A.4.1 행렬 (review)

■ 행렬(Matrix)

a_{ij} : $\boldsymbol{A}$의 i번째 행과 j번째 열에 있는 숫자를 (i, j)번째 요소/원소(element)

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

■ 벡터(vector)

$m = 1$ 행벡터

$n = 1$ 열벡터

1

A.4.1 행렬 (review)

– 두 벡터 $\boldsymbol{a}_{m\times 1}$, $\boldsymbol{b}_{n\times 1}$의 곱

내적(inner product): $m = n$이면

$$\boldsymbol{a}^T\boldsymbol{b} = a_1b_1 + a_2b_2 + \cdots + a_mb_m = \sum_{i=1}^{m} a_ib_j$$

$\boldsymbol{a}^T\boldsymbol{b} = 0$ 이면 $\boldsymbol{a}$와 $\boldsymbol{b}$는 직교

외적(outer product): 모든 m, n에 대해

$$\boldsymbol{a}\boldsymbol{b}^T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} (b_1\, b_2 \cdots b_n) = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_mb_1 & a_mb_2 & \cdots & a_2m_n \end{pmatrix} = (a_ib_j)$$

2

A.4.1 행렬 (review)

■ 특수한 행렬

1) 정방행렬(square matrix):

2) 대각행렬(diagonal matrix): $D = diag(d_1, d_2, \ldots, d_m)$

3) 단위행렬(identity matrix): $I = diag(1, 1, \ldots, 1)$

4) 대칭행렬(symmetric matrix): $A^T = A$

▪ 멱등행렬(idempotent matrix): $A^2 = AA = A$

▪ 직교행렬 : $A^{-1} = A^T$

3

A.4.1 행렬 (review)

■ 행렬의 연산

3) 행렬의 전치(transpose)

$B = A^T$이면 모든 i, j에 대해 $b_{ij} = a_{ji}$

$(A^T)^T = A$

$(AB)^T = B^T A^T$

4) 행렬의 곱(multiplication)

$A_{m \times r} B_{r \times n} = C_{m \times n}$

c_{ij}는 A의 i번째 행벡터와 B의 j번째 열벡터의 내적 – 1

$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}$

일반적으로 $AB \neq BA$

4

A.4.1 행렬 (review)

■ 선형종속 및 행렬의 계수

$n \times 1$인 m개의 벡터 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$와 m개의 실수 $c_1, c_2, \cdots, c_m$

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_m\boldsymbol{a}_m = \boldsymbol{0}$$

모든 c_i가 0인 경우가 아니어도 성립하면 m개의 벡터: **선형종속**

만약 모든 c_i가 0인 경우에만 성립하면 m개의 벡터: **선형독립**

- A의 **계수**는 선형독립의 관계에 있는 열벡터(행벡터)들의 최대수

$r(A_{n\times m}) = MIN(m,n)$

: $A_{n\times m}$의 계수는 m이나 n의 두 수 중 작은 수를 초과하지 못함

5

A.4.1 행렬 (review)

■ 역행렬(inverse marix)

$A_{m\times m}$에 대해 $BA = AB = I$ 의 조건이 만족하면

$B_{m\times m}$는 A의 역행렬. $B = A^{-1}$ 이라 표시

(1) $(AB)^{-1} = B^{-1}A^{-1}$

(2) $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$

(3) $(A^T)^{-1} = (A^{-1})^T$

(4) $\det(A^{-1}) = 1/(\det(A))$: $\det(A)$는 A의 행렬식

행렬 $A_{m\times m}$의 역행렬이 존재하기 위해서는

- $r(A) = m$
- $\det(A) \neq 0$

6

A.4.1 행렬 (review)

- 연립방정식은 행렬과 벡터를 이용하여 표현가능

  $\boldsymbol{Ax} = \boldsymbol{b}$ 로 표현

  좌변과 우변에 $\boldsymbol{A}^{-1}$를 곱하여 구할 수 있음

  $\boldsymbol{A}^{-1}\boldsymbol{Ax} = \boldsymbol{Ix} = \boldsymbol{A}^{-1}\boldsymbol{b}$

