MACHINE LEARNING | 기계 학습

1.1.2 지식기반 방식에서 기계 학습으로의 대전환

- 인공지능의 주도권 전환
  - 지식기반 → 기계 학습
  - 기계 학습: 데이터 중심 접근방식

그림 1-3 기계 학습으로 만든 최첨단 인공지능 제품들

1.1.3 기계 학습 개념

- 간단한 기계 학습 예제
  - 가로축은 시간, 세로축은 이동체의 위치
  - 관측한 4개의 점이 데이터

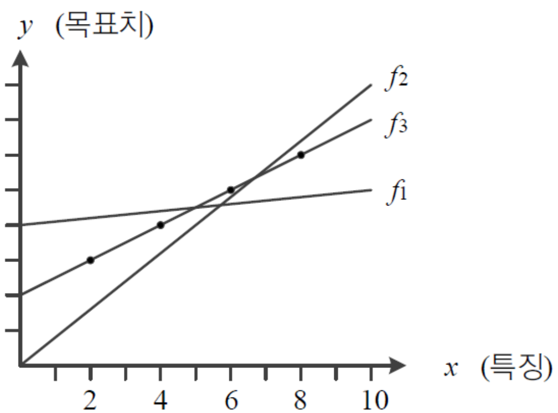


그림 1-4 간단한 기계 학습 예제

- 예측prediction 문제
  - 임의의 시간이 주어지면 이때 이동체의 위치는?
  - 회귀regression 문제와 분류classification 문제로 나뉨
    - 회귀는 목표치가 실수, 분류는 부류값 ([그림 1-4]는 회귀 문제)

1.1.3 기계 학습 개념

훈련집합

- 가로축은 특징, 세로축은 목표치
- 관측한 4개의 점이 훈련집합을 구성함

$$\text{훈련집합: } \mathbb{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}, \quad \mathbb{Y} = \{y_1, y_2, \cdots, y_n\} \tag{1.1}$$

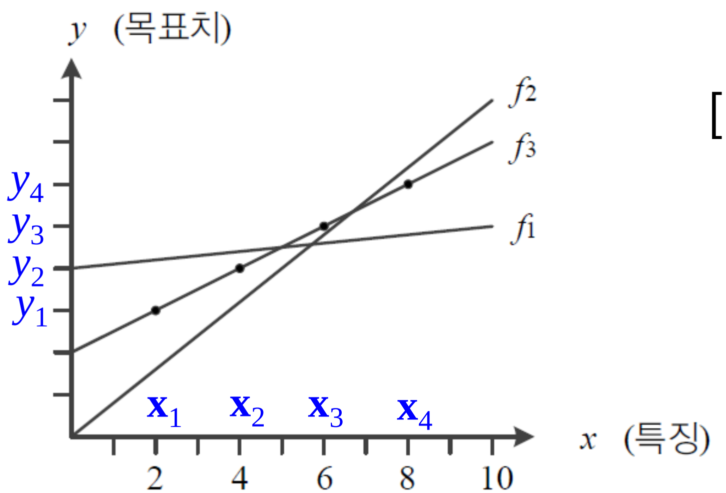


그림 1-4 간단한 기계 학습 예제

[그림 1-4] 예제의 훈련집합

$$\mathbb{X} = \{\mathbf{x}_1 = (2.0), \mathbf{x}_2 = (4.0), \mathbf{x}_3 = (6.0), \mathbf{x}_4 = (8.0)\}$$
$$\mathbb{Y} = \{y_1 = 3.0,\ y_2 = 4.0,\ y_3 = 5.0,\ y_4 = 6.0\}$$

1.1.3 기계 학습 개념

- 데이터를 어떻게 모델링할 것인가
  - 눈대중으로 보면 직선을 이루므로 직선을 선택하자 → 모델로 직선을 선택한 셈
  - 직선 모델의 수식
    - 2개의 매개변수 w와 b

$$y = wx + b \tag{1.2}$$

- 기계 학습은
  - 가장 정확하게 예측할 수 있는, 즉 최적의 매개변수를 찾는 작업
  - 처음에는 최적값을 모르므로 임의의 값에서 시작하고, 점점 성능을 개선하여 최적에 도달
  - [그림 1-4]의 예에서는 f_1에서 시작하여 $f_1 \rightarrow f_2 \rightarrow f_3$
    - 최적인 f_3은 $w=0.5$와 $b=2.0$

1.1.3 기계 학습 개념

- 학습을 마치면,
  - 예측에 사용
  - 예) 10.0 순간의 이동체 위치를 알고자 하면, $f_3(10.0)=0.5*10.0+2.0=7.0$이라 예측함

- 기계 학습의 궁극적인 목표
  - 훈련집합에 없는 새로운 샘플에 대한 오류를 최소화 (새로운 샘플 집합: 테스트 집합)
  - 테스트 집합에 대한 높은 성능을 일반화generalization 능력이라 부름

1.2.1 1차원과 2차원 특징 공간

- 1차원 특징 공간
- 2차원 특징 공간
  - 특징 벡터 표기
    - $\mathbf{x}=(x_1,x_2)^{\mathrm{T}}$
  - 예시
    - $\mathbf{x}$=(몸무게,키)$^{\mathrm{T}}$, y=장타율
    - $\mathbf{x}$=(체온,두통)$^{\mathrm{T}}$, y=감기 여부

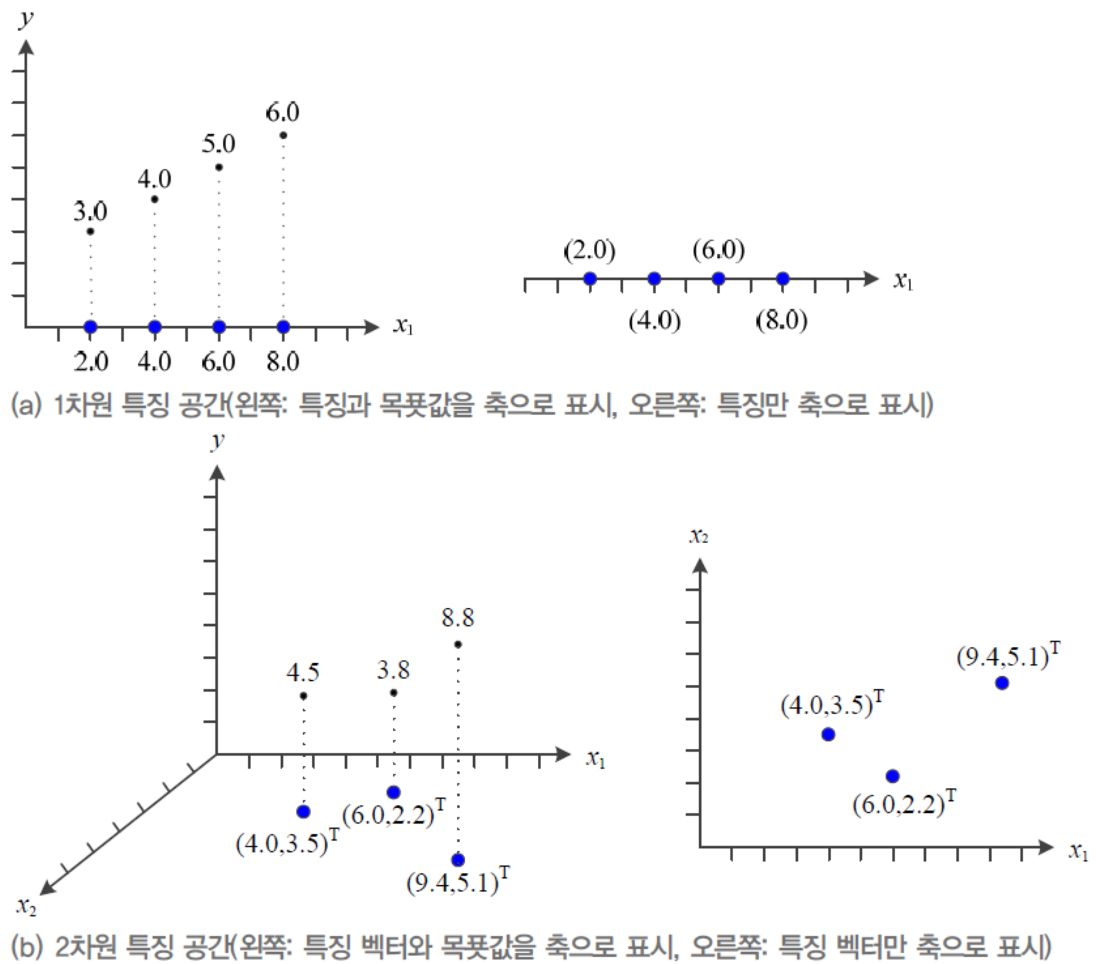
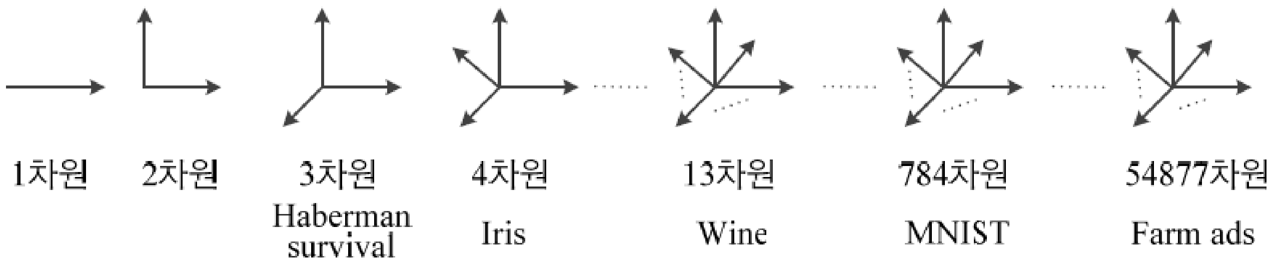


그림 1-5 특징 공간과 데이터의 표현

1.2.2 다차원 특징 공간

다차원 특징 공간 예제



Haberman survival: $\mathbf{x} = (\text{나이}, \text{수술년도}, \text{양성 림프샘 개수})^T$

Iris: $\mathbf{x} = (\text{꽃받침 길이}, \text{꽃받침 너비}, \text{꽃잎 길이}, \text{꽃잎 너비})^T$

Wine: $\mathbf{x} = (\text{Alcohol}, \text{Malic acid}, \text{Ash}, \text{Alcalinity of ash}, \text{Magnesium}, \text{Total phenols}, \text{Flavanoids}, \text{Nonflavanoid phenols}$
$\text{Proanthocyanins}, \text{Color intensity}, \text{Hue}, \text{OD280 / OD315 of diluted wines}, \text{Proline})^T$

MNIST: $\mathbf{x} = (\text{화소}1, \text{화소}2, \cdots, \text{화소}784)^T$

Farm ads: $\mathbf{x} = (\text{단어}1, \text{단어}2, \cdots, \text{단어}54877)^T$

그림 1-6 다차원 특징 공간

1.2.2 다차원 특징 공간

- d-차원 데이터
  - 특징 벡터 표기: $\mathbf{x}=(x_1,x_2,\ldots,x_d)^T$
- d-차원 데이터를 위한학습 모델
  - 직선 모델을 사용하는 경우 매개변수 수=d+1

$$y = \underline{w_1}x_1 + \underline{w_2}x_2 + \cdots + \underline{w_d}x_d + \underline{b} \qquad (1.3)$$

  - 2차 곡선 모델을 사용하면 매개변수 수가 크게 증가
    - 매개변수 수=d^2+d+1
    - 예) Iris 데이터: d=4이므로 21개의 매개변수
    - 예) MNIST 데이터: d=784이므로 615,441개의 매개변수

$$y = \underline{w_1}x_1^2 + \underline{w_2}x_2^2 + \cdots + \underline{w_d}x_d^2 + \underline{w_{d+1}}x_1x_2 + \cdots + \underline{w_{d^2}}x_{d-1}x_d + \underline{w_{d^2+1}}x_1 + \cdots + \underline{w_{d^2+d}}x_d + \underline{b} \qquad (1.5)$$

1.3.2 데이터베이스의 중요성

- Iris 데이터베이스는 통계학자인 피셔 교수가 1936년에 캐나다 동부 해안의 가스페 반도에 서식하는 3종의 붓꽃(setosa, versicolor, virginica)을 50송이씩 채취하여 만들었다[Fisher1936]. 150개 샘플 각각에 대해 꽃받침 길이, 꽃받침 너비, 꽃잎 길이, 꽃잎 너비를 측정하여 기록하였다. 따라서 4차원 특징 공간이 형성되며 목푯값은 3종을 숫자로 표시함으로써 1, 2, 3 값 중의 하나이다. http://archive.ics.uci.edu/ml/datasets/Iris에 접속하여 내려받을 수 있다.

Sepal length ◆	Sepal width ◆	Petal length ◆	Petal width ◆	Species ◆
5.2	3.5	1.4	0.2	*I. setosa*
4.9	3.0	1.4	0.2	*I. setosa*
4.7	3.2	1.3	0.2	*I. setosa*
4.6	3.1	1.5	0.2	*I. setosa*
7.0	3.2	4.7	1.4	*I. versicolor*
6.4	3.2	4.5	1.5	*I. versicolor*
6.9	3.1	4.9	1.5	*I. versicolor*
5.5	2.3	4.0	1.3	*I. versicolor*
6.3	3.3	6.0	2.5	*I. virginica*
5.8	2.7	5.1	1.9	*I. virginica*
7.1	3.0	5.9	2.1	*I. virginica*
6.3	2.9	5.6	1.8	*I. virginica*

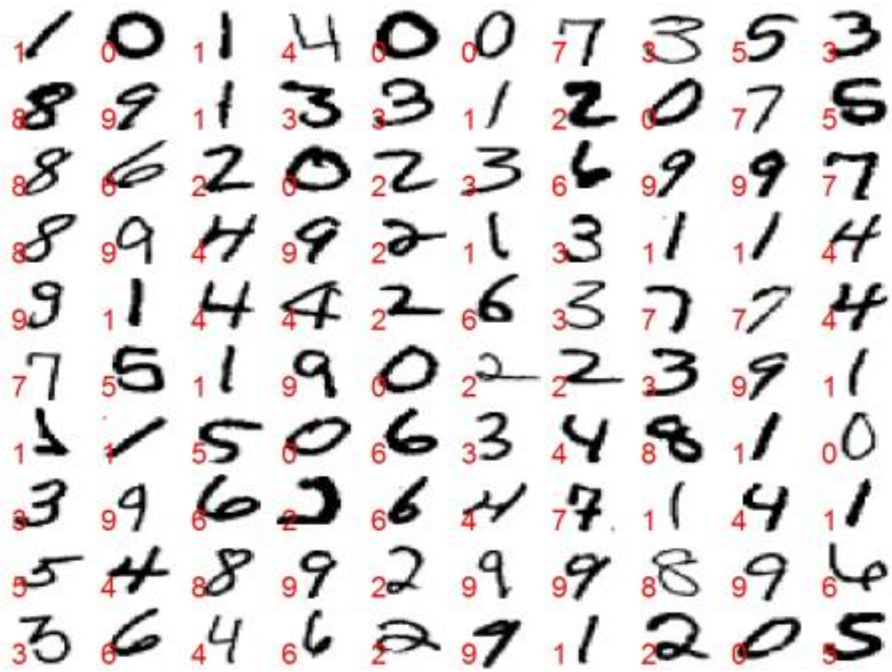

Setosa


Versicolor


Virginica

1.3.2 데이터베이스의 중요성

- MNIST 데이터베이스는 미국표준국(NIST)에서 수집한 필기 숫자 데이터베이스로, 훈련집합 60,000자, 테스트집합 10,000자를 제공한다. http://yann.lecun.com/exdb/mnist에 접속하면 무료로 내려받을 수 있으며, 1988년부터 시작한 인식률 경쟁 기록도 볼 수 있다. 2017년 8월 기준으로는 [Ciresan2012] 논문이 0.23%의 오류율로 최고 자리를 차지하고 있다. 테스트집합에 있는 10,000개 샘플에서 단지 23개만 틀린 것이다.



