$$(x_1 \otimes x_2)(y_1, y_2) = y_1^T x_1 x_2^T y_2 = (y_1^T x_1)(y_2^T x_2)$$

로 identified될 수 있다.

이러한 관점에서 텐서를 바라보자. 텐서는 스칼라들의 체가 연관된 공간에서의 cartesian 곱으로부터 오는 다중선형사상(multilinear map)이다. 만약 $\{e_{1j}\}$와 $\{e_{2j}\}$가 $\mathbb{R}^N$과 $\mathbb{R}^M$의 정규 기저들이라고 하자. 그러면 $\{e_{1j} \otimes e_{2j}\}$는 $\mathbb{R}^N \times \mathbb{R}^M$의 기저가 된다. 이런 방식으로 우리는 이 기저를 전통적인 행렬에도 적용할 수 있다.

$$\mathbf{A}(e_{1j}^T, e_{2k}) = e_{1j}^Y \mathbf{A} e_{2k} = \mathbf{A}(j, k).$$

Definition 1.3.3 (힐버트 공간에서의 텝서곱)**.** $x_1 \in \mathcal{H}_1$과 $x_2 \in \mathcal{H}_2$를 두 실 힐버트 공간에서의 원소들이라고 하자. 그러면 텐서곱 $x_1 \otimes x_2 : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{R}$은 겹선형 사상으로 어떤 $(y_1, y_2) \times \mathcal{H}_1 \times \mathcal{H}_2$에 대해

$$(x_1 \otimes x_2)(y_1, y_2) = \langle x_1, y_1 \rangle_{\mathcal{H}_1} \langle x_2, y_2 \rangle_{\mathcal{H}_2}$$

로 정의된다. □

Example 1.3.3 (행렬의 텐서곱 예시)**.**

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 2 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 6 & 4 & 4 & 8 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \otimes \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 & 3 & 4 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 & 3 & 4 \\ 1 & 2 & 1 & 2 & 2 & 4 \\ 3 & 4 & 3 & 4 & 6 & 8 \end{bmatrix}.$$

□

1.3.5 데카르트 곱(Cartesian product)

두 개의 집합 A, B가 있을 때, 이들의 **데카르트 곱(Cartesian product)** $A \times B$는

$$A \times B = \{(a, b) | a \in A \text{ and } b \in B\}$$

로 정의된다.

1.3.6 노름(norm)

1.3.6.1 벡터 노름(vector norm)

Definition 1.3.4 (노름과 노름공간)**.** 벡터공간 X에서 다음 세 조건들이 만족되면 함수 $\|\cdot\|$을 **노름(norm)**이라 하고 또한 벡터공간 X를 **노름공간(normed space)**라 한다.

1. 임의의 $\mathbf{x} \in X$, where $\|\mathbf{x}\| \geq 0$이며 $\|\mathbf{x}\| = 0$이기 위한 필요충분조건은 $\mathbf{x} = \mathbf{0}$이다.

2. 임의의 $\mathbf{x}, \mathbf{y} \in X$에 대해

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

가 성립한다.

3. 임의의 스칼라 α와 임의의 $\mathbf{x} \in X$에 대해

$$\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$$

가 성립한다.

□

1.3.6.2 행렬 노름(matrix norm)

1.4 행렬의 분해(matrix decomposition)

1.4.1 고유값 분해(eigenvalue decomposition)

고유값 분해는 행렬 A가 $n \times n$ 정방행렬일 때만 적용 가능하다.

1.4.2 스펙트럼 분해(spectral decomposition)

$p \times p$ 대칭행렬 A에 대한 **스펙트럼 분해(spectral decomposition)**는 다음과 같다. $p \times p$ 대칭행렬 A는 직교행렬 P에 의해 **대각화(diagonalization)**된다고 한다.

$$A = P\Lambda P^T = \sum_{i=1}^{p} \lambda_i e_i e_i^T.$$

이때 $PP^T = P^T P = I$를 만족하는 직교행렬 P는 $P = [e_1, \dots, e_p]$로 이루어지며, Λ는 A의 고유값(eigenvalue)들로만 이루어진 대각행렬(diagonal matrix)

$$\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{bmatrix}$$

이다. 대각행렬 Λ는 $P^T AP = \Lambda$이다.