■ 대각합(Trace)

: $\boldsymbol{A}_{m\times m}$의 대각합은 $tr(\boldsymbol{A})$로 표시

$$tr(\boldsymbol{A}) = \sum_{i=1}^{m} a_{ii}$$

(1) $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$

(2) $\boldsymbol{A}$가 멱등행렬이면 $\boldsymbol{A}$ $r(\boldsymbol{A}) = tr(\boldsymbol{A})$

7

A.4.1 행렬 (review)

■ 행렬식(determinant)

: $\boldsymbol{A}_{m\times m}$의 행렬식은 $\det(\boldsymbol{A})$ 또는 $|\boldsymbol{A}|$로 표시

여인수(cofactor): $c_{ij} = (-1)^{i+j}|\boldsymbol{A}_{ij}|$

$$adj\boldsymbol{A} = \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{pmatrix}$$

행렬 $\boldsymbol{A}$의 역행렬 $\boldsymbol{A}^{-1}$이 존재하면

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}(adj\boldsymbol{A})$$

$\boldsymbol{A}_{m\times m}$의 역행렬 $\boldsymbol{A}^{-1}$이 존재하기 위해서는 $|\boldsymbol{A}| \neq 0$

8

A.4.1 행렬 (review)

■ 고유값 및 고유벡터 (Eigenvalue and Eigenvector)

$A_{m \times m}$와 영벡터가 아닌 벡터 $\boldsymbol{x}$에 대해 다음이 만족 될 때

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

λ: 고유값(eigenvalue)

$\boldsymbol{x}$: λ에 대응하는 고유벡터(eigenvector)

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$$

특성방정식(characteristic equation)

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$$

특성방정식은 λ의 m차 방정식이며 m개의 근이 고유값 $\lambda_1, \lambda_2, \cdots, \lambda_m$

각각의 λ_i에 대해 $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$를 만족하는 $\boldsymbol{x}_i$가 λ_i에 대응하는 고유벡터

9

A.4.1 행렬 (review)

(1) A의 모든 원소가 실수라도 행렬의 고유값은 실수가 아닐 수 있음

(2) A의 역행렬이 존재하기 위해서는 고유값은 하나라도 0이면 안됨

(3) A의 모든 원소가 실수이고 **대칭행렬**이라면 고유값은 반드시 실수이고 대응하는 고유벡터도 원소가 모두 실수인 실수벡터

(4) A가 **실수대칭행렬**이면

$$P^T A P = D$$

를 만족시키는 직교행렬 P가 존재, $D = diag(\lambda_1, \lambda_2, \cdots, \lambda_m)$

$$A = PDP^T$$

행렬 P의 열은 D의 대각원소인 고유값에 해당하는 고유벡터

10

A.4.2 벡터미분법 (review)

■ 벡터미분법

스칼라 a와 $n\times 1$벡터 $\boldsymbol{x}$

$$\frac{\partial a}{\partial \boldsymbol{x}} = \begin{pmatrix} \partial a/\partial x_1 \\ \partial a/\partial x_2 \\ \vdots \\ \partial a/\partial x_n \end{pmatrix}$$

1) 벡터 $\boldsymbol{a}$와 $\boldsymbol{x}$의 내적

$$\boldsymbol{a}^T\boldsymbol{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

$$\frac{\partial(\boldsymbol{a}^T\boldsymbol{x})}{\partial \boldsymbol{x}} = \frac{\partial(\boldsymbol{x}^T\boldsymbol{a})}{\partial \boldsymbol{x}} = \begin{pmatrix} \partial(\boldsymbol{a}^T\boldsymbol{x})/\partial x_1 \\ \partial(\boldsymbol{a}^T\boldsymbol{x})/\partial x_2 \\ \vdots \\ \partial(\boldsymbol{a}^T\boldsymbol{x})/\partial x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{a}$$

11

A.4.2 벡터미분법 (review)

2) 이차형식(quadratic form)

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = (x_1 x_2 \cdots x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\,x_i x_j$$

