Linear Algebra for Statistics

Chapter 5

Instructor: Seokho Lee (이석호)

Hankuk University of Foreign Studies

Chapter 5 특이한 행렬

5.1 대칭행렬(Symmetric Matrix)

Definition

정방행렬 $A_{n\times n}$ 에 대하여

$$A^T = A \tag{5.1}$$

즉,

$$a_{ij} = a_{ji}, \quad (\forall\, i, j = 1, \ldots, n) \tag{5.2}$$

이면, A 는 대칭행렬(symmetric matrix)이다.

Theorem (정리 5.1)

대칭행렬에 대해 다음이 성립한다.

(1) A, B 가 대칭행렬이면 $(AB)^T = BA$

(2) $A_{n\times n}$ 에 대하여, $A^T A$ 는 항상 대칭행렬이다.

(3) $A_{n\times n}$ 에 대하여, $\dfrac{A^T + A}{2}$ 는 항상 대칭행렬이다.

(4) $A_{n\times n}$ 가 대칭행열이면, 최대로 $n(n+1)/2$ 개의 서로 다른 원소를 포함할 수 있다.

5.2 반대칭행렬(Skew-Symmetric Matrix)

Definition

행렬 $A_{n\times n}$ 에 대하여

$$A^T = -A \tag{5.3}$$

즉,

$$a_{ij} = -a_{ji}, \quad (\forall\, i, j = 1, \ldots, n) \tag{5.4}$$

을 만족하면, A 는 반대칭행렬(skew-symmetric matrix)이다.

Theorem (정리 5.2)

반대칭행렬에 대해 다음이 성립한다.

(1) *A 가 반대칭행렬이면 $a_{11} = a_{22} = \cdots = a_{nn} = 0$ 이다.*

(2) *A 가 반대칭행렬이면 $\mathrm{tr}(A) = 0$ 이다.*

(3) *정방행렬 $A_{n\times n}$ 에 대하여, $\dfrac{A^T - A}{2}$ 는 항상 반대칭행렬이다.*

(4) *$A_{n\times n}$ 가 반대칭행열이면, 임의의 벡터 $\mathbf{x}_{n\times 1}$ 에 대하여, $\mathbf{x}^T A\mathbf{x} = 0$ 이다.*

Proof.

(1) : $A^T = -A$ 이므로 $A^T + A = 0$ 이다. 따라서

$$\begin{aligned} a_{ij} + a_{ji} &= 0, \quad \text{if } i \neq j \qquad (5.5) \\ 2a_{ii} &= 0, \quad \text{if } i = j \qquad (5.6) \end{aligned}$$

이다. 따라서 (1)이 성립한다.

(2) : 위에 의해 자동으로 성립

(3), (4) : 연습문제 □

Proposition (5.3)

임의의 정방행렬 A 는 대칭행렬과 반대칭행렬의 합으로 표현할 수 있다.

Proof.

임의의 정방행렬 A 를 아래와 같이 표현 가능하다.

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T). \tag{5.7}$$

□

5.3 항등행렬(Identity Matrix)

Definition

정방행렬 I_n 이 아래와 같이 정의될 때,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{5.8}$$

이를 항등행렬(identity matrix)이라고 한다.

Remark 항등행렬 I_n 에 대해 다음이 성립한다.

- I_n 은 대칭행렬이다.
- 임의의 행렬 $A_{m\times n}$ 에 대하여 $AI_n = A, I_mA = A$ 가 성립한다.

5.4 역행렬(Inverse Matrix)

Definition

행렬 $A_{n \times n}$ 에 대하여

$$AA^{-1} = A^{-1}A = I_n \tag{5.9}$$

을 만족하는 A^{-1} 이 존재하면, A^{-1} 은 A 의 역행렬이라 부른다.

Example

다음 행렬 A 에 대하여 A^{-1} 가 A 의 역행렬이 되는지 확인하라.

- $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$

  (풀이)

$$AA^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (1)(1) + (1)(0) & (1)(-1) + (1)(1) \\ (0)(1) + (1)(0) & (-1)(0) + (1)(1) \end{pmatrix} = I_2$$

$$A^{-1}A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (1)(1) + (-1)(0) & (1)(1) + (-1)(1) \\ (1)(0) + (1)(0) & (0)(1) + (1)(1) \end{pmatrix} = I_2$$

- $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

Example

A^{-1} 가 A 의 역행렬이 되도록 상수 a, b 를 정하라.

$$A = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix}$$

5.5 삼각행렬(Triangular Matrix)

Definition

정방행렬 $A_{n\times n}$ 에 대하여

$$a_{ij} = 0 \quad \text{for } i > j \tag{5.10}$$

이면 이 행렬을 상삼각행렬(upper triangular matrix)이라 한다.

Definition

정방행렬 $A_{n\times n}$ 에 대하여

$$a_{ij} = 0 \quad \text{for } i < j \tag{5.11}$$

이면 이 행렬을 하삼각행렬(lower triangular matrix)이라 한다.

Example

다음은 상삼각행렬과 하삼각행렬의 예이다.

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

5.6 합벡터(Summing Vector)

행렬의 모든 원소들을 더하거나, 행렬의 행방향 혹은 열방향의 원소들을 더하는 역할을 하는 벡터를 **합벡터(summing vector)**라고 하며 아래와 같이 정의한다.

$$\mathbf{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n\times 1}.$$

(1) 행렬의 행방향 원소의 합을 구할 때:

$$A\mathbf{1}_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{1\cdot} \\ a_{2\cdot} \\ \vdots \\ a_{m\cdot} \end{pmatrix}. \tag{5.12}$$

여기서 $a_{i\cdot} = \sum_{j=1}^{n} a_{ij}$ 를 의미한다.