1.4.3 특이값분해(SVD)

특이값분해(singular value decomposition, SVD)는 $m \times n$ 직사각형 행렬 A에 대해 스펙트럼 분해를 일반화한 것이다. A의 특이값 분해는 다음과 같다.

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

이 때

- U: A의 left singular vector로 이루어진 $m \times m$ 직교행렬 (orthogonal matrix)
- Σ: 주 대각성분이 $\sqrt{\lambda_i}$로 이루어진 $m \times n$ 직사각 대각행렬 (diagonal matrix)
- V: A의 right singular vector로 이루어진 $n \times n$ 직교행렬 (orthogonal matrix)

행렬 A의 계수 (rank)가 k라고 할 때,

- $U = [u_1, \dots, u_k, \dots u_m]$는 AA^T를 고유값분해 (eigenvalue decomposition)로 직교대각화하여 얻은 $m \times m$ 직교행렬 (orthogonal matrix)이며, 특히 $[u_1, \dots, u_k]$를 **좌특이벡터 (left signular vector)**라고 한다.
- $V = [v_1, \dots, v_k, \dots, v_n]$는 $A^T A$를 고유값분해로 직교대각화하여 얻은 $n \times n$ 직교행렬이며, 특히 $[v_1, v_2, \dots, v_k]$를 **우특이벡터 (right signular vector)**라고 한다.
- Σ는 $A^T A$의 0이 아닌 고유값이 $\lambda_1, \lambda_2, \dots, \lambda_k$일 때 $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$를 대각성분으로 가지고 나머지 성분을 0으로 갖는 $m \times n$ 직사각 **대각행렬 (diagonal matrix)**이다.

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & \sqrt{\lambda_k} & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

즉 A를 다시 쓰면

$$A = \begin{bmatrix} u_1 & \cdots & u_k & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & \sqrt{\lambda_k} & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ \vdots \\ V_k^T \\ \vdots \\ V_n^T \end{bmatrix}$$

이다. 위 식에서 **특이값(singular value)**는 $\sigma_i^2 = \lambda_i$ 로부터 $\sigma_i = \sqrt{\lambda_i}$가 된다. 참고로 U, V가 직교행렬이면 $UU^T = I, VV^T = I$가 성립한다.

1.4.4 특이값분해와 고유값분해의 관계(the relationship between spectral decomposition and eigenvalue decomposition)

$m \times n$ 행렬 A의 특이값분해의 U는 AA^T의 고유벡터이고, V는 A^TA의 고유벡터이며, A의 0이 아닌 특이값들의 제곱 $\Sigma\Sigma^T, \Sigma^T\Sigma$는 AA^T, A^TA의 고유값과 같음을 알 수 있다. 참고로 $\sigma_i = \sqrt{\lambda_i}$ 이므로 $\Sigma\Sigma^T$ 또는 $\Sigma^T\Sigma = \lambda_i$이다.

$$\begin{aligned} U &= AA^T \\ &= (U\Sigma V^T)(U\Sigma V^T)^T \\ &= (U\Sigma V^T)(V\Sigma^T U^T) \\ &= U(\Sigma\Sigma^T)U^T \end{aligned}$$

$$\begin{aligned} V &= A^TA \\ &= (U\Sigma V^T)^T(U\Sigma V^T) \\ &= (V\Sigma^T U^T)(U\Sigma V^T) \\ &= V(\Sigma^T\Sigma)V^T \end{aligned}$$

즉 $u_1, \ldots, u_k, \ldots u_m$ 는 range(A) 의 직교정규벡터, $v_1, \ldots, v_k, \ldots, v_n$ 는 $\mathcal{N}(A)^\perp$ 의 직교정규벡터이다.