$\boldsymbol{A}$가 대칭행렬이면 $a_{ij} = a_{ji},\ i,j = 1,2,\cdots,n$

$$\partial(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x})/\partial\boldsymbol{x} = 2\boldsymbol{A}\boldsymbol{x}$$

12

A.4.2 벡터미분법 (review)

예)

행렬 $\boldsymbol{X}_{n\times k}$, 벡터 $\boldsymbol{b}_{k\times 1}$

$\boldsymbol{X}^T\boldsymbol{X}$: $k\times k$ 대칭행렬

$\boldsymbol{b}^T(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{b}$: 이차형식

$$\frac{\partial \boldsymbol{b}^T(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{b}}{\partial \boldsymbol{b}} = 2(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{b} \qquad \frac{\partial(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\boldsymbol{A}\boldsymbol{x}$$

$\boldsymbol{b}^T\boldsymbol{X}^T\boldsymbol{y} = \boldsymbol{b}^T(\boldsymbol{X}^T\boldsymbol{y})$: $\boldsymbol{X}^T\boldsymbol{y}$ 와 $\boldsymbol{b}$의 내적

$$\frac{\partial \boldsymbol{b}^T\boldsymbol{X}^T\boldsymbol{y}}{\partial \boldsymbol{b}} = \frac{\partial \boldsymbol{b}^T(\boldsymbol{X}^T\boldsymbol{y})}{\partial \boldsymbol{b}} = \boldsymbol{X}^T\boldsymbol{y} \qquad \frac{\partial(\boldsymbol{x}^T\boldsymbol{a})}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

13

A.4.3 확률벡터의 기대값과 분산 (review)

■ 확률벡터의 기대값과 분산

- **확률벡터(random vector)**: 벡터를 이루고 있는 요소가 확률변수

  $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^T$

- **기대값벡터**

  $E(\boldsymbol{y}) = (E(y_1), E(y_2), \cdots, E(y_n))^T$

- **분산-공분산행렬** (줄여서 분산행렬)

$$Var(\boldsymbol{y}) = E[(\boldsymbol{y}-E(\boldsymbol{y}))(\boldsymbol{y}-E(\boldsymbol{y}))^T] = \begin{pmatrix} Var(y_1) & Cov(y_1,y_2) & \cdots & Cov(y_1,y_n) \\ Cov(y_2,y_1) & Var(y_2) & \cdots & \vdots \\ \vdots & & \ddots & \\ Cov(y_n,y_1) & & \cdots & Var(y_n) \end{pmatrix}$$

$-Cov(y_i, y_j) = Cov(y_j, y_i)$이므로 대칭행렬

14

A.4.3 확률벡터의 기대값과 분산 (review)

$$E(\epsilon_i) = 0, \quad Var(\epsilon_i) = \sigma^2, \quad i = 1, \ldots, n$$
$$Cov(\epsilon_i, \epsilon_j) = 0, \quad i \neq j$$

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$E(\boldsymbol{\epsilon}) = \begin{pmatrix} E(\epsilon_1) \\ E(\epsilon_2) \\ \vdots \\ E(\epsilon_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}$$

$$Var(\boldsymbol{\epsilon}) = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 \boldsymbol{I}$$

15

A.4.1 행렬 (review)

- $H = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$

$$(\boldsymbol{X}^T\boldsymbol{X})^T = \boldsymbol{X}^T\boldsymbol{X}$$

$$\begin{aligned} \left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\right]^T &= (\boldsymbol{X}^T)^T\left[(\boldsymbol{X}^T\boldsymbol{X})^{-1}\right]^T\boldsymbol{X}^T \\ &= \boldsymbol{X}\left[(\boldsymbol{X}^T\boldsymbol{X})^T\right]^{-1}\boldsymbol{X}^T \\ &= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \qquad \text{(대칭행렬)} \end{aligned}$$

$$\begin{aligned} \left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\right]\left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\right] &= \boldsymbol{X}\left[(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{X}\right](\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \\ &= \boldsymbol{X}\boldsymbol{I}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \\ &= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \qquad \text{(멱등행렬)} \end{aligned}$$

16