1.3.4 데이터 가시화

- 4차원 이상의 초공간은 한꺼번에 가시화 불가능
- 여러 가지 가시화 기법
  - 2개씩 조합하여 여러 개의 그래프 그림

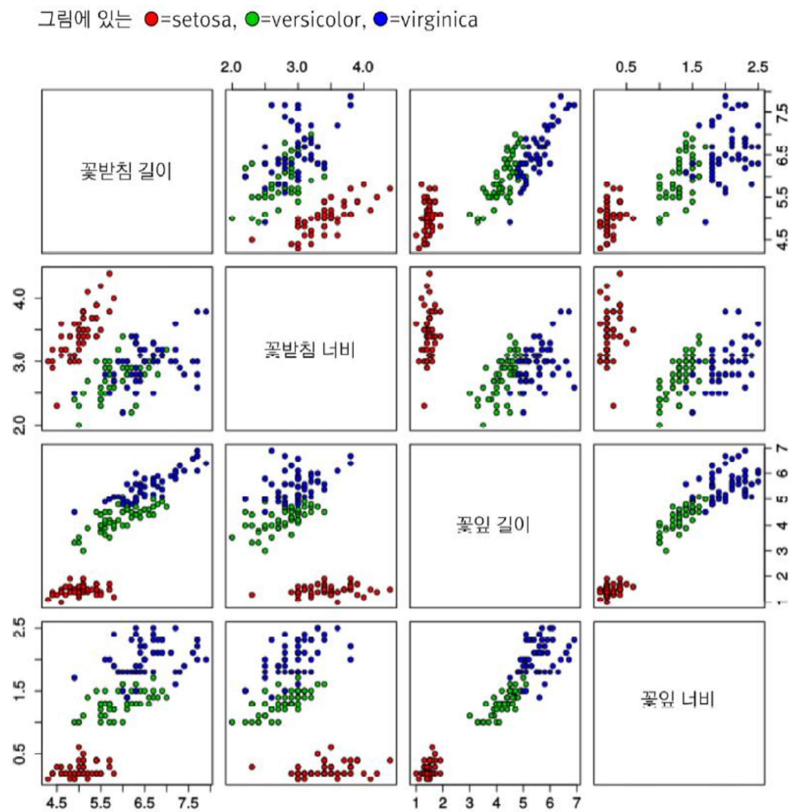


그림 1-10 고차원 특징 공간의 가시화

1.4 간단한 기계 학습의 예

선형 회귀 문제

- [그림 1-4]: 식 (1.2)의 직선 모델을 사용하므로 두 개의 매개변수 $\Theta = (w, b)^{\mathrm{T}}$

$$y = wx + b \tag{1.2}$$

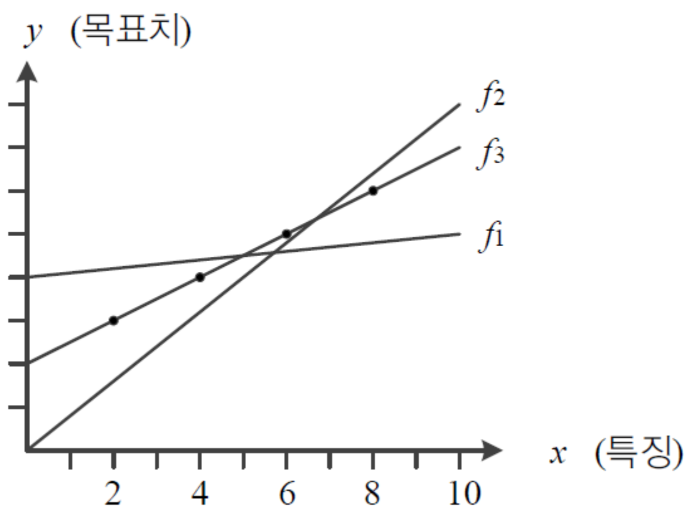


그림 1-4 간단한 기계 학습 예제

1.4 간단한 기계 학습의 예

- 목적 함수objective function (또는 비용 함수cost function)
  - 식 (1.8)은 선형 회귀를 위한 목적 함수
    - $f_\Theta(\mathbf{x}_i)$는 예측함수의 출력, y_i는 예측함수가 맞추어야 하는 목푯값이므로 $f_\Theta(\mathbf{x}_i)$-y_i는 오차
    - 식 (1.8)을 평균제곱오차MSE(mean squared error)라 부름

$$J(\Theta) = \frac{1}{n}\sum_{i=1}^{n}(f_\Theta(\mathbf{x}_i) - y_i)^2 \tag{1.8}$$

  - 처음에는 최적 매개변수 값을 알 수 없으므로 난수로 $\Theta_1 = (w_1, b_1)^\mathrm{T}$ 설정 → $\Theta_2 = (w_2, b_2)^\mathrm{T}$ 로 개선 → $\Theta_3 = (w_3, b_3)^\mathrm{T}$ 로 개선 → Θ_3는 최적해 $\hat{\Theta}$
    - 이때 $J(\Theta_1) > J(\Theta_2) > J(\Theta_3)$

1.4 간단한 기계 학습의 예

- [예제 1-1]
  - 훈련집합

    $\mathbb{X} = \{x_1 = (2.0), x_2 = (4.0),\ x_3 = (6.0),\ x_4 = (8.0)\}$,

    $\mathbb{Y} = \{y_1 = 3.0,\ y_2 = 4.0,\ y_3 = 5.0,\ y_4 = 6.0\}$
  - 초기 직선의 매개변수 $\Theta_1 = (0.1, 4.0)^{\mathrm{T}}$라 가정

$$
\begin{aligned}
\mathbf{x}_1, \mathrm{y}_1 &\longrightarrow \left(f_{\Theta_1}(2.0) - 3.0\right)^2 = \left((0.1 * 2.0 + 4.0) - 3.0\right)^2 = 1.44 \\
\mathbf{x}_2, y_2 &\longrightarrow \left(f_{\Theta_1}(4.0) - 4.0\right)^2 = \left((0.1 * 4.0 + 4.0) - 4.0\right)^2 = 0.16 \\
\mathbf{x}_3, y_3 &\longrightarrow \left(f_{\Theta_1}(6.0) - 5.0\right)^2 = \left((0.1 * 6.0 + 4.0) - 5.0\right)^2 = 0.16 \\
\mathbf{x}_4, y_4 &\longrightarrow \left(f_{\Theta_1}(8.0) - 6.0\right)^2 = \left((0.1 * 8.0 + 4.0) - 6.0\right)^2 = 1.44
\end{aligned}
\quad \longrightarrow \quad J(\Theta_1) = 0.8
$$

1.4 간단한 기계 학습의 예

[예제 1-1] 훈련집합

- Θ_1을 개선하여 $\Theta_2 = (0.8, 0.0)^{\mathrm{T}}$가 되었다고 가정

$$\mathbf{x}_1, y_1 \rightarrow \left(f_{\Theta_2}(2.0) - 3.0\right)^2 = \left((0.8 * 2.0 + 0.0) - 3.0\right)^2 = 1.96$$
$$\mathbf{x}_2, y_2 \rightarrow \left(f_{\Theta_2}(4.0) - 4.0\right)^2 = \left((0.8 * 4.0 + 0.0) - 4.0\right)^2 = 0.64$$
$$\mathbf{x}_3, y_3 \rightarrow \left(f_{\Theta_2}(6.0) - 5.0\right)^2 = \left((0.8 * 6.0 + 0.0) - 5.0\right)^2 = 0.04$$
$$\mathbf{x}_4, y_4 \rightarrow \left(f_{\Theta_2}(8.0) - 6.0\right)^2 = \left((0.8 * 8.0 + 0.0) - 6.0\right)^2 = 0.16$$

$$\longrightarrow J(\Theta_2) = 0.7$$

- Θ_2를 개선하여 $\Theta_3 = (0.5, 2.0)^{\mathrm{T}}$가 되었다고 가정
- 이때 $J(\Theta_3) = 0.0$이 되어 Θ_3은 최적값 $\hat{\Theta}$ 이 됨

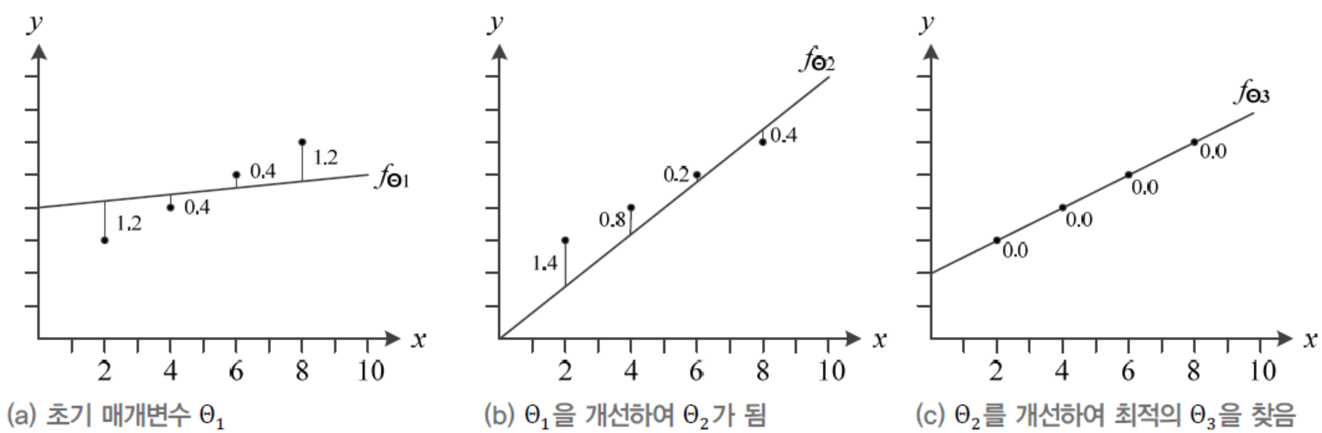


그림 1-11 기계 학습에서 목적함수의 역할

1.4 간단한 기계 학습의 예

- 기계 학습이 할 일을 공식화하면,

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} J(\Theta) \qquad (1.9)$$

  - 기계 학습은 작은 개선을 반복하여 최적해를 찾아가는 수치적 방법으로 식 (1.9)를 풂

- 알고리즘 형식으로 쓰면,

알고리즘 1-1 기계 학습 알고리즘

입력 : 훈련집합 $\mathbb{X}$와 $\mathbb{Y}$
출력 : 최적의 매개변수 $\hat{\Theta}$

난수를 생성하여 초기 해 Θ_1을 설정한다.
t=1
while ($J(\Theta_t)$가 0.0에 충분히 가깝지 않음) // 수렴 여부 검사
  $J(\Theta_t)$가 작아지는 방향 $\Delta\Theta_t$를 구한다. // $\Delta\Theta_t$는 주로 미분을 사용하여 구함
  $\Theta_{t+1} = \Theta_t + \Delta\Theta_t$
  t=t+1
$\hat{\Theta} = \Theta_t$

1.4 간단한 기계 학습의 예

- 좀더 현실적인 상황
  - 지금까지는 데이터가 선형을 이루는 아주 단순한 상황을 고려함
  - 실제 세계는 선형이 아니며 잡음이 섞임 → 비선형 모델이 필요

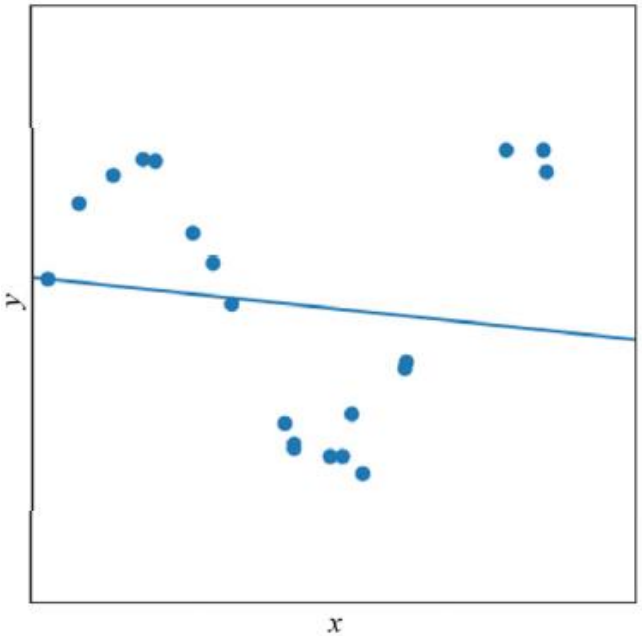


그림 1-12 선형 모델의 한계

1.5.1 과소적합과 과잉적합

- [그림 1.13]의 1차 모델은 과소적합
  - 모델의 '용량이 작아' 오차가 클 수밖에 없는 현상
- 비선형 모델을 사용하는 대안
  - [그림 1-13]의 2차, 3차, 4차, 12차는 다항식 곡선을 선택한 예
  - 1차(선형)에 비해 오차가 크게 감소함

