

# 분석적 사고

전종준  
서울시립대학교 통계학과

September 10, 2018

## Abstract

키워드: 선형대수, 미적분학, 회귀분석, 탐색적 자료분석

# 1 자료와 선형대수

## 2 벡터와 벡터공간

숙제 제출: 여러 함수의 정의

### 2.1 자료와 벡터

#### 자료의 표현 1

- 한 사람의 혈압:  $x$
- 한 사람의 나이와 혈압:  $(x, y)$
- 한 사람의 나이와 혈압과 신장과,..... (총 100가지의 변수를 고려하면 이를 어떻게 표현할 것인가?)

#### 자료의 표현 2

- 한 사람의 성별
- 한 사람의 혈액형
- 한 사람의 국적

#### 자료의 연산 1

- 두 사람의 혈압:  $x$
- 두 사람의 나이와 혈압:  $(x, y)$
- 두 사람의 나이와 혈압과 신장과,..... (총 100가지의 변수를 고려하면 이를 어떻게 표현할 것인가?)
- 두 사람의 나이와 혈압과 신장과,..... 의 평균: 연산을 어떻게 할 것인가?
- 필요한 연산을 정의해 보시오.

#### 자료의 연산

- 열 사람의 성별과 비율
- 열 사람의 혈액형과 비율
- 열 사람의 국적과 비율

### 벡터 공간

- 연산이 정의되어 있고 (addition, scalar product)
- 연산의 항등원 역원이 정의되어 있고
- 연산 순서에 대한 법칙이 정해져 있음 (교환, 결합, 분배)
- 이런 규칙이 정해진 시스템에서는 평균이라는 값을 논리적 오류 없이 정의할 수 있음.

### R에서 벡터 연산하기

- R 다운로드/ Rstudio 설치 하기
- 다음 예제 따라하기
  - 벡터의 생성
  - 벡터의 연산 ( 더하기, 스칼라 곱하기)
  - (예제를 만들어서) 네 사람의 정보에 대해 평균구하기

### 자료의 차원과 공간

- $\mathbf{X} = (X_1, X_2) \in \mathbb{R}^2$  ( $X_1$ : 키,  $X_2$ : 체중)
- $\mathbf{X} = (X_1, X_2, X_3) \in \mathbb{R}^3$  ( $X_3 = X_1 - X_2$ )

\* 위 두가지 예에서 자료의 특성에 대해서 이야기 해보자.

\* 위 두가지 예에서 자료의 속한 공간에 대해서 이야기 해보자 (숫자 혹은 그림)

- $\mathcal{X} = \{\mathbf{X} \in \mathbb{R}^2 : \mathbf{X} = \beta_1(1, 0) + \beta_2(0, 1), \beta_1, \beta_2 \in \mathbb{R}\}$

$\mathbf{X} = \beta_1(1, 0) + \beta_2(0, 1)$  과 다른방식으로도  $\mathcal{X}$ 의 구성이 가능한가?

### 기저

- 좌표의 표현: 어떤 점  $(1, 2)$ 을 나타내는 방식을 이야기 해보자. 어떤 사람은  $x$ 축으로 1만큼  $y$ 축 방향으로 2만큼 움직였다고 한다. 또 다른 사람은  $45^\circ$  방향으로  $\sqrt{2}$  만큼 움직인후에  $y$  축 방향으로 1만큼 움직였다고 한다. 여러가지 방법으로  $(1, 2)$ 로 가는 방법을 생각해보자.
- 기저와 방향:
- 3차원에서 기저를 생각해보자
- $p$  차원의 좌표와 기저를 생각해보자.
- $p$  차원 좌표를 하나의  $p$  차원 데이터라고 하고 예를 들어보자. 일반적으로  $p > 1$  인 경우 이를 다변량자료(multivariate data)라고 부른다.

**생각해볼 문제** Due date: 17, Sep

- $\mathcal{D} = \{1, 2, 3, 4, 5\}$  에서  $\mathcal{R} = \{0, 1\}$  으로 가는 함수  $f$ 가 데이터라고 하자. 이 데이터는 몇 차원 데이터인가?
- $\mathcal{D} = \{1, 2, 3, 4, 5\}$  에서  $\mathbb{R}$  로 가는 함수를 데이터라고 하자. 이 데이터는 몇 차원 데이터인가?
- $f: \mathbb{R} \mapsto \mathbb{R}$  을 데이터라고 하자. 이 데이터는 몇 차원 데이터인가?
- $f(x) = ax + b$  ( $a, b$  는 상수)라고 하고 이 함수를 데이터라고 하자. 이 데이터는 몇 차원 데이터인가?

### 3 행렬과 데이터