1.4.5 특이값 분해에 대한 추가 설명 (additional explanation about spectral decomposition)

여기서는 Boyd 교수의 강의노트[1]를 참고하였다. 잠시 편의를 위해 U 는 $m \times n$ 행렬, Σ는 $n \times n$ 행렬이라고 하자. 그러면 특이값 분해는

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

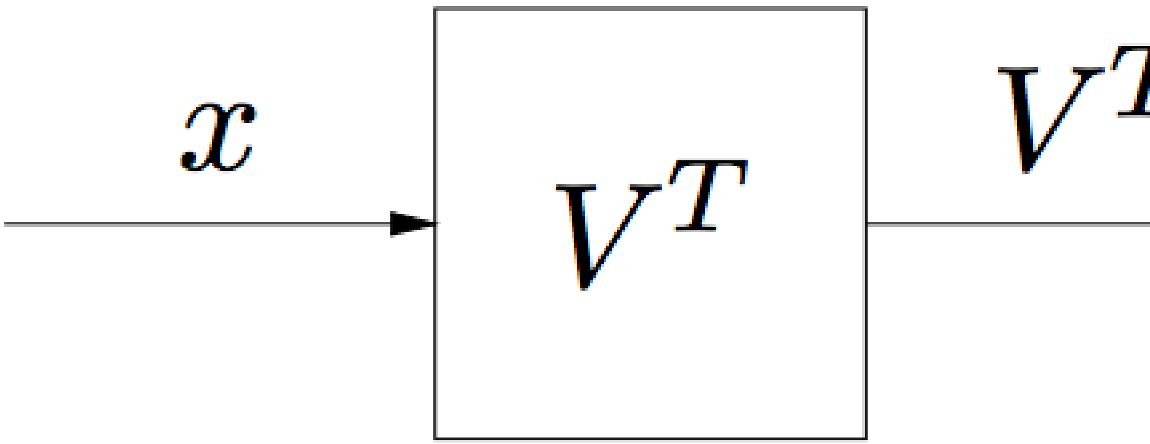


그림 1.3: SVD 그림.

가 된다. 선형 사상 (mapping) $y = Ax$는 다음과 같이 분해할 수 있다.

- x를 input direction들 $v_1, \ldots, v_n$ 을 따라 계수들을 계산한다.

[1] https://web.stanford.edu/class/archive/ee/ee263/ee263.1082/notes/ee263coursereader.pdf

- σ_i는 척도계수
- 이것들을 다시 output directions $u_1, \ldots, u_n$을 따라 재구성한다.

대칭행렬 A에 대한 고유값 분해와 달라지는 점은 input direction들과 output direction들이 다르다는 것이다.

- v_1는 input direction으로 가장 민감하다.(most sensitive, highest gain)
- u_1은 output direction으로 가장 민감하다.
- $Av_1 = \sigma_1 u_1$이다.

Example 1.4.1 (SVD의 기하학적 의미 예)**.** $A = \mathbb{R}^{2\times 2}$이며 $\Sigma = \mathrm{diag}(1, 0.5)$인 경우를 생각해보자. 이 경우 x를 v_1, v_2를 따라 풀면 $v_1^T x = 0.5, v_2^T x = 0.6$ 즉 $x = 0.5v_1 + 0.6v_2$이며, $Ax = (v_1^T x)\sigma_1 u_1 + (v_2^T x)\sigma_2 u_2 = (0.5)(1)u_1 + (0.6)(0.5)u_2$ 이다. □

1.4.6 특이값 분해의 기하학적 의미(geometrical meaning of spectral decomposition)

위키피디아[2]를 참고하자.

1.5 기저(basis)

1.5.1 리츠 기저(Riesz basis)