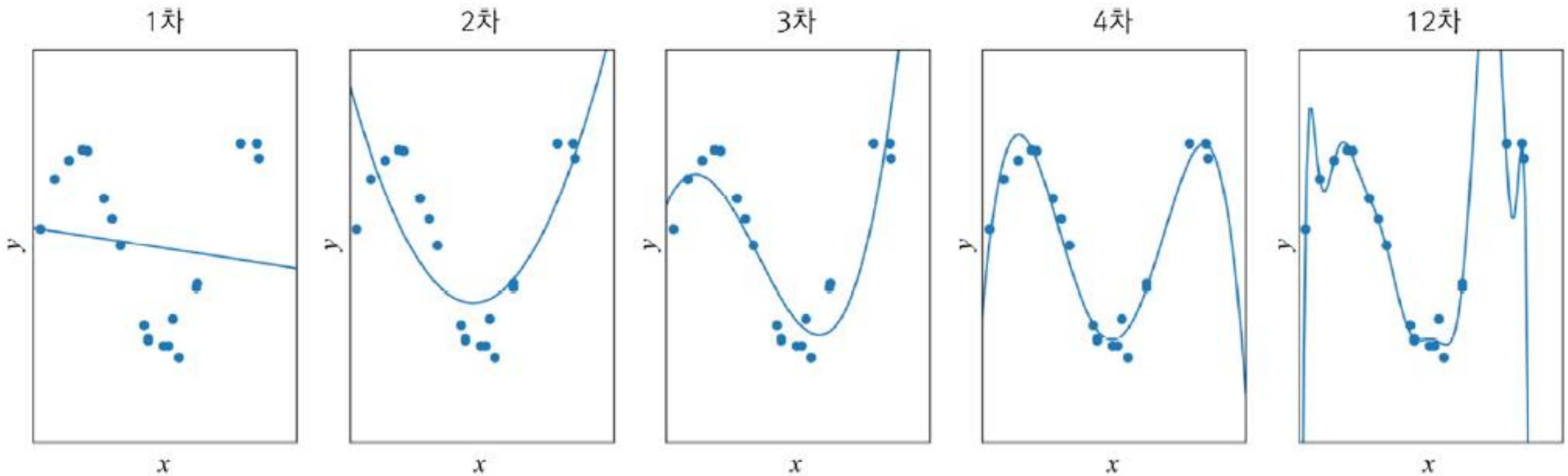


그림 1-13 과소적합과 과잉적합 현상

1.5.1 과소적합과 과잉적합

과잉적합

- 12차 다항식 곡선을 채택한다면 훈련집합에 대해 거의 완벽하게 근사화함
- 하지만 '새로운' 데이터를 예측한다면 큰 문제 발생
  - x_0에서 빨간 막대 근방을 예측해야 하지만 빨간 점을 예측
- 이유는 '용량이 크기' 때문. 학습 과정에서 잡음까지 수용 → 과잉적합 현상
- 적절한 용량의 모델을 선택하는 모델 선택 작업이 필요함

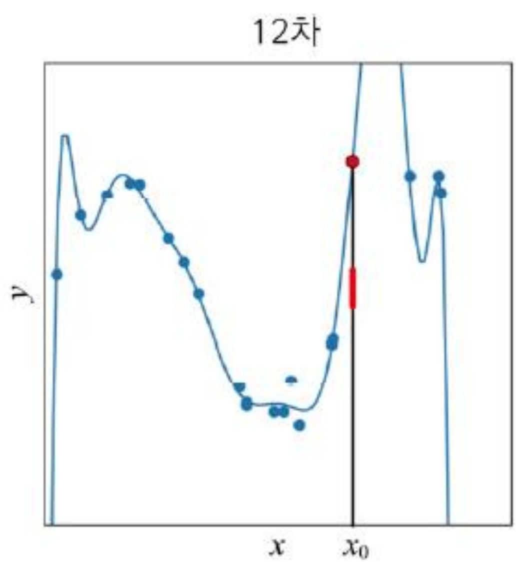


그림 1-14 과잉적합되었을 때 부정확한 예측 현상

1.5.2 바이어스와 분산

- 1차~12차 다항식 모델의 비교 관찰
  - 1~2차는 훈련집합과 테스트집합 모두 낮은 성능
  - 12차는 훈련집합에 높은 성능을 보이나 테스트집합에서는 낮은 성능 → 낮은 일반화 능력
  - 3~4차는 훈련집합에 대해 12차보다 낮겠지만 테스트집합에는 높은 성능 → 높은 일반화 능력

1.5.2 바이어스와 분산

- 훈련집합을 여러 번 수집하여 1차~12차에 적용하는 실험
  - 2차는 매번 큰 오차 → 바이어스가 큼. 하지만 비슷한 모델을 얻음 → 낮은 분산
  - 12차는 매번 작은 오차 → 바이어스가 작음. 하지만 크게 다른 모델을 얻음 → 높은 분산
  - 일반적으로 용량이 작은 모델은 바이어스는 크고 분산은 작음. 복잡한 모델은 바이어스는 작고 분산은 큼
  - 바이어스와 분산은 트레이드오프 관계

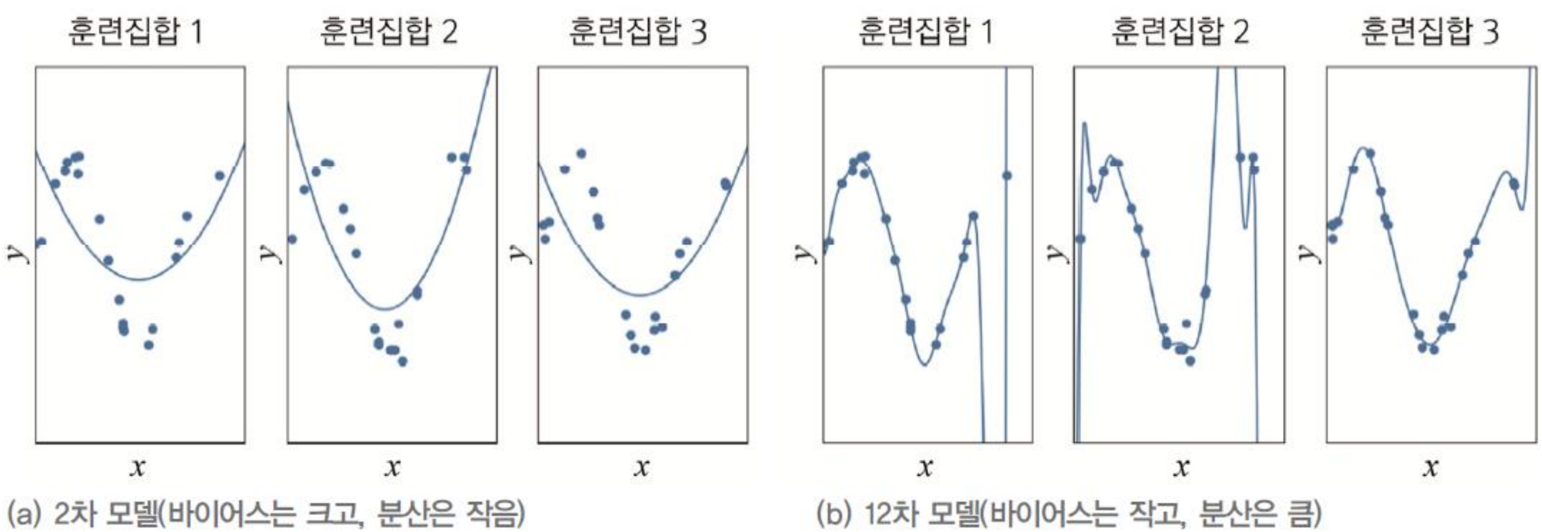


그림 1-15 모델의 바이어스와 분산 특성

1.5.2 바이어스와 분산

- 기계 학습의 목표
  - 낮은 바이어스와 낮은 분산을 가진 예측기 제작이 목표. 즉 왼쪽 아래 상황

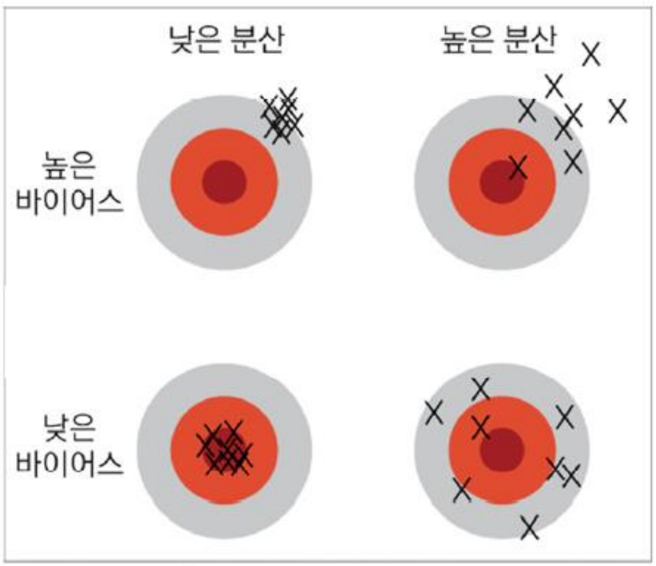


그림 1-16 바이어스와 분산

  - 하지만 바이어스와 분산은 트레이드오프 관계
  - 따라서 바이어스 희생을 최소로 유지하며 분산을 최대로 낮추는 전략 필요

1.5.3 검증집합과 교차검증을 이용한 모델 선택 알고리즘

검증집합을 이용한 모델 선택

- 훈련집합과 테스트집합과 다른 별도의 검증집합을 가진 상황

알고리즘 1-2 검증집합을 이용한 모델 선택

입력 : 모델집합 Ω, 훈련집합, 검증집합, 테스트집합
출력 : 최적 모델과 성능

```
for (Ω에 있는 각각의 모델)
        모델을 훈련집합으로 학습시킨다.
        검증집합으로 학습된 모델의 성능을 측정한다.   // 검증 성능 측정
가장 높은 성능을 보인 모델을 선택한다.
테스트집합으로 선택된 모델의 성능을 측정한다.
```

1.5.3 검증집합과 교차검증을 이용한 모델 선택 알고리즘

교차검증cross validation

- 비용 문제로 별도의 검증집합이 없는 상황에 유용한 모델 선택 기법
- 훈련집합을 등분하여, 학습과 평가 과정을 여러 번 반복한 후 평균 사용

알고리즘 1-3 교차검증에 의한 모델 선택

입력 : 모델집합 Ω, 훈련집합, 테스트집합, 그룹 개수 k
출력 : 최적 모델과 성능

```
훈련집합을 k개의 그룹으로 등분한다.
for (Ω에 있는 각각의 모델)
    for (i=1 to k)
        i번째 그룹을 제외한 k-1개 그룹으로 모델을 학습시킨다.
        학습된 모델의 성능을 i번째 그룹으로 측정한다.
    k개 성능을 평균하여 해당 모델의 성능으로 취한다.
가장 높은 성능을 보인 모델을 선택한다.
테스트집합으로 선택된 모델의 성능을 측정한다.
```

1.7.1 지도 방식에 따른 유형

- 지도 학습
  - 특징 벡터 $\mathbb{X}$와 목푯값 $\mathbb{Y}$가 모두 주어진 상황
  - 회귀와 분류 문제로 구분
- 비지도 학습
  - 특징 벡터 $\mathbb{X}$는 주어지는데 목푯값 $\mathbb{Y}$가 주어지지 않는 상황
  - 군집화 과업 (고객 성향에 따른 맞춤 홍보 응용 등)

1.7.1 지도 방식에 따른 유형

- 강화 학습
  - 목푯값이 주어지는데, 지도 학습과 다른 형태임
  - 예) 바둑
    - 수를 두는 행위가 샘플인데, 게임이 끝나면 목푯값 하나가 부여됨
      - 이기면 1, 패하면 -1을 부여
    - 게임을 구성한 샘플들 각각에 목푯값을 나누어 주어야 함

1. 행렬의 개념

정의 6-1 행렬(Matrix): $A = [a_{ij}]$

n, m이 양의 정수일 때 n행, m열로 나열된 실수의 2차원 배열

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = [a_{ij}]$$

- 가로줄을 행Row, 세로줄을 열Column, 행 크기와 열 크기로 행렬의 크기를 말함

$$A = [a_{ij}] = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$$

- 행렬의 크기는 3행 4열, 3×4 (3-by-4) 행렬이라고 함
- a_{ij}는 행렬 A의 i행, j열 원소를 의미
- 행렬 A의 i번째 행 : $\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{bmatrix}$ 행렬 A의 j번째 열 : $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{nj} \end{bmatrix}$

1. 행렬의 개념

예제 6-1

행렬 A에 대해 다음을 답하라.

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 10 & 8 & 6 & 4 \\ 9 & 11 & 13 & 15 \\ 12 & 14 & 16 & 2 \end{bmatrix}$$

(1) 두 번째 행 (2) 첫 번째 열 (3) a_{24} (4) a_{33}

풀이

(1) $\begin{bmatrix} 10 & 8 & 6 & 4 \end{bmatrix}$ (2) $\begin{bmatrix} 1 \\ 10 \\ 9 \\ 12 \end{bmatrix}$ (3) $a_{24} = 4$ (4) $a_{33} = 13$

2. 행렬의 연산

- 행렬에서 가능한 연산 : 덧셈, 뺄셈, 스칼라곱, 곱셈

❖ 행렬의 덧셈과 뺄셈

- 두 행렬의 크기가 같아야만 연산 가능

  (두 행렬의 행과 열의 크기가 각각 같음)

정의 6-2 행렬의 덧셈과 뺄셈

두 행렬 A, B에서 같은 자리에 있는 원소들끼리 더하거나 빼는 연산

– 덧셈 표현: $A+B$

– 뺄셈 표현: $A-B$

2. 행렬의 연산

$n \times m$ 크기의 행렬 A와 B가 각각 $A = \begin{bmatrix} a_{11} & a_{22} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$, $B = \begin{bmatrix} b_{11} & b_{22} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$ 일 때,

두 행렬의 덧셈과 뺄셈 연산은 다음과 같이 수행한다.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2m}+b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nm}+b_{nm} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1m}-b_{1m} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2m}-b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}-b_{n1} & a_{n2}-b_{n2} & \cdots & a_{nm}-b_{nm} \end{bmatrix}$$

2. 행렬의 연산

예제 6-2

다음 행렬 A, B를 이용해 주어진 문제를 연산하라.

$$A=\begin{bmatrix} 2 & 1 & 4 & 0 \\ -7 & 3 & 6 & 1 \\ 8 & -4 & -2 & 3 \\ 1 & 9 & 4 & -2 \end{bmatrix} \qquad B=\begin{bmatrix} 7 & -1 & 2 & 4 \\ 2 & 8 & -5 & 1 \\ -3 & -4 & 2 & -2 \\ 9 & 6 & -2 & 0 \end{bmatrix}$$

(1) $A+B$ (2) $A-B$ (3) $B-A$

풀이

(1) $A+B=\begin{bmatrix} 2 & 1 & 4 & 0 \\ -7 & 3 & 6 & 1 \\ 8 & -4 & -2 & 3 \\ 1 & 9 & 4 & -2 \end{bmatrix}+\begin{bmatrix} 7 & -1 & 2 & 4 \\ 2 & 8 & -5 & 1 \\ -3 & -4 & 2 & -2 \\ 9 & 6 & -2 & 0 \end{bmatrix}=\begin{bmatrix} 9 & 0 & 6 & 4 \\ -5 & 11 & 1 & 2 \\ 5 & -8 & 0 & 1 \\ 10 & 15 & 2 & -2 \end{bmatrix}$

(2) $A-B=\begin{bmatrix} 2 & 1 & 4 & 0 \\ -7 & 3 & 6 & 1 \\ 8 & -4 & -2 & 3 \\ 1 & 9 & 4 & -2 \end{bmatrix}-\begin{bmatrix} 7 & -1 & 2 & 4 \\ 2 & 8 & -5 & 1 \\ -3 & -4 & 2 & -2 \\ 9 & 6 & -2 & 0 \end{bmatrix}=\begin{bmatrix} -5 & 2 & 2 & -4 \\ -9 & -5 & 11 & 0 \\ 11 & 0 & -4 & 5 \\ -8 & 3 & 6 & -2 \end{bmatrix}$

(3) $B-A=\begin{bmatrix} 7 & -1 & 2 & 4 \\ 2 & 8 & -5 & 1 \\ -3 & -4 & 2 & -2 \\ 9 & 6 & -2 & 0 \end{bmatrix}-\begin{bmatrix} 2 & 1 & 4 & 0 \\ -7 & 3 & 6 & 1 \\ 8 & -4 & -2 & 3 \\ 1 & 9 & 4 & -2 \end{bmatrix}=\begin{bmatrix} 5 & -2 & -2 & 4 \\ 9 & 5 & -11 & 0 \\ -11 & 0 & 4 & -5 \\ 8 & -3 & -6 & 2 \end{bmatrix}$

2. 행렬의 연산

❖ 행렬의 스칼라곱

정의 6-3 행렬의 스칼라곱(Scalar Multiplication): $kA = Ak = [ka_{ij}]$

행렬 A에 실수 k를 곱하는 연산

- 행렬의 각 원소마다 그 실수 값을 곱함

$$kA = k\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1m} \\ ka_{21} & ka_{22} & \cdots & ka_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nm} \end{bmatrix}$$

예제 6-3

다음을 연산하라.

$$-4\begin{bmatrix} -1 & 2 & 5 \\ 3 & 2 & -7 \\ 8 & 3 & 1 \end{bmatrix}$$

풀이

$$-4\begin{bmatrix} -1 & 2 & 5 \\ 3 & 2 & -7 \\ 8 & 3 & 1 \end{bmatrix} = \begin{bmatrix} -4\times-1 & -4\times2 & -4\times5 \\ -4\times3 & -4\times2 & -4\times-7 \\ -4\times8 & -4\times3 & -4\times1 \end{bmatrix} = \begin{bmatrix} 4 & -8 & -20 \\ -12 & -8 & 28 \\ -32 & -12 & -4 \end{bmatrix}$$

2. 행렬의 연산

❖ 행렬의 곱셈

정의 6-4 행렬의 곱셈

$n \times m$ 행렬 A와 $r \times s$ 행렬 B가 있고 $m = r$일 때, $n \times s$ 행렬 $A \cdot B = [c_{ij}]$

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{r1} & b_{r2} & \dots & b_{rs} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1s} \\ \dots & \dots & \dots & \dots & \dots \\ c_{i1} & \dots & c_{ij} & \dots & c_{is} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n1} & \dots & c_{nj} & \dots & c_{ns} \end{bmatrix}$$

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}$$

- 곱셈 연산 수행

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{ms} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1m}b_{m1} & a_{11}b_{12} + a_{12}b_{22} + \dots + a_{1m}b_{m2} & \dots & a_{11}b_{1s} + a_{12}b_{2s} + \dots + a_{1m}b_{ms} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2m}b_{m1} & a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2m}b_{m2} & \dots & a_{21}b_{1s} + a_{22}b_{2s} + \dots + a_{2m}b_{ms} \\ \dots & \dots & \dots & \dots \\ a_{n1}b_{11} + a_{n2}b_{21} + \dots + a_{nm}b_{m1} & a_{n1}b_{12} + a_{n2}b_{22} + \dots + a_{nm}b_{m2} & \dots & a_{n1}b_{1s} + a_{n2}b_{2s} + \dots + a_{nm}b_{ms} \end{bmatrix}$$

2. 행렬의 연산

- A의 i번째 행과 행렬 B의 j번째 열이 서로 대응하여 연산되기 때문에 행렬 A의 열 크기와 행렬 B의 행 크기가 같아야 함

연산

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \quad \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{ms} \end{bmatrix}$$

- 행렬 A의 크기가 $n \times m$ 이고, 행렬 B의 크기가 $m \times s$ 일 때 곱 AB의 결과로 나오는 행렬의 크기는 $n \times s$ 임

예제 6-4

행렬 A, B, C가 다음과 같을 때, 연산이 가능한 것을 골라 연산하라.

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 2 \\ 9 & 3 & 8 \\ 4 & 7 & 5 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 6 & 1 \\ 3 & 5 & 2 \end{bmatrix}$$

(1) AB (2) BA (3) AC (4) CA (5) BC (6) CB

2. 행렬의 연산

풀이

(1) AB는 연산이 가능하지 않다. 행렬 A는 2×2 행렬, 행렬 B는 3×3 행렬로 행렬 A의 열 크기(2)와 행렬 B의 행 크기(3)가 서로 다르기 때문이다.

(2) BA는 연산이 가능하지 않다. 행렬 B는 3×3 행렬, 행렬 A는 2×2 행렬로 행렬 B의 열 크기(3)와 행렬 A의 행 크기(2)가 서로 다르기 때문이다.

(3) AC는 연산이 가능하다. 행렬 A는 2×2 행렬, 행렬 C는 2×3 행렬로 행렬 A의 열 크기(2)와 행렬 B의 행 크기(2)가 같기 때문이다.

$$\therefore\ AC=\begin{bmatrix}1&3\\4&6\end{bmatrix}\times\begin{bmatrix}0&6&1\\3&5&2\end{bmatrix}$$

$$=\begin{bmatrix}1\times0+3\times3 & 1\times6+3\times5 & 1\times1+3\times2\\4\times0+6\times3 & 4\times6+6\times5 & 4\times1+6\times2\end{bmatrix}=\begin{bmatrix}9&21&7\\18&54&16\end{bmatrix}$$

(4) CA는 연산이 가능하지 않다. 행렬 C는 2×3 행렬, 행렬 A는 2×2 행렬로 행렬 C의 열 크기(3)와 행렬 A의 행 크기(2)가 서로 다르기 때문이다.

2. 행렬의 연산

(5) BC는 연산이 가능하지 않다. 행렬 B는 3×3 행렬, 행렬 C는 2×3 행렬로 행렬 B의 열 크기(3)와 행렬 A의 행 크기(2)가 서로 다르기 때문이다.

(6) CB는 연산이 가능하다. 행렬 C는 2×3 행렬, 행렬 B는 3×3 행렬로 행렬 C의 열 크기(3)와 행렬 B의 행 크기(3)가 같기 때문이다.

$$\therefore\ CB = \begin{bmatrix} 0 & 6 & 1 \\ 3 & 5 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 2 \\ 9 & 3 & 8 \\ 4 & 7 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 0\times1+6\times9+1\times4 & 0\times0+6\times3+1\times7 & 0\times2+6\times8+1\times5 \\ 3\times1+5\times9+2\times4 & 3\times0+5\times3+2\times7 & 3\times2+5\times8+2\times5 \end{bmatrix}$$

$$= \begin{bmatrix} 58 & 25 & 53 \\ 56 & 29 & 56 \end{bmatrix}$$

2. 행렬의 연산

정리 6-1 행렬 연산의 성질

(1) $A+B=B+A$

(2) $A+(B+C)=(A+B)+C$

(3) $A+O=O+A=A$

(4) $A+(-A)=(-A)+A=O$

(5) $(-1)A=-A$

(6) $k(A+B)=kA+kB$

(7) $(k+l)A=kA+lA$

(8) $(kl)A=k(lA)$

(9) $k(AB)=(kA)B=A(kB)$

(10) $IA=A=AI$

※ O: 영행렬

I : 단위행렬

3. 행렬의 종류

정의 6-5 영행렬(Zero Matrix: O)

$n \times m$ 행렬 $A = [a_{ij}]$가 있을 때, 모든 i, j에 대하여 $a_{ij} = 0$인 행렬

$$O = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

정의 6-6 n차 정사각행렬(n-square Matrix)

$n \times m$ 행렬 $A = [a_{ij}]$가 있을 때, $m = n$인 행렬

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

3. 행렬의 종류

정의 6-7 대각행렬(Diagonal Matrix)

n차 정사각행렬에서 대각원소 $a_{11}, a_{22}, \dots, a_{nn}$ 이외의 모든 원소가 0인 행렬

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

정의 6-8 단위행렬(항등행렬, Unit Matrix, Identity Matrix: I)

대각행렬에서 대각원소가 모두 1인 행렬

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

3. 행렬의 종류

- 단위행렬은 행렬의 곱셈에서 $AI=IA=A$ 이기 때문에 항등행렬 이라고도 함
- 단위행렬과의 곱셈 연산은 항상 교환법칙이 성립

예제 6-5

행렬 $A=\begin{bmatrix} 4 & 6 & -1 \\ -8 & 3 & 2 \\ 2 & 8 & 4 \end{bmatrix}$ 와 단위행렬 I를 다음과 같이 곱셈 연산하라.

(1) AI (2) IA

풀이

(1) $AI=\begin{bmatrix} 4 & 6 & -1 \\ -8 & 3 & 2 \\ 2 & 8 & 4 \end{bmatrix}\times\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$=\begin{bmatrix} 4\times1+6\times0+(-1)\times0 & 4\times0+6\times1+(-1)\times0 & 4\times0+6\times0+(-1)\times1 \\ (-8)\times1+3\times0+2\times0 & (-8)\times0+3\times1+2\times0 & (-8)\times0+3\times0+2\times1 \\ 2\times1+8\times0+4\times0 & 2\times0+8\times1+4\times0 & 2\times0+8\times0+4\times1 \end{bmatrix}$$

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

(2) $IA=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\times\begin{bmatrix} 4 & 6 & -1 \\ -8 & 3 & 2 \\ 2 & 8 & 4 \end{bmatrix}$

$$=\begin{bmatrix} 1\times4+0\times(-8)+0\times2 & 1\times6+0\times3+0\times8 & 1\times(-1)+0\times2+0\times4 \\ 0\times4+1\times(-8)+0\times2 & 0\times6+1\times3+0\times8 & 0\times(-1)+1\times2+0\times4 \\ 0\times4+0\times(-8)+1\times2 & 0\times6+0\times3+1\times8 & 0\times(-1)+0\times2+1\times4 \end{bmatrix}$$

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

3. 행렬의 종류

정의 6-9 전치행렬(Transpose Matrix: A^T)

$n \times m$ 행렬 $A = [a_{ij}]$가 있을 때, 행과 열을 바꾼 $m \times n$ 행렬

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \text{일 때, } A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}$$

정의 6-10 대칭행렬(Symmetric Matrix)

n 차 정사각행렬 $A = [a_{ij}]$가 있을 때, $A^T = A$인 행렬

3. 행렬의 종류

예제 6-6

다음 행렬의 전치행렬을 구하고, 대칭행렬인지 구별하라.

(1) $A=\begin{bmatrix} 1 & 0 & -2 & 6 \\ -3 & 5 & 2 & -4 \\ 0 & 4 & 4 & 8 \\ 9 & -1 & 3 & 7 \end{bmatrix}$ (2) $B=\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \end{bmatrix}$

풀이

(1) $A^T=\begin{bmatrix} 1 & -3 & 0 & 9 \\ 0 & 5 & 4 & -1 \\ -2 & 2 & 4 & 3 \\ 6 & -4 & 8 & 7 \end{bmatrix}$ $A^T \neq A$ 이므로 대칭행렬이 아니다.

(2) $B^T=\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \end{bmatrix}$ $B^T=B$ 이므로 대칭행렬이다.

3. 행렬의 종류

정의 6-11 부울행렬(Boolean Matrix)

행렬의 모든 원소가 부울값(0과 1)으로만 구성된 행렬

- 부울행렬 : 원소 간의 관계를 표현하거나 관계를 합성하는 데에 유용하게 사용되는 행렬, 0과 1로만 표현되기 때문에 일반 행렬과 다른 연산 방식을 사용

3. 행렬의 종류

정리 6-2 부울행렬 연산자

행렬 $A=[a_{ij}]$와 $B=[b_{ij}]$에 대해

(1) 합[Join]: $A \vee B=[a_{ij} \vee b_{ij}]$

(2) 교차[Meet]: $A \wedge B=[a_{ij} \wedge b_{ij}]$

(3) 부울곱[Boolean Product]: $A \odot B$

$n \times m$ 부울행렬 $A=[a_{ij}]$와 $m \times s$ 부울행렬 $B=[b_{ij}]$가 있을 때, $n \times s$ 부울행렬 $A \odot B=[c_{ij}]$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \odot \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{ms} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & b_{1j} & \cdots & c_{1s} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{is} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nj} & \cdots & c_{ns} \end{bmatrix}$$

$$c_{ij}=(a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \vee \ldots \vee (a_{im} \wedge b_{mj})$$

3. 행렬의 종류

- 부울행렬의 합
  - 논리합(∨) 연산과 같은 방식으로 연산
  - 행렬 A의 원소인 a_{ij}와 행렬 B의 원소인 b_{ij}중 하나라도 1이면 합 연산의 결과는 1이 됨

$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$이고 $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$이라면, 두 행렬의 합 연산은 다음과 같다.

$$A \vee B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vee \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \vee 0 & 0 \vee 0 \\ 0 \vee 1 & 1 \vee 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- 부울행렬의 교차
  - 논리곱(∧) 연산과 방식이 같음
  - 행렬 A의 원소인 a_{ij}와 행렬 B의 원소인 b_{ij}모두 1인 경우에만 교차 연산의 결과가 1이 됨

$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$이고 $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$인 예를 이용해 두 행렬의 교차 연산을 수행하면 다음과 같다.

$$A \wedge B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \wedge \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \wedge 0 & 0 \wedge 0 \\ 0 \wedge 1 & 1 \wedge 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

3. 행렬의 종류

- 부울행렬의 부울곱
  - 행렬의 곱셈 방식과 논리합, 논리곱의 연산을 적용하여 수행

$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 이고 $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ 이라면, 두 행렬의 부울곱 연산은 다음과 같다.

$$A \odot B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \odot \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} (1 \wedge 0) \vee (0 \wedge 1) & (1 \wedge 0) \vee (0 \wedge 1) \\ (0 \wedge 0) \vee (1 \wedge 1) & (0 \wedge 0) \vee (1 \wedge 1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

예제 6-7

다음을 연산하라.