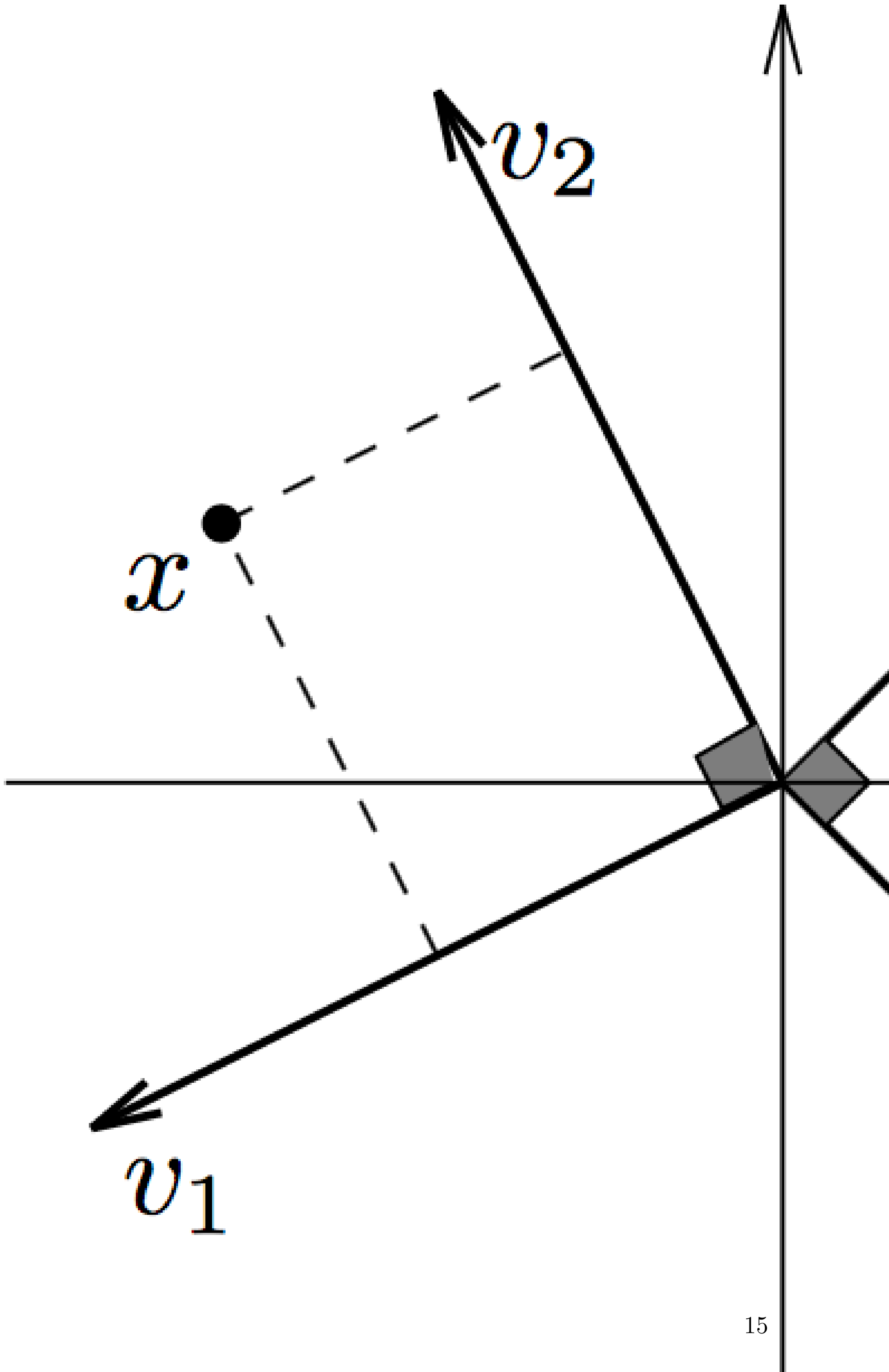
In the Hilbert space $L_2[0,1]$, an unconditional basis is called a **Riesz basis** if it is "almost normalized". This means that there exist real, positive, non-zero consts m and M so that

$$0 < m \leq \|\phi_i\| \leq M < \infty.$$

A Riesz basis is characterized by two Riesz constants A and B, so that for all $f = \sum_i s_i \phi_i \in L_2[0,1]$,

$$A^2\|f\|^2 \leq \sum_{i\in\mathbb{Z}} s_i^2 \leq B^2\|f\|^2.$$

[2] https://en.wikipedia.org/wiki/Singular_value_decomposition


v_2
x
v_1

(Jensen의 Noise reduction and wavelet thresholding으로부터)

There exists $\phi_0(x) \in \mathcal{V}_1$ such that $\{\phi_0(x-k)|k \in \mathcal{Z}\}$ forms a Riesz basis of $\mathcal{V}_1$, i.e, there exists $0 < A \leq B < \infty$ such that

$$A\|c_k\|^2 \leq \|\sum_k c_k\phi_0(x-k)\|^2 \leq B\|c_k\|^2$$

for all $\{c_k\} \in l^2$, where A and B do not depend on the c_k

(동익이형 박사논문 47쪽)

1.5.2 Radial basis function

Radial funtion이란 거리에만 의존하는 함수를 의미한다. 어떤 함수에 대한 근사 모델을 radial function의 선형조합으로 표현할 수 있다.