(1) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \vee \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ (2) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \wedge \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (3) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \odot \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$

풀이

3. 행렬의 종류

$$(1)\quad \begin{bmatrix}1&1&0\\0&1&0\\0&1&1\end{bmatrix} \vee \begin{bmatrix}1&1&1\\0&0&1\\1&0&1\end{bmatrix} = \begin{bmatrix}1\vee 1&1\vee 1&0\vee 1\\0\vee 0&1\vee 0&0\vee 1\\0\vee 1&1\vee 0&1\vee 1\end{bmatrix} = \begin{bmatrix}1&1&1\\0&1&1\\1&1&1\end{bmatrix}$$

$$(2)\quad \begin{bmatrix}0&1&0\\0&1&1\end{bmatrix} \wedge \begin{bmatrix}1&1&1\\0&0&1\end{bmatrix} = \begin{bmatrix}0\wedge 1&1\wedge 1&0\wedge 1\\0\wedge 0&1\wedge 0&1\wedge 1\end{bmatrix} = \begin{bmatrix}0&1&0\\0&0&1\end{bmatrix}$$

$$(3)\quad \begin{bmatrix}1&0&0\\1&1&1\end{bmatrix} \odot \begin{bmatrix}0&1\\1&1\\1&0\end{bmatrix} = \begin{bmatrix}(1\wedge 0)\vee(0\wedge 1)\vee(0\wedge 1)&(1\wedge 1)\vee(0\wedge 1)\vee(0\wedge 0)\\(1\wedge 0)\vee(1\wedge 1)\vee(1\wedge 1)&(1\wedge 1)\vee(1\wedge 1)\vee(1\wedge 0)\end{bmatrix}$$

$$= \begin{bmatrix}0\vee 0\vee 0&1\vee 0\vee 0\\0\vee 1\vee 1&1\vee 1\vee 0\end{bmatrix} = \begin{bmatrix}0&1\\1&1\end{bmatrix}$$

정리 6-3 부울행렬 연산의 특징

(1) $A\vee A = A$ $\qquad$ $A\wedge A = A$

(2) $A\vee B = B\vee A$ $\qquad$ $A\wedge B = B\wedge A$

(3) $(A\vee B)\vee C = A\vee (B\vee C)$ $\qquad$ $(A\wedge B)\wedge C = A\wedge (B\wedge C)$

$A\odot (B\odot C) = (A\odot B)\odot C$

(4) $A\vee (B\wedge C) = (A\vee B)\wedge (A\vee C)$ $\qquad$ $A\wedge (B\vee C) = (A\wedge B)\vee (A\wedge C)$

4. 행렬식

정의 6-12 행렬식(Determinant: $|A|$ 또는 $\det(A)$)

n차 정사각행렬에 대응하는 함수

$$|A| = \det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

4. 행렬식

❖ 2차, 3차 정사각행렬에 대한 기본 행렬식

정의 6-13 2차, 3차 정사각행렬에 대한 행렬식

- 2차 정사각행렬 $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 의 행렬식

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

- 3차 정사각행렬 $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$ 의 행렬식

$$\det(B) = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \begin{matrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{matrix}$$

$$= (b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{32}b_{21}) - (b_{13}b_{22}b_{31} + b_{23}b_{32}b_{11} + b_{33}b_{21}b_{12})$$

4. 행렬식

예제 6-8

다음 정사각행렬의 행렬식을 구하라.

(1) $A = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}$ (2) $B = \begin{bmatrix} 3 & -1 & -2 \\ -4 & 2 & 1 \\ 1 & 4 & -3 \end{bmatrix}$

풀이

(1) $\det(A) = \begin{vmatrix} 3 & 1 \\ 2 & -1 \end{vmatrix} = 3 \times (-1) - 1 \times 2 = -3 - 2 = -5$

(2) $\det(B) = \begin{vmatrix} 3 & -1 & -2 \\ -4 & 2 & 1 \\ 1 & 4 & -3 \end{vmatrix}$

$$= [(3 \times 2 \times (-3) + (-1) \times 1 \times 1 + (-2) \times 4 \times (-4)]$$
$$- [(-2) \times 2 \times 1 + 1 \times 4 \times 3 + (-3) \times (-4) \times (-1)]$$
$$= 17$$

4. 행렬식

❖ 3차 이상의 정사각행렬에 대한 행렬식

- 소행렬과 소행렬식
  - 3차 이상의 정사각행렬의 행렬식은 행렬을 작게 분할한 소행렬을 이용함

정의 6-14 소행렬(Minor Matrix: M_{ij})

n차 정사각행렬에서 i번째 행과 j번째 열을 제거해서 얻은 $(n-1)\times(n-1)$행렬

예를 들어 행렬 $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$이 있을 때, 소행렬 M_{11}은 행렬 A에서 1행과 1열을 제외한 나머지 부분, 즉, $M_{11} = \begin{bmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{bmatrix}$이다. 소행렬 M_{32}는 행렬 A에서 3행과 2열을 제외하여 얻은 행렬 $M_{32} = \begin{bmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 13 & 15 & 16 \end{bmatrix}$이다. 이와 같이 어떤 행렬 A에 대한 소행렬 M_{ij}는 행렬 A보다 행과 열의 크기가 하나씩 작다.

정의 6-15 소행렬식($\det(M_{ij})$)

n차 정사각행렬의 소행렬 M_{ij}에 대한 소행렬식

예제 6-9

정사각행렬 $A = \begin{bmatrix} 5 & 1 & 3 \\ 2 & 6 & 4 \\ 1 & 3 & 6 \end{bmatrix}$의 가능한 소행렬을 모두 구하고, 각각의 행렬식을 구하라.

풀이

$M_{11} = \begin{bmatrix} 6 & 4 \\ 3 & 6 \end{bmatrix}$: 1행과 1열을 제외한 원소들로 구성

$$\det(M_{11}) = 6 \times 6 - 4 \times 3 = 24$$

$M_{12} = \begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}$: 1행과 2열을 제외한 원소들로 구성

$$\det(M_{12}) = 2 \times 6 - 4 \times 1 = 8$$

$M_{13} = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}$: 1행과 3열을 제외한 원소들로 구성

$$\det(M_{13}) = 2 \times 3 - 6 \times 1 = 0$$

4. 행렬식

$$M_{21} = \begin{bmatrix} 1 & 3 \\ 3 & 6 \end{bmatrix}$$: 2행과 1열을 제외한 원소들로 구성

$$\det(M_{21}) = 1 \times 6 - 3 \times 3 = -3$$

$$M_{22} = \begin{bmatrix} 5 & 3 \\ 1 & 6 \end{bmatrix}$$: 2행과 2열을 제외한 원소들로 구성

$$\det(M_{22}) = 5 \times 6 - 3 \times 1 = 27$$

$$M_{23} = \begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix}$$: 2행과 3열을 제외한 원소들로 구성

$$\det(M_{23}) = 5 \times 3 - 1 \times 1 = 14$$

$$M_{31} = \begin{bmatrix} 1 & 3 \\ 6 & 4 \end{bmatrix}$$: 3행과 1열을 제외한 원소들로 구성

$$\det(M_{31}) = 1 \times 4 - 6 \times 3 = -14$$

$$M_{32} = \begin{bmatrix} 5 & 3 \\ 2 & 4 \end{bmatrix}$$: 3행과 2열을 제외한 원소들로 구성

$$\det(M_{32}) = 5 \times 4 - 2 \times 3 = 14$$

$$M_{33} = \begin{bmatrix} 5 & 1 \\ 2 & 6 \end{bmatrix}$$: 3행과 3열을 제외한 원소들로 구성

$$\det(M_{33}) = 5 \times 6 - 1 \times 2 = 28$$

4. 행렬식

- 여인수와 여인수행렬

정의 6-16 여인수(Cofactor: A_{ij}), 여인수행렬(Cofactor Matrix: $[A_{ij}]$)

n차 정사각행렬 $A=[a_{ij}]$에서 원소 a_{ij}에 관련된 계수와 그 계수들의 행렬

$$A_{ij}=(-1)^{i+j}\det(M_{ij})$$

- 여인수는 행렬식을 구하는 식에서 행렬 A의 원소 a_{ij}의 계수가 되는 수로 소행렬식에 의해 결정되며, 여인수행렬 내에서의 위치에 따라 부호가 정해짐

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

[그림 6-1] 여인수행렬에서 각 원소의 부호

4. 행렬식

예제 6-10

[예제 6−9]에서 사용된 행렬 $A = \begin{bmatrix} 5 & 1 & 3 \\ 2 & 6 & 4 \\ 1 & 3 & 6 \end{bmatrix}$의 각 원소에 대한 여인수를 구하여 여인수행렬을 구하라.

풀이

$$A_{11} = (-1)^{1+1}\det(M_{11}) = \det(M_{11}) = \begin{vmatrix} 6 & 4 \\ 3 & 6 \end{vmatrix} = 6 \times 6 - 4 \times 3 = 24$$

$$A_{12} = (-1)^{1+2}\det(M_{12}) = -\det(M_{12}) = -\begin{vmatrix} 2 & 4 \\ 1 & 6 \end{vmatrix} = -(2 \times 6 - 4 \times 1) = -8$$

$$A_{13} = (-1)^{1+3}\det(M_{13}) = \det(M_{13}) = \begin{vmatrix} 2 & 6 \\ 1 & 3 \end{vmatrix} = 2 \times 3 - 6 \times 1 = 0$$

$$A_{21} = (-1)^{2+1}\det(M_{21}) = -\det(M_{21}) = -\begin{vmatrix} 1 & 3 \\ 3 & 6 \end{vmatrix} = -(1 \times 6 - 3 \times 3) = 3$$

4. 행렬식

$$A_{22} = (-1)^{2+2}\det(M_{22}) = \det(M_{22}) = \begin{vmatrix} 5 & 3 \\ 1 & 6 \end{vmatrix} = 5\times 6 - 3\times 1 = 27$$

$$A_{23} = (-1)^{2+3}\det(M_{23}) = -\det(M_{23}) = -\begin{vmatrix} 5 & 1 \\ 1 & 3 \end{vmatrix} = -(5\times 3 - 1\times 1) = -14$$

$$A_{31} = (-1)^{3+1}\det(M_{31}) = \det(M_{31}) = \begin{vmatrix} 1 & 3 \\ 6 & 4 \end{vmatrix} = 1\times 4 - 6\times 3 = -14$$

$$A_{32} = (-1)^{3+2}\det(M_{32}) = -\det(M_{32}) = -\begin{vmatrix} 5 & 3 \\ 2 & 4 \end{vmatrix} = -(5\times 4 - 2\times 3) = -14$$

$$A_{33} = (-1)^{3+3}\det(M_{33}) = \det(M_{33}) = \begin{vmatrix} 5 & 1 \\ 2 & 6 \end{vmatrix} = 5\times 6 - 1\times 2 = 28$$

$$\therefore\ [A_{ij}] = \begin{bmatrix} 24 & -8 & 0 \\ 3 & 27 & -14 \\ -14 & -14 & 28 \end{bmatrix}$$

4. 행렬식

- 여인수를 이용한 행렬식

정의 6-17 여인수를 이용한 행렬식

n차 정사각행렬 A에 대한 행렬식은

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} : \ i\text{행을 선택한 경우}$$
$$= a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} : \ j\text{열을 선택한 경우}$$

- 여인수를 이용한 행렬식의 원리
  - 행렬식을 구해야 하는 n차 정사각행렬에서 행이나 열 중에서 하나를 선택하여 해당하는 원소의 여인수와 곱한 후 그 결과를 더하여 구하는 방식

예제 6-11

[예제 6-9]에서 사용된 3차 정사각행렬 $A = \begin{bmatrix} 5 & 1 & 3 \\ 2 & 6 & 4 \\ 1 & 3 & 6 \end{bmatrix}$ 의 행렬식을 구하라.

풀이

• 1행을 선택했을 경우

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$
$$= 5 \cdot (-1)^{1+1}\det(M_{11}) + 1 \cdot (-1)^{1+2}\det(M_{12}) + 3 \cdot (-1)^{1+3}\det(M_{13})$$

4. 행렬식

$$= 5 \cdot \begin{vmatrix} 6 & 4 \\ 3 & 6 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 2 & 4 \\ 1 & 6 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & 6 \\ 1 & 3 \end{vmatrix}$$
$$= 5(6 \cdot 6 - 4 \cdot 3) - (2 \cdot 6 - 4 \cdot 1) + 3(2 \cdot 3 - 6 \cdot 1) = 112$$

- 2행을 선택했을 경우

$$\det(A) = a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23}$$
$$= 2 \cdot (-1)^{2+1}\det(M_{21}) + 6 \cdot (-1)^{2+2}\det(M_{22}) + 4 \cdot (-1)^{2+3}\det(M_{23})$$
$$= (-2) \cdot \begin{vmatrix} 1 & 3 \\ 3 & 6 \end{vmatrix} + 6 \cdot \begin{vmatrix} 5 & 3 \\ 1 & 6 \end{vmatrix} + (-4) \cdot \begin{vmatrix} 5 & 1 \\ 1 & 3 \end{vmatrix}$$
$$= (-2)(1 \cdot 6 - 3 \cdot 3) + 6(5 \cdot 6 - 3 \cdot 1) - 4(5 \cdot 3 - 1 \cdot 1) = 112$$

- 3행을 선택했을 경우

$$\det(A) = a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33}$$
$$= 1 \cdot (-1)^{3+1}\det(M_{31}) + 3 \cdot (-1)^{3+2}\det(M_{32}) + 6 \cdot (-1)^{3+3}\det(M_{33})$$
$$= 1 \cdot \begin{vmatrix} 1 & 3 \\ 6 & 4 \end{vmatrix} + (-3) \cdot \begin{vmatrix} 5 & 3 \\ 2 & 4 \end{vmatrix} + 6 \cdot \begin{vmatrix} 5 & 1 \\ 2 & 6 \end{vmatrix}$$
$$= (1 \cdot 4 - 3 \cdot 6) - 3(5 \cdot 4 - 3 \cdot 2) + 6(5 \cdot 6 - 1 \cdot 2) = 112$$

4. 행렬식

- 1열을 선택했을 경우

$$\det(A) = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31}$$

$$= 5 \cdot (-1)^{1+1}\det(M_{11}) + 2 \cdot (-1)^{2+1}\det(M_{21}) + 1 \cdot (-1)^{3+1}\det(M_{31})$$

$$= 5 \cdot \begin{vmatrix} 6 & 4 \\ 3 & 6 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 1 & 3 \\ 3 & 6 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 6 & 4 \end{vmatrix}$$

$$= 5(6 \cdot 6 - 4 \cdot 3) - 2(1 \cdot 6 - 3 \cdot 3) + (1 \cdot 4 - 3 \cdot 6) = 112$$

- 2열을 선택했을 경우

$$\det(A) = a_{12}A_{12} + a_{22}A_{22} + a_{32}A_{32}$$

$$= 1 \cdot (-1)^{1+2}\det(M_{12}) + 6 \cdot (-1)^{2+2}\det(M_{22}) + 3 \cdot (-1)^{3+2}\det(M_{32})$$

$$= (-1) \cdot \begin{vmatrix} 2 & 4 \\ 1 & 6 \end{vmatrix} + 6 \cdot \begin{vmatrix} 5 & 3 \\ 1 & 6 \end{vmatrix} + (-3) \cdot \begin{vmatrix} 5 & 3 \\ 2 & 4 \end{vmatrix}$$

$$= -(2 \cdot 6 - 4 \cdot 1) + 6(5 \cdot 6 - 3 \cdot 1) - 3(5 \cdot 4 - 3 \cdot 2) = 112$$

- 3열을 선택했을 경우

$$\det(A) = a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33}$$

$$= 3 \cdot (-1)^{1+3}\det(M_{13}) + 4 \cdot (-1)^{2+3}\det(M_{23}) + 6 \cdot (-1)^{3+3}\det(M_{33})$$

$$= 3 \cdot \begin{vmatrix} 2 & 6 \\ 1 & 3 \end{vmatrix} + (-4) \cdot \begin{vmatrix} 5 & 1 \\ 1 & 3 \end{vmatrix} + 6 \cdot \begin{vmatrix} 5 & 1 \\ 2 & 6 \end{vmatrix}$$

$$= 3(2 \cdot 3 - 6 \cdot 1) - 4(5 \cdot 3 - 1 \cdot 1) + 6(5 \cdot 6 - 1 \cdot 2) = 112$$

4. 행렬식

예제 6-12

4차 정사각행렬 $C = \begin{bmatrix} 1 & -5 & 2 & -3 \\ 4 & 2 & -1 & 1 \\ -2 & -4 & 5 & 3 \\ 1 & 0 & -2 & -1 \end{bmatrix}$ 의 행렬식을 구하라.

풀이

1과 0이 포함되어 있는 4행을 선택한다.

$$\begin{aligned} \det(C) &= c_{41}A_{41} + c_{42}A_{42} + c_{43}A_{43} + c_{44}A_{44} \\ &= 1A_{41} + 0A_{42} + (-2)A_{43} + (-1)A_{44} \\ &= A_{41} - 2A_{43} - A_{44} \\ &= (-1)^{4+1}\det(M_{41}) - 2(-1)^{4+3}\det(M_{43}) - (-1)^{4+4}\det(M_{44}) \\ &= -\det(M_{41}) + 2\det(M_{43}) - \det(M_{44}) \end{aligned}$$

$\det(M_{41})$, $\det(M_{43})$, $\det(M_{44})$을 각각 구하면 다음과 같다.

4. 행렬식

- $M_{41} = \begin{bmatrix} -5 & 2 & -3 \\ 2 & -1 & 1 \\ -4 & 5 & 3 \end{bmatrix}$ 2행을 선택하여 $\det(M_{41})$을 구하면

$$\det(M_{41}) = p_{21}A_{21} + p_{22}A_{22} + p_{23}A_{23}$$

$$= 2(-1)^{2+1}\det(M_{21}) + (-1)(-1)^{2+2}\det(M_{22}) + 1(-1)^{2+3}\det(M_{23})$$

$$= (-2)\begin{vmatrix} 2 & -3 \\ 5 & 3 \end{vmatrix} - \begin{vmatrix} -5 & -3 \\ -4 & 3 \end{vmatrix} - \begin{vmatrix} -5 & 2 \\ -4 & 5 \end{vmatrix}$$

$$= (-2)(2 \cdot 3 - (-3) \cdot 5) - ((-5) \cdot 3 - (-3) \cdot (-4))$$

$$-((-5) \cdot 5 - 2 \cdot (-4)) = 2$$

- $M_{43} = \begin{bmatrix} 1 & -5 & -3 \\ 4 & 2 & 1 \\ -2 & -4 & 3 \end{bmatrix}$ 3열을 선택하여 $\det(M_{43})$을 구하면

$$\det(M_{43}) = q_{13}A_{13} + q_{23}A_{23} + q_{33}A_{33}$$

$$= (-3)(-1)^{1+3}\det(M_{13}) + 1(-1)^{2+3}\det(M_{23}) + 3(-1)^{3+3}\det(M_{33})$$

$$= (-3)\begin{vmatrix} 4 & 2 \\ -2 & -4 \end{vmatrix} - \begin{vmatrix} 1 & -5 \\ -2 & -4 \end{vmatrix} + 3\begin{vmatrix} 1 & -5 \\ 4 & 2 \end{vmatrix}$$

$$= (-3)(4 \cdot (-4) - 2 \cdot (-2)) - (1 \cdot (-4) - (-5) \cdot (-2))$$

$$+3(1 \cdot 2 - (-5) \cdot 4) = 116$$

4. 행렬식

- $M_{44} = \begin{bmatrix} 1 & -5 & 2 \\ 4 & 2 & -1 \\ -2 & -4 & 5 \end{bmatrix}$ 1행을 선택하여 $\det(M_{43})$을 구하면

$$\det(M_{44}) = r_{11}A_{11} + r_{12}A_{12} + r_{13}A_{13}$$

$$= 1(-1)^{1+1}\det(M_{11}) + (-5)(-1)^{1+2}\det(M_{12}) + 2(-1)^{1+3}\det(M_{13})$$

$$= \begin{vmatrix} 2 & -1 \\ -4 & 5 \end{vmatrix} + 5\begin{vmatrix} 4 & -1 \\ -2 & 5 \end{vmatrix} + 2\begin{vmatrix} 4 & 2 \\ -2 & -4 \end{vmatrix}$$

$$= (2 \cdot 5 - (-1) \cdot (-4)) + 5(4 \cdot 5 - (-1) \cdot (-2))$$

$$+ 2(4 \cdot (-4) - 2 \cdot (-2)) = 72$$

$$\therefore \det(C) = -\det(M_{41}) + 2\det(M_{43}) - \det(M_{44})$$

$$= -2 + 2 \cdot 116 - 72 = 158$$

5. 역행렬

정의 6-18 역행렬(Inverse Matrix: A^{-1})

정사각행렬 A에 대해 $AB = BA = I$를 만족하게 하는 행렬 B

$$AA^{-1} = A^{-1}A = I$$

예제 6-13

$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$의 역행렬을 구하라.

풀이

2차 정사각행렬 A의 역행렬 $A^{-1} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$라고 할 때,

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} w+2y & x+2z \\ w+3y & x+3z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$w+2y=1,\ x+2z=0,\ w+3y=0,\ x+3z=1$

$(w+2y)-(w+3y)=1-0 \qquad \therefore\ y=-1,\ w=3$

$(x+2z)-(x+3z)=0-1 \qquad \therefore\ z=1,\ x=-2$

$\therefore\ A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

5. 역행렬

정의 6-19 행렬식을 이용한 역행렬

$$A^{-1} = \frac{1}{\det(A)}[A_{ij}]^T \quad (\text{단, } \det(A) \neq 0)$$

정의 6-20 수반행렬(Adjoint Matrix: $[A_{ij}]^T$)

여인수행렬 $[A_{ij}]$에 대한 전치행렬

예제 6-14

행렬 $A = \begin{bmatrix} -2 & 2 & -3 \\ 1 & 0 & 1 \\ -4 & 3 & -5 \end{bmatrix}$의 역행렬을 구하라.

풀이

행렬 A의 행렬식을 구하기 위해 2행 선택

$$\det(A) = a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} = A_{21} + A_{23}$$

$$= (-1)^{2+1}\det(M_{21}) + (-1)^{2+3}\det(M_{23})$$

$$= -\begin{vmatrix} 2 & -3 \\ 3 & -5 \end{vmatrix} - \begin{vmatrix} -2 & 2 \\ -4 & 3 \end{vmatrix} = -[2 \cdot (-5) - (-3) \cdot 3] - [(-2) \cdot 3 - 2 \cdot (-4)] = -1$$

5. 역행렬

여인수행렬 $[A_{ij}]$를 구하기 위해

$$A_{11} = (-1)^{1+1}\det(M_{11}) = \begin{vmatrix} 0 & 1 \\ 3 & -5 \end{vmatrix} = 0 \cdot (-5) - 1 \cdot 3 = -3$$

$$A_{12} = (-1)^{1+2}\det(M_{12}) = -\begin{vmatrix} 1 & 1 \\ -4 & -5 \end{vmatrix} = -[1 \cdot (-5) - 1 \cdot (-4)] = 1$$

$$A_{13} = (-1)^{1+3}\det(M_{13}) = \begin{vmatrix} 1 & 0 \\ -4 & 3 \end{vmatrix} = 1 \cdot 3 - 0 \cdot (-4) = 3$$

$$A_{21} = (-1)^{2+1}\det(M_{21}) = -\begin{vmatrix} 2 & -3 \\ 3 & -5 \end{vmatrix} = -[2 \cdot (-5) - (-3) \cdot 3] = 1$$

$$A_{22} = (-1)^{2+2}\det(M_{22}) = \begin{vmatrix} -2 & -3 \\ -4 & -5 \end{vmatrix} = (-2) \cdot (-5) - (-3) \cdot (-4) = -2$$

$$A_{23} = (-1)^{2+3}\det(M_{23}) = -\begin{vmatrix} -2 & 2 \\ -4 & 3 \end{vmatrix} = -[(-2) \cdot 3 - 2 \cdot (-4)] = -2$$

$$A_{31} = (-1)^{3+1}\det(M_{31}) = \begin{vmatrix} 2 & -3 \\ 0 & 1 \end{vmatrix} = 2 \cdot 1 - (-3) \cdot 0 = 2$$

$$A_{32} = (-1)^{3+2}\det(M_{32}) = -\begin{vmatrix} -2 & -3 \\ 1 & 1 \end{vmatrix} = -[(-2) \cdot 1 - (-3) \cdot 1] = -1$$

5. 역행렬

$$A_{33} = (-1)^{3+3}\det(M_{33}) = \begin{vmatrix} -2 & 2 \\ 1 & 0 \end{vmatrix} = (-2)\cdot 0 - 2\cdot 1 = -2$$

여인수행렬 $[A_{ij}] = \begin{bmatrix} -3 & 1 & 3 \\ 1 & -2 & -2 \\ 2 & -1 & -2 \end{bmatrix}$ 이므로, 수반행렬 $[A_{ij}]^T = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -2 & -1 \\ 3 & -2 & -2 \end{bmatrix}$

$$\therefore\ A^{-1} = -\begin{bmatrix} -3 & 1 & 2 \\ 1 & -2 & -1 \\ 3 & -2 & -2 \end{bmatrix} = \begin{bmatrix} 3 & -1 & -2 \\ -1 & 2 & 1 \\ -3 & 2 & 2 \end{bmatrix}$$

정의 6-21 가역행렬(Invertible Matrix), 특이행렬(Singular Matrix)

- 가역행렬: $\det(A) \neq 0$ 인 행렬, 역행렬이 존재하는 행렬
- 특이행렬: $\det(A) = 0$ 인 행렬, 역행렬이 존재하지 않는 행렬

5. 역행렬

예제 6-15

다음 행렬이 가역행렬인지 특이행렬인지 구분하고, 가역행렬이라면 역행렬을 구하라.

(1) $A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & -3 \\ 4 & 1 & 1 \end{bmatrix}$ (2) $B = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 5 & 6 & 0 \end{bmatrix}$

풀이

(1) $\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = A_{11} - A_{12} + 2A_{13}$

$$= (-1)^{1+1}\det(M_{11}) - (-1)^{1+2}\det(M_{12}) + 2(-1)^{1+3}\det(M_{13})$$

$$= \begin{vmatrix} 1 & -3 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 2 & -3 \\ 4 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 1 \\ 4 & 1 \end{vmatrix}$$

$$= (1 \cdot 1 - (-3) \cdot 1) + (2 \cdot 1 - (-3) \cdot 4) + 2(2 \cdot 1 - 1 \cdot 4) = 14$$

따라서 $\det(A) \neq 0$이므로 역행렬을 구할 수 있다. 즉 가역행렬이다. 여인수행렬을 구하면,

$$A_{11} = (-1)^{1+1}\det(M_{11}) = \begin{vmatrix} 1 & -3 \\ 1 & 1 \end{vmatrix} = 1 \cdot 1 - (-3) \cdot 1 = 4$$

$$A_{12} = (-1)^{1+2}\det(M_{12}) = -\begin{vmatrix} 2 & -3 \\ 4 & 1 \end{vmatrix} = -(2 \cdot 1 - (-3) \cdot 4) = -14$$

$$A_{13} = (-1)^{1+3}\det(M_{13}) = \begin{vmatrix} 2 & 1 \\ 4 & 1 \end{vmatrix} = 2 \cdot 1 - 1 \cdot 4 = -2$$

$$A_{21} = (-1)^{2+1}\det(M_{21}) = -\begin{vmatrix} -1 & 2 \\ 1 & 1 \end{vmatrix} = -((-1) \cdot 1 - 2 \cdot 1) = 3$$

5. 역행렬

$$A_{22}=(-1)^{2+2}\det(M_{22})=\begin{vmatrix}1 & 2\\4 & 1\end{vmatrix}=1\cdot 1-2\cdot 4=-7$$

$$A_{23}=(-1)^{2+3}\det(M_{23})=-\begin{vmatrix}1 & -1\\4 & 1\end{vmatrix}=-(1\cdot 1-(-1)\cdot 4)=-5$$

$$A_{31}=(-1)^{3+1}\det(M_{31})=\begin{vmatrix}-1 & 2\\1 & -3\end{vmatrix}=(-1)\cdot(-3)-2\cdot 1=1$$

$$A_{32}=(-1)^{3+2}\det(M_{32})=-\begin{vmatrix}1 & 2\\2 & -3\end{vmatrix}=-(1\cdot(-3)-2\cdot 2)=7$$

$$A_{33}=(-1)^{3+3}\det(M_{33})=\begin{vmatrix}1 & -1\\2 & 1\end{vmatrix}=1\cdot 1-(-1)\cdot 2=3$$

여인수행렬 $[A_{ij}]=\begin{bmatrix}4 & -14 & -2\\3 & -7 & -5\\1 & 7 & 3\end{bmatrix}$, 수반행렬 $[A_{ij}]^T=\begin{bmatrix}4 & 3 & 1\\-14 & -7 & 7\\-2 & -5 & 3\end{bmatrix}$

$$\therefore\ A^{-1}=\frac{1}{14}\begin{bmatrix}4 & 3 & 1\\-14 & -7 & 7\\-2 & -5 & 3\end{bmatrix}$$

(2) $\det(B)=b_{11}A_{11}+b_{12}A_{12}+b_{13}A_{13}=A_{11}+2A_{12}$

$$=(-1)^{1+1}\det(M_{11})+2(-1)^{1+2}\det(M_{12})$$

$$=\begin{vmatrix}4 & 0\\6 & 0\end{vmatrix}-2\begin{vmatrix}3 & 0\\5 & 0\end{vmatrix}=(4\cdot 0-0\cdot 6)-2(3\cdot 0-0\cdot 5)=0$$

$\therefore\ \det(A)=0$이므로 역행렬을 구할 수 없다. 즉 특이행렬이다.