- Gaussian

$$\phi(r) = e^{-(\epsilon r)^2}$$

- Multiquadric

$$\phi(r) = \sqrt{1+(\epsilon r)^2}$$

- Inverse quadratic

$$\phi(r) = \frac{1}{1+(\epsilon r)^2}$$

- Inverse multiquadric

$$\phi(r) = \frac{1}{\sqrt{1+(\epsilon r)^2}}$$

1.6 공간(space)

이 부분은 전체적으로 (**?**)의 정의와 내용들을 따라간다.

1.6.1 군과 장 (groups and fields)

수학에서 **군 (group)**은 연산과 함께 정의되는 원소들의 집합으로 원소들이 closure, associativity, identity and invertability를 만족해야 한다. 대충 말해서 **장 (field)**은 합, 차, 곱, 몫의 개념을 갖고 있는 대수적 구조이다.

1.6.2 벡터 공간 (vector spaces)

벡터 공간은 스칼라들의 장 (field) 에서 정의된다. 이런 스칼라들은 실수가 될 수도 있고, 복소수가 될 수도 있다. 벡터 공간은 덧셈과 곱셈을 적용할 수 있는 수학적 실재 (mathematical entity) 들의 집합이다. 예를 들면 숫자 공간 $\mathbb{R}$, $\mathbb{C}$ 등은 벡터 공간이 된다. 숫자 뿐만 아니라 함수, 다항식, 선형 연산자, 벡터 등등이 수학적 실재가 될 수 있다.

벡터 공간 (vector space) V 는 다음 공리들을 만족하는 원소들 $\mathbf{x}$ (벡터라 부른다) 의 collection이다.

1. V 는 덧셈 하에서 교환 가능한 가환군 (commutative group, 교환 법칙이 성립하는 군) 이다.
   i. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x} \in V$ for any $\mathbf{x}, \mathbf{y} \in V$ (closure)
   ii. $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z} \in V$ for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ (associativity)
   iii. 다음과 같은 $\mathbf{0}$ 벡터 (항등원) 가 존재해 모든 $\mathbf{x} \in V$ 에 대해 $\mathbf{x}+\mathbf{0}=\mathbf{x}$를 만족한다. (identity)
   iv. 다음과 같은 덧셈에 대한 역원 $-\mathbf{x}$이 존재해 $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$이 된다. (invertibility)
2. V 는 어떤 장 $\mathbb{F}$ 에 대해 다음의 공리를 만족하며, 이런 원소들을 **스칼라 (scalar)**라고 부른다.
   i. V 는 스칼라 곱에 대해 닫혀 있다.

$$\alpha\mathbf{x} \in V \text{ for arbitrary } \mathbf{x} \in V, \alpha \in \mathbb{F}.$$

   ii. 스칼라 곱은 V 와 $\mathbb{F}$ 의 원소들에 대해 **분배가능 (distributive)**하다.

$$\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y} \in V, (\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{y} \in V.$$

   iii. 스칼라 곱은 결합법칙이 성립한다. $\alpha(\beta\mathbf{x})=\beta(\alpha\mathbf{x})$.
   iv. Zero scalar $0 \in \mathbb{F}$ 에 대한 곱은 항등원 $0\mathbf{x}=\mathbf{0} \in V$ 을 생산한다.