2.1.1 벡터와 행렬

벡터

- 샘플을 특징 벡터로feature vector 표현
- 예) Iris 데이터에서 꽃받침의 길이, 꽃받침의 너비, 꽃잎의 길이, 꽃잎의 너비라는 4개의 특징이 각각 5.1, 3.5, 1.4, 0.2인 샘플

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}$$

- 여러 개의 특징 벡터를 첨자로 구분

$$\mathbf{x}_1 = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 4.9 \\ 3.0 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 4.7 \\ 3.2 \\ 1.3 \\ 0.2 \end{pmatrix}, \cdots, \mathbf{x}_{150} = \begin{pmatrix} 5.9 \\ 3.0 \\ 5.1 \\ 1.8 \end{pmatrix}$$

2.1.1 벡터와 행렬

행렬

- 여러 개의 벡터를 담음
- 훈련집합을 담은 행렬을 설계행렬이라 부름
- 예) Iris 데이터에 있는 150개의 샘플을 설계 행렬 $\mathbf{X}$로 표현

$$\mathbf{X} = \begin{pmatrix} 5.1 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 3.2 & 1.3 & 0.2 \\ 4.6 & 3.1 & 1.5 & 0.2 \\ \vdots & \vdots & \vdots & \vdots \\ 6.2 & 3.4 & 5.4 & 2.3 \\ 5.9 & 3.0 & 5.1 & 1.8 \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \\ \vdots & \vdots & \vdots & \vdots \\ x_{149,1} & x_{149,2} & x_{149,3} & x_{149,4} \\ x_{150,1} & x_{150,2} & x_{150,3} & x_{150,4} \end{pmatrix}$$

행row

열column

2.1.1 벡터와 행렬

- 행렬 $\mathbf{A}$의 전치행렬 $\mathbf{A}^{\mathrm{T}}$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \quad \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}$$

예를 들어, $\mathbf{A} = \begin{pmatrix} 3 & 4 & 1 \\ 0 & 5 & 2 \end{pmatrix}$라면 $\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \\ 1 & 2 \end{pmatrix}$

  - Iris의 설계 행렬을 전치행렬 표기에 따라 표현하면,

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_{150}^{\mathrm{T}} \end{pmatrix}$$

2.1.1 벡터와 행렬

- 행렬을 이용하면 수학을 간결하게 표현할 수 있음
  - 예) 다항식의 행렬 표현

$$\begin{aligned} f(\mathbf{x}) &= f(x_1, x_2, x_3) \\ &= 2x_1x_1 - 4x_1x_2 + 3x_1x_3 + x_2x_1 + 2x_2x_2 + 6x_2x_3 - 2x_3x_1 + 3x_3x_2 + 2x_3x_3 + 2x_1 + 3x_2 - 4x_3 + 5 \\ &= (x_1 \quad x_2 \quad x_3)\begin{pmatrix} 2 & -4 & 3 \\ 1 & 2 & 6 \\ -2 & 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (2 \quad 3 \quad -4)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + 5 \\ &= \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c \end{aligned}$$

- 특수한 행렬들

$$\text{정사각행렬}\begin{pmatrix} 2 & 0 & 1 \\ 1 & 21 & 5 \\ 4 & 5 & 12 \end{pmatrix}, \quad \text{대각행렬}\begin{pmatrix} 50 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8 \end{pmatrix},$$

$$\text{단위행렬}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \text{대칭행렬}\begin{pmatrix} 1 & 2 & 11 \\ 2 & 21 & 5 \\ 11 & 5 & 1 \end{pmatrix}$$

2.1.1 벡터와 행렬

행렬 연산

- 행렬 곱셈 $\mathbf{C} = \mathbf{AB}$, 이때 $c_{ij} = \sum_{k=1,s} a_{ik} b_{kj}$ (2.1)

2*3 행렬 $\mathbf{A} = \begin{pmatrix} 3 & 4 & 1 \\ 0 & 5 & 2 \end{pmatrix}$와 3*3행렬 $\mathbf{B} = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 0 & 5 \\ 4 & 5 & 1 \end{pmatrix}$을 곱하면 2*3 행렬 $\mathbf{C} = \mathbf{AB} = \begin{pmatrix} 14 & 5 & 24 \\ 13 & 10 & 27 \end{pmatrix}$

  - 교환법칙 성립하지 않음: $\mathbf{AB} \neq \mathbf{BA}$
  - 분배법칙과 결합법칙 성립: $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$이고 $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$

- 벡터의 내적

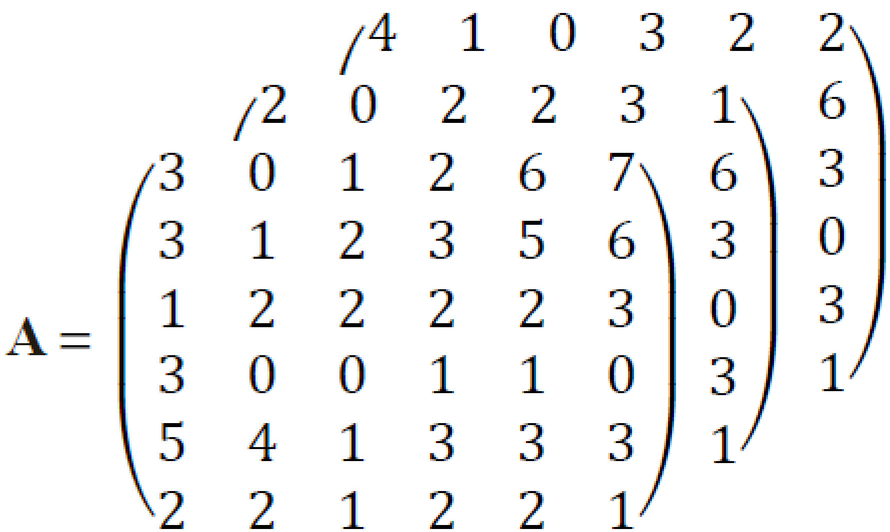
벡터의 내적 $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = \sum_{k=1,d} a_k b_k$ (2.2)

$\mathbf{x}_1 = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}$와 $\mathbf{x}_2 = \begin{pmatrix} 4.9 \\ 3.0 \\ 1.4 \\ 0.2 \end{pmatrix}$의 내적 $\mathbf{x}_1 \cdot \mathbf{x}_2$는 37.49

2.1.1 벡터와 행렬

텐서

- 3차원 이상의 구조를 가진 숫자 배열
- 예) 3차원 구조의 RGB 컬러 영상



2.1.2 놈과 유사도

- 벡터와 행렬의 크기를 놈으로 측정
  - 벡터의 p차 놈

$$p\text{차 놈: } \|\mathbf{x}\|_p = \left(\sum_{i=1,d} |x_i|^p\right)^{\frac{1}{p}} \tag{2.3}$$

$$\text{최대 놈: } \|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|, \cdots, |x_d|) \tag{2.4}$$

  - 예) $\mathbf{x} = (3 \ \ -4 \ \ 1)$ 일 때, 2차 놈은 $\|\mathbf{x}\|_2 = (3^2 + (-4)^2 + 1^2)^{1/2} = 5.099$

  - 행렬의 프로베니우스 놈

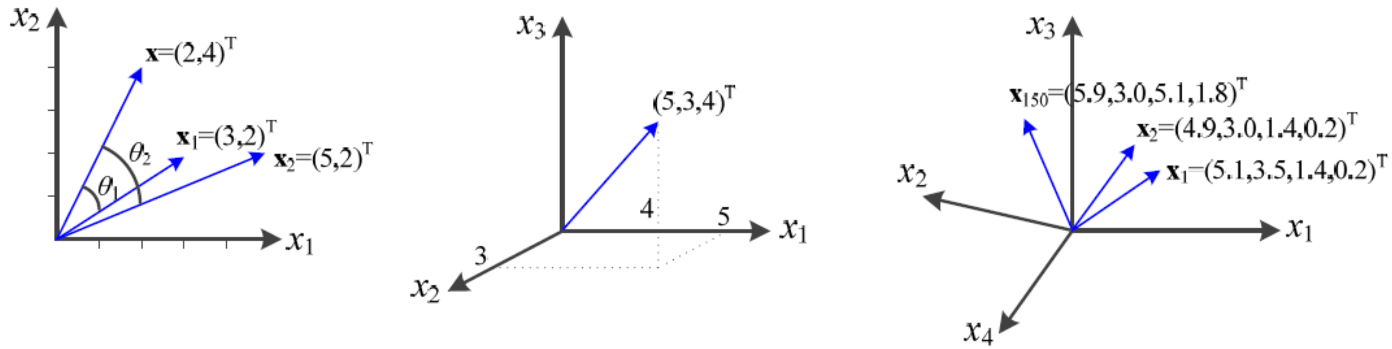
$$\text{프로베니우스 놈: } \|\mathbf{A}\|_F = \left(\sum_{i=1,n} \sum_{j=1,m} a_{ij}^2\right)^{\frac{1}{2}} \tag{2.6}$$

예를 들어, $\left\| \begin{pmatrix} 2 & 1 \\ 6 & 4 \end{pmatrix} \right\|_F = \sqrt{2^2 + 1^2 + 6^2 + 4^2} = 7.550$

2.1.2 놈과 유사도

유사도와 거리

- 벡터를 기하학적으로 해석

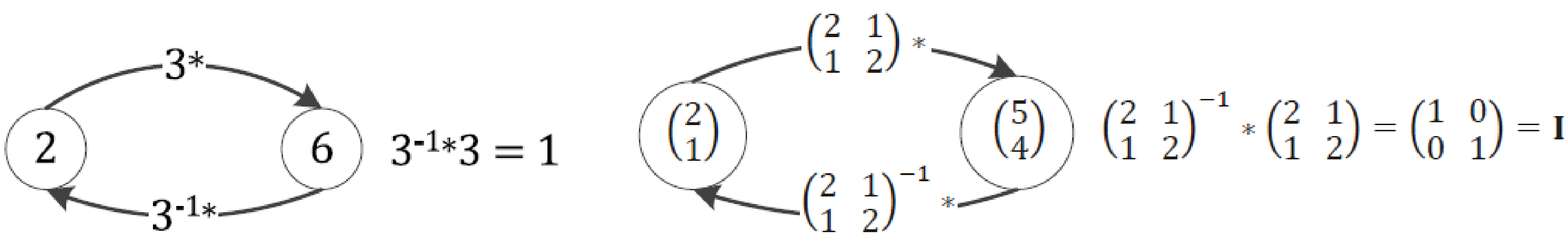


(a) 2차원 벡터 (b) 3차원 벡터 (c) 4차원 벡터(Iris 데이터)

그림 2-2 벡터를 기하학적으로 해석

2.1.5 역행렬

■ 역행렬의 원리



(a) 역수의 원리

(b) 역행렬의 원리

그림 2-9 역행렬

- 정사각행렬 **A**의 역행렬 $\mathbf{A}^{-1}$

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

- 예를 들어, $\begin{pmatrix} 2 & 1 \\ 6 & 4 \end{pmatrix}$의 역행렬은 $\begin{pmatrix} 2 & -0.5 \\ -3 & 1 \end{pmatrix}$

2.1.5 역행렬

정리

정리 2-1 다음 성질은 서로 필요충분조건이다.

- $\mathbf{A}$는 역행렬을 가진다. 즉, 특이행렬이 아니다.
- $\mathbf{A}$는 최대계수를 가진다.
- $\mathbf{A}$의 모든 행이 선형독립이다.
- $\mathbf{A}$의 모든 열이 선형독립이다.
- $\mathbf{A}$의 행렬식은 0이 아니다.
- $\mathbf{A}^{\mathrm{T}}\mathbf{A}$는 양의 정부호positive definite 대칭 행렬이다.
- $\mathbf{A}$의 고윳값은 모두 0이 아니다.