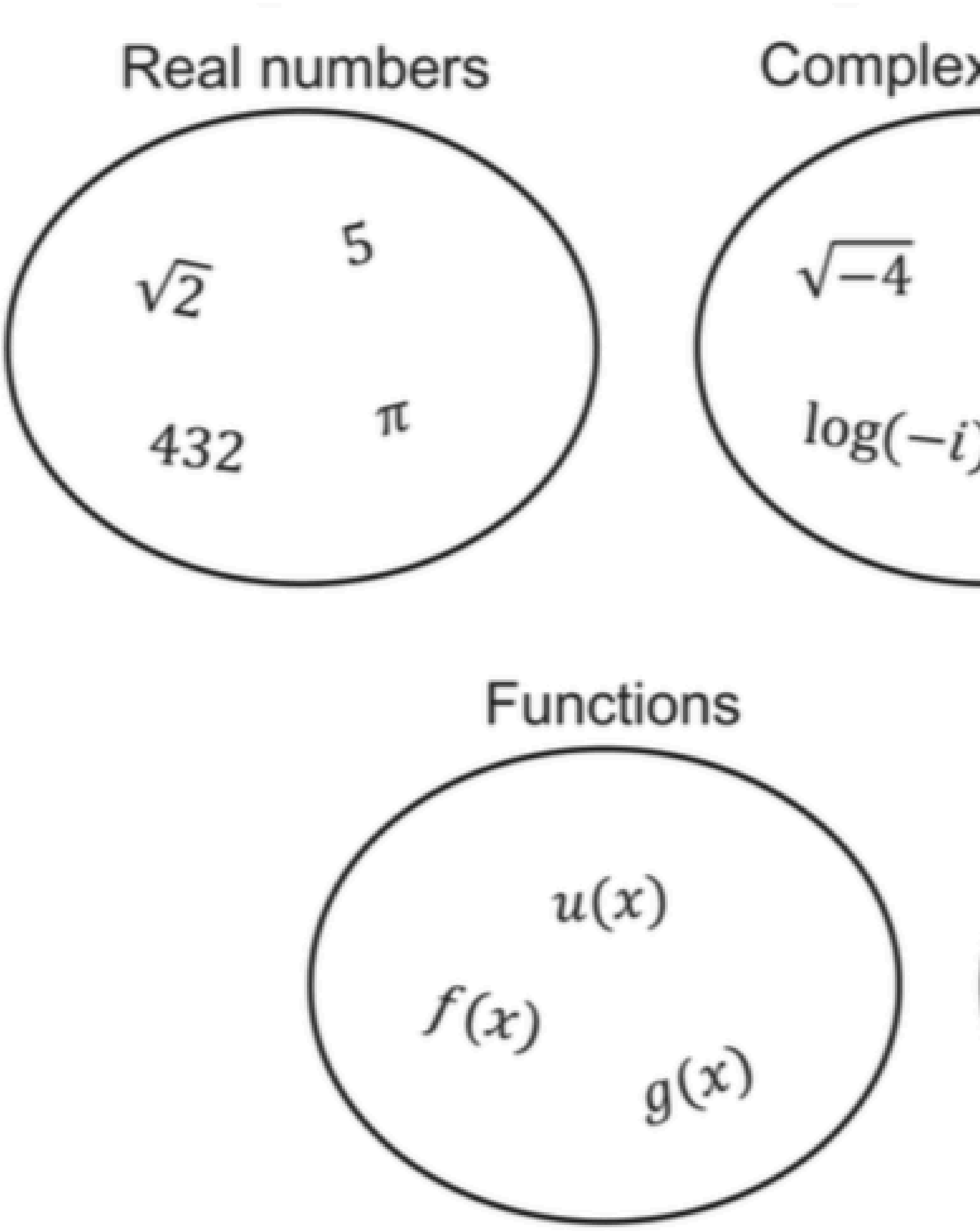


그림 1.5: 벡터 공간은 다양한 종류의 수학적 실재(mathematical entity)로부터 만들어진다.

v. Unit scalar $1 \in \mathbb{F}$ 는 $1\mathbf{x} = \mathbf{x}$ 성질을 갖는다.

여기서부터는 벡터 공간의 주요 성질들에 대해 다룬다.

1.6.3 내적 (inner products)

Definition 1.6.1 (내적)**.** 내적은 $\mathbf{x}$, $\mathbf{y}$ 의 순서쌍을 스칼라 (좀 더 일반적으로는 복소수) 에 mapping하는 것이고 $\langle \mathbf{x}, \mathbf{y} \rangle$ 로 쓴다. 이 mapping은 다음 규칙들을 만족해야 한다.

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$. (여기서 별표는 complex conjugate를 의미한다.)

2. $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha^* \langle \mathbf{x}, \mathbf{z} \rangle + \beta^* \langle \mathbf{y}, \mathbf{z} \rangle$ (여기서 α, β 는 어떤 복소수이다.)