2.1.5 역행렬

- 행렬 **A**의 행렬식 *det*(**A**)

$$\left.\begin{aligned} det\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= ad - bc \\ det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= aei + bfg + cdh - ceg - bdi - afh \end{aligned}\right\} \qquad (2.15)$$

예를 들어 $\begin{pmatrix} 2 & 1 \\ 6 & 4 \end{pmatrix}$의 행렬식은 $2*4-1*6=2$

- 기하학적 의미
  - 2차원에서는 2개의 행 벡터가 이루는 평행사변형의 넓이
  - 3차원에서는 3개의 행 벡터가 이루는 평행사각기둥의 부피

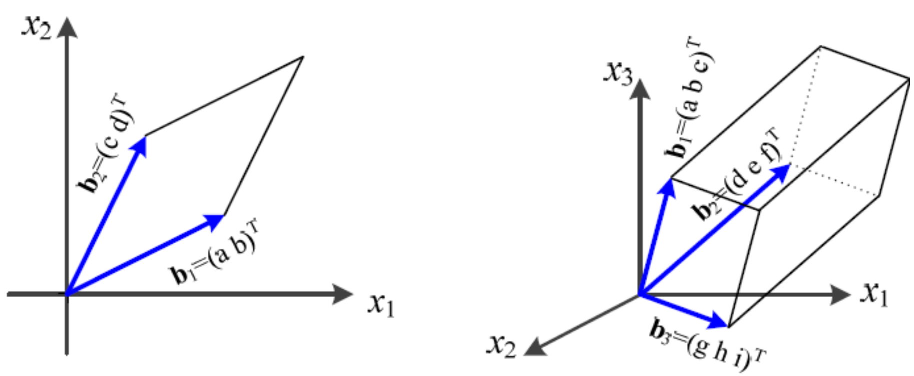


그림 2-10 행렬식의 기하학적 해석

2.2.1 확률 기초

확률변수random variable

- 예) 윷

그림 2-13 윷을 던졌을 때 나올 수 있는 다섯 가지 경우(왼쪽부터 도, 개, 걸, 윷, 모)

- 다섯 가지 경우 중 한 값을 갖는 확률변수 x
- x의 정의역은 {도, 개, 걸, 윷, 모}

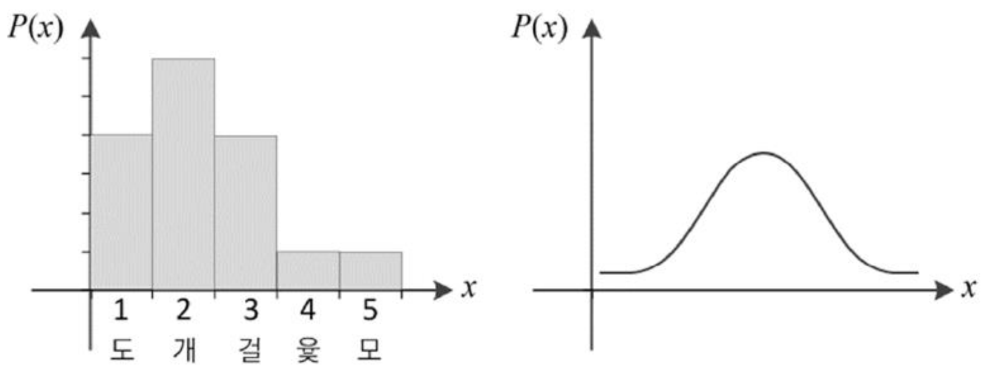
○를 배, ●를 등이라 할 때, 네 쪽의 윷짝을 던졌을 때 나타날 수 있는 총 경우의 수는 다음과 같다. -.

1) ○○○○	2) ○○○●	3) ○○●○	4) ○○●●	5) ○●○○	6) ○●○●
7) ○●●○	8) ○●●●	9) ●○○○	10) ●○○●	11) ●○●○	12) ●○●●
13) ●●○○	14) ●●○●	15) ●●●○	16) ●●●●		

2.2.1 확률 기초

확률분포

$$P(x = \text{도}) = \frac{4}{16}, P(x = \text{개}) = \frac{6}{16}, P(x = \text{걸}) = \frac{4}{16}, P(x = \text{윷}) = \frac{1}{16}, P(x = \text{모}) = \frac{1}{16}$$



(a) 이산인 경우의 확률질량함수

(b) 연속인 경우의 확률밀도함수

그림 2-14 확률분포

확률벡터random vector

- 예) Iris에서 확률벡터 $\mathbf{x}$는 4차원 $\mathbf{x} = (x_1, x_2, x_3, x_4)^\mathrm{T} = (\text{꽃받침 길이}, \text{꽃받침 너비}_1, \text{꽃잎}$

2.2.1 확률 기초

■ 간단한 확률실험 장치

- 주머니에서 번호를 뽑은 다음, 번호에 따라 해당 병에서 공을 뽑고 색을 관찰함
- 번호를 y, 공의 색을 x라는 확률변수로 표현하면 정의역은 $y \in$ {①,②,③}, $x \in$ {파랑, 하양}

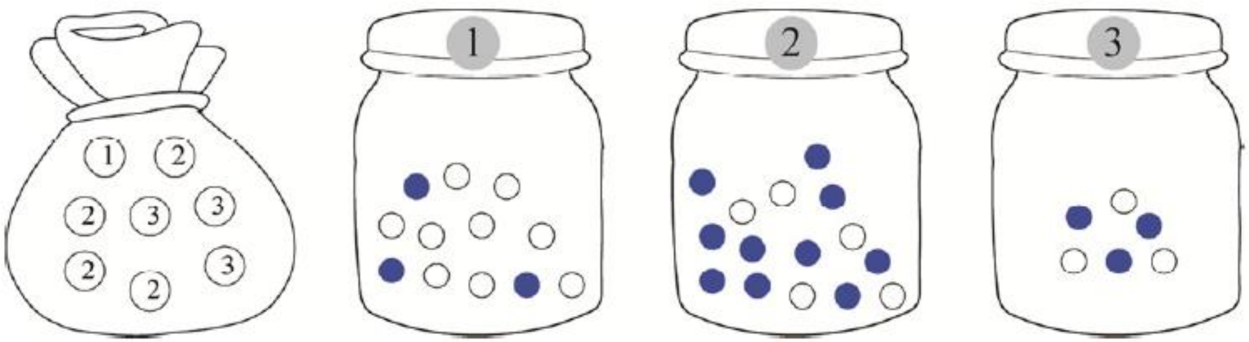


그림 2-15 확률 실험

2.2.1 확률 기초

곱 규칙과 합 규칙

- ①번 카드를 뽑을 확률은 $P(y=①)=P(①)=1/8$
- 카드는 ①번, 공은 하양일 확률은 $P(y=①,x=\text{하양})=P(①,\text{하양})$ ← 결합확률

$$P(y=①, x=\text{하양}) = P(x=\text{하양}|y=①)P(y=①) = \frac{9}{12}\frac{1}{8} = \frac{3}{32}$$

- 곱 규칙

$$\text{곱 규칙: } P(y,x) = P(x|y)P(y) \tag{2.23}$$

- 하얀 공이 뽑힐 확률

$$P(\text{하양}) = P(\text{하양}|①)P(①) + P(\text{하양}|②)P(②) + P(\text{하양}|③)P(③)$$
$$= \frac{9}{12}\frac{1}{8} + \frac{5}{15}\frac{4}{8} + \frac{3}{6}\frac{3}{8} = \frac{43}{96}$$

- 합 규칙

$$\text{합 규칙: } P(x) = \sum_y P(y,x) = \sum_y P(x|y)P(y) \tag{2.24}$$

2.2.2 베이즈 정리와 기계 학습

- 베이즈 정리 : 사전지식에서 사후 확률을 구할 수 있는 기법
- $P(y,x) = P(x|y)P(y) = P(x,y) = P(y|x)P(x)$

$$\longrightarrow \quad P(y|x) = \frac{P(x|y)P(y)}{P(x)} \qquad (2.26)$$

- "하얀 공이 나왔다는 사실만 알고 어느 병에서 나왔는지 모르는데, 어느 병인지 추정하라."

$$\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x) \qquad (2.27)$$

2.2.2 베이즈 정리와 기계 학습

- 베이즈 정리 (식 (2.26))
  - 베이즈 정리를 적용하면, $\hat{y} = \underset{y}{\operatorname{argmax}}\, P(y|x = \text{하양}) = \underset{y}{\operatorname{argmax}} \dfrac{P(x = \text{하양}|y)P(y)}{P(x = \text{하양})}$
  - 세 가지 경우에 대해 확률을 계산하면,

$$P(①|\text{하양}) = \frac{P(\text{하양}|①)P(①)}{P(\text{하양})} = \frac{\frac{9}{12}\frac{1}{8}}{\frac{43}{96}} = \frac{9}{43}$$

$$P(②|\text{하양}) = \frac{P(\text{하양}|②)P(②)}{P(\text{하양})} = \frac{\frac{5}{15}\frac{4}{8}}{\frac{43}{96}} = \frac{16}{43}$$

③번 병일 확률이 가장 높음

$$P(③|\text{하양}) = \frac{P(\text{하양}|③)P(③)}{P(\text{하양})} = \frac{\frac{3}{6}\frac{3}{8}}{\frac{43}{96}} = \frac{18}{43}$$

- 베이즈 정리의 해석

$$\overbrace{P(y|x)}^{\text{사후확률}} = \frac{\overbrace{P(x|y)}^{\text{우도}}\ \overbrace{P(y)}^{\text{사전확률}}}{P(x)}$$

2.2.2 베이즈 정리와 기계 학습

기계 학습에 적용

- 예) Iris 데이터 분류 문제
  - 특징 벡터 **x**, 부류 $y \in$ {setosa, versicolor, virginica}
  - 분류 문제를 argmax로 표현하면 식 (2.29)

$$\hat{y} = \underset{y}{\operatorname{argmax}} P(y|\mathbf{x}) \tag{2.29}$$

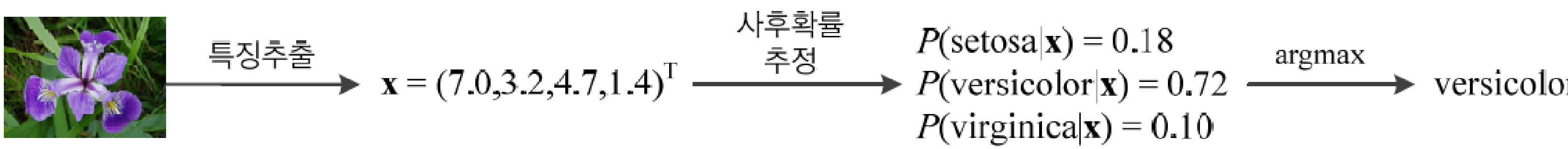


그림 2-16 붓꽃의 부류 예측 과정

- 사후확률 $P(y|\mathbf{x})$를 직접 추정하는 일은 아주 단순한 경우를 빼고 불가능
- 따라서 베이즈 정리를 이용하여 추정함
  - 사전확률은 식 (2.30)으로 추정

$$\text{사전확률:} \quad P(y = c_i) = \frac{n_i}{n} \tag{2.30}$$

2.2.4 평균과 분산

- 데이터의 요약 정보로서 평균과 분산

$$\left.\begin{aligned} \text{평균} \quad \mu &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ \text{분산} \quad \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \end{aligned}\right\} \tag{2.36}$$

- 평균 벡터와 공분산 행렬

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \tag{2.37}$$

$$\boldsymbol{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i-\boldsymbol{\mu})(\mathbf{x}_i-\boldsymbol{\mu})^{\mathrm{T}} \tag{2.39}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & & \sigma_{2d} \\ \vdots & & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_{dd} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & & \sigma_{2d} \\ \vdots & & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_d^2 \end{pmatrix}$$

2.2.4 평균과 분산

평균 벡터와 공분산 행렬 예제

예제 2-7

Iris 데이터베이스의 샘플 중 8개만 가지고 공분산 행렬을 계산하자.

$$\mathbb{X} = \left\{ \mathbf{x}_1 = \begin{pmatrix} 5.1 \\ 3.5 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 4.9 \\ 3.0 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 4.7 \\ 3.2 \\ 1.3 \\ 0.2 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 4.6 \\ 3.1 \\ 1.5 \\ 0.2 \end{pmatrix}, \mathbf{x}_5 = \begin{pmatrix} 5.0 \\ 3.6 \\ 1.4 \\ 0.2 \end{pmatrix}, \mathbf{x}_6 = \begin{pmatrix} 5.4 \\ 3.9 \\ 1.7 \\ 0.4 \end{pmatrix}, \mathbf{x}_7 = \begin{pmatrix} 4.6 \\ 3.4 \\ 1.4 \\ 0.3 \end{pmatrix}, \mathbf{x}_8 = \begin{pmatrix} 5.0 \\ 3.4 \\ 1.5 \\ 0.2 \end{pmatrix} \right\}$$

먼저 평균벡터를 구하면 $\boldsymbol{\mu} = (4.9125,\ 3.3875,\ 1.45,\ 0.2375)^{\mathrm{T}}$이다. 첫 번째 샘플 $\mathbf{x}_1$을 식 (2.39)에 적용하면 다음과 같다.

$$(\mathbf{x}_1 - \boldsymbol{\mu})(\mathbf{x}_1 - \boldsymbol{\mu})^{\mathrm{T}} = \begin{pmatrix} 0.1875 \\ 0.1125 \\ -0.05 \\ -0.0375 \end{pmatrix} (0.1875 \quad 0.1125 \quad -0.05 \quad -0.0375)$$

$$= \begin{pmatrix} 0.0325 & 0.0211 & -0.0094 & -0.0070 \\ 0.0211 & 0.0127 & -0.0056 & -0.0042 \\ -0.0094 & -0.0056 & 0.0025 & 0.0019 \\ -0.0070 & -0.0042 & 0.0019 & 0.0014 \end{pmatrix}$$

나머지 7개 샘플도 같은 계산을 한 다음, 결과를 모두 더하고 8로 나누면 다음과 같은 공분산 행렬을 얻는다.

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.0661 & 0.0527 & 0.0181 & 0.0083 \\ 0.0527 & 0.0736 & 0.0181 & 0.0130 \\ 0.0181 & 0.0181 & 0.0125 & 0.0056 \\ 0.0083 & 0.0130 & 0.0056 & 0.0048 \end{pmatrix}$$

2.3.1 매개변수 공간의 탐색

- [알고리즘 2-3]은 기계 학습이 사용하는 전형적인 알고리즘
  - 라인 3에서는 목적함수가 작아지는 방향을 주로 미분으로 찾아냄

알고리즘 2-3 기계 학습이 사용하는 전형적인 탐색 알고리즘(1장의 [알고리즘 1-1]과 같음)

입력 : 훈련집합 $\mathbb{X}$와 $\mathbb{Y}$

출력 : 최적해 $\hat{\mathbf{\Theta}}$

1 난수를 생성하여 초기해 $\mathbf{\Theta}$을 설정한다.
2 repeat
3     $J(\mathbf{\Theta})$가 작아지는 방향 $d\mathbf{\Theta}$를 구한다.
4     $\mathbf{\Theta} = \mathbf{\Theta} + d\mathbf{\Theta}$
5 until(멈춤 조건)
6 $\hat{\mathbf{\Theta}} = \mathbf{\Theta}$