3. 모든 $\mathbf{x}$ 에 대해 $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ 이다.

4. $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ 은 $\mathbf{x} = \mathbf{0}$ 과 필요충분조건이다.

□

특별히 내적을 갖고 있는 벡터 공간을 **내적 공간 (inner product space)** 이라고 부른다.

내적 공간에서는 우리는 항상 **내적 노름 (inner product norm)** 을 다음과 같이 정의할 수 있다.

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Proposition 1.6.2 (내적 공간의 노름과 관련된 성질들)**.** 내적 공간은 그 공간 상에서 노름 $\|\cdot\|$ 을 정의하며 다음의 설질들을 갖는다.

- $\|ax\| = |a|\|x\|$
- $|\langle x, y \rangle| \leq \|x\|\|y\|$ *(Cauchy-Schwarz inequality)*
- $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality)*
- $d(x, y) = \|x - y\|$ *is a metric (distance), i.e.,* $d(x, y) = d(y, x) > 0$ *if* $x \neq y$, $d(x, x) = 0$ *and* $d(x, y) \leq d(x, z) + d(z, y)$.

□

1.6.4 직교성(orthogonality)

다음은 (?)책에 있는 직교성의 정의다.

Definition 1.6.3 (직교)**.** 내적 공간에 있는 두 원소들 $\mathbf{x}, \mathbf{y}$는 $\langle x\mathbf{x}, \mathbf{y}\rangle = 0$일 때 **직교 (orthogonal)**한다고 말한다. □

1.6.4.1 완비 벡터 공간(complete vector spaces)

벡터 공간이 유한차원일 때, 공간의 완비성(completeness)은 같은 공간의 다른 larger orthonormal set에 포한되어 있지 않는 orthonormal set을 찾음으로써 증명할 수 있다. 예를 들면 3차원 공간에서는 linear combination이 그 공간의 모든 벡터를 표현할 수 있는 three orthonormal vector의 set을 찾기만 하면 되는 것이다. 그러나 우리가 무한 차원 공간을 고려할 때, 무한한 숫자의 orthonormal vector를 고려하는 것은 쉽지 않다. 사실 무한한 숫자의 vector의 linear combination은 때대로 같은 공간에 포함된 벡터를 표현하는 데 충분치 않을 수도 있다. 이러한 모호함을 해결하기 위해 무한 차원 공간에서 완비성을 생각하는 것이다.

Definition 1.6.4 (벡터의 Cauchy sequence)**.** 벡터의 수열 $\{\mathbf{x}_1, \mathbf{x}_2, ...\}$가 모든 $\epsilon > 0$에 대해 적단한 근사 숫자 N이 존재해 모든 $m, n > N$에 대해 $\|\mathbf{x}_m - \mathbf{x}_n\| < \epsilon$을 만족한다면 이를 벡터의 **코시 수열(Cauchy sequence)**이라고 한다. 쉽게 얘기하자면 $\mathbf{x}_m$과 $\mathbf{x}_n$이 $m, n \to \infty$ 함체 따라 가까어지는 수열을 코시 수열이라 부르는 것이다. □

Definition 1.6.5 (코시 수열의 수렴)**.** 벡터의 무한 수열 $\{\mathbf{x}_1, \mathbf{x}_2, ...\}$가 있을 때, 만약 $\mathbf{x}$가 존재해 $\|\mathbf{x}_n - \mathbf{x}\| \to 0$을 만족한다면 이 수열이 수렴(convergent)한다고 한다. □

그러나 그 역은 성립하지 않는다. 코시 수열이라고 해서 다 수렴하는 건 아니다. 완비성을 정의해 이를 해결할 수 있는 것이다.

Definition 1.6.6 (벡터 공간의 완비성)**.** 만약 어떤 벡터 공간의 모든 코시 수열이 수렴한다면, 우리는 그 공간을 **완비(complete)**라고 부른다. □

다음은 (?)의 10장에 나오는 예이다.

Example 1.6.1 (유클리드 공간에서의 완비성)**.** 유클리드 거리 하에서 개구간 $(0, 1)$은 complete metric space가 아니지만, 닫힌 구간 $[0, 1]$은 complete metric space가 된다. 유클리드 공간 $\mathbb{R}^d$는 complete하다. □