(2) 행렬의 열방향 원소의 합을 구할 때:

$$\mathbf{1}_m^T A = (1, 1, \cdots, 1) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{\cdot 1}, a_{\cdot 2}, \dots, a_{\cdot n}).$$

여기서 $a_{\cdot j} = \sum_{i=1}^{m} a_{ij}$ 를 의미한다.

(3) 행렬의 모든 원소의 합을 구할 때:

$$\mathbf{1}_m^T A \mathbf{1}_n = (1, 1, \cdots, 1) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = (a_{\cdot 1}, a_{\cdot 2}, \dots, a_{\cdot n}) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = a_{\cdot\cdot}$$

여기서 $a_{\cdot\cdot} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}$ 를 의미한다.

Example

아래 항들을 합벡터를 사용하여 표현하라.

(1) n

(2) $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

(3) $J_n = \{1\}_{n\times n}$

(풀이)

(1) $n = \mathbf{1}_n^T \mathbf{1}_n$

(2) $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \mathbf{y}^T \mathbf{1}_n = \mathbf{1}_n^T \mathbf{y}$

(3) $J_n = \{1\}_{n\times n} = \mathbf{1}_n \mathbf{1}_n^T = \mathbf{1}_n \otimes \mathbf{1}_n^T$

5.7 기본벡터(Elementary Vector)

Definition

원소

$$e_j = \begin{cases} 0 & \text{if } j \neq i \\ 1 & \text{if } j = i \end{cases} \qquad (j = 1, \ldots, n)$$

로 구성되는 벡터 $\mathbf{e}_i \in \mathbb{R}^n$ 를 기본벡터(elementary vector)라고 부르며 다음과 같은 형태를 지닌다.

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{번째 원소} \tag{5.13}$$

Example

아래는 $\mathbb{R}^3$ 에서 정의되는 모든 기본벡터를 나타낸다.

$$\mathbf{e}_1 = (1, 0, 0)^T, \quad \mathbf{e}_2 = (0, 1, 0)^T, \quad \mathbf{e}_3 = (0, 0, 1)^T$$

Remark 기본벡터는 다음과 같이 활용된다.

(1) 행렬의 i 번째 열을 선택하는 경우:

$$A\mathbf{e}_i = \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mi} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = \tilde{\mathbf{a}}_i$$

(2) 행렬의 i 번째 행을 선택하는 경우:

$$\mathbf{e}_i^T A = (0, \cdots, 1, \cdots, 0) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{i1}, a_{i2}, \cdots, a_{in}) = \mathbf{a}_i^T$$

(3) 행렬의 (i,j) 번째 원소를 선택하는 경우:

$$\mathbf{e}_i^T A \mathbf{e}_j = (0, \cdots, 1, \cdots, 0) \begin{pmatrix} a_{11} & \cdots & a_{ij} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = (a_{i1}, \cdots, a_{ij}, \cdots, a_{in}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ij}$$

Example

다음 행렬에서 기본벡터를 이용하여 2행 3열 원소를 선택하는 식을 표현하라

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 \\ 2 & 1 & 1 & 1 \\ 0 & 3 & -1 & 4 \\ 7 & -2 & -3 & 8 \end{pmatrix}$$

(풀이)

$$a_{23} = \mathbf{e}_2^T A \mathbf{e}_3 = 1$$

5.8 멱등행렬(Idempotent Matrix)

Definition

정방행렬 $A_{n\times n}$ 에 대하여,

$$AA = A \tag{5.14}$$

가 성립하면 A 를 멱등행렬(idempotent matrix)이라 부른다.

Proposition (5.4)

I_n 은 멱등행렬이다.

Proof.

항등행렬의 정의에 의해 $I_nI_n = I_n$. □

Proposition (5.5)

$\frac{1}{n}J_n$ 은 멱등행렬이다.

Proof.

$\frac{1}{n}J_n = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ 이고 $\mathbf{1}_n^T\mathbf{1}_n = n$ 이므로,

$$\left(\frac{1}{n}J_n\right)\left(\frac{1}{n}J_n\right) = \left(\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)\left(\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right) = \frac{1}{n^2}\mathbf{1}_n\underbrace{\mathbf{1}_n^T\mathbf{1}_n}_{=n}\mathbf{1}_n^T = \frac{n}{n^2}\mathbf{1}_n\mathbf{1}_n^T = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T = \frac{1}{n}J_n$$

□

Proposition (5.6)

A 가 멱등행렬이면 임의의 자연수 k 에 대해, $A^k = A$ 이다.

Proof.

$A^2 = AA = A, A^3 = A^2A = AA = A,$ 등 순차적으로 적용하면 됨. □

Proposition (5.7)

A 가 멱등행렬이면 $I_n - A$ 도 멱등행렬이다.

Proof.

$(I_n - A)(I_n - A) = I_n - A - A + A^2 = I_n - A - A + A = I_n - A.$ □

Proposition (5.8)

임의의 행렬 $X_{m\times n}$ 에 대해, $X(X^TX)^{-1}X^T$ 는 멱등행렬이다.

Proof.

$$\left(X(X^TX)^{-1}X^T\right)\left(X(X^TX)^{-1}X^T\right) = X\underbrace{(X^TX)^{-1}X^TX}_{=I_n}(X^TX)^{-1}X^T = X(X^TX)^{-1}X^T$$

□

Example

다음 행렬이 멱등행렬이 되기 위한 상수 c 값을 정하라.

$$\begin{pmatrix} 2 & c & -4 \\ -1 & 3 & 4 \\ 1 & c & -3 \end{pmatrix} \tag{5.15}$$

5.9 행렬의 이차형식(Quadratic Form)

다항식은 변수 x 와 이의 승수들의 합으로 이루어진 함수를 말하여 아래의 모습을 한다.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p. \tag{5.16}$$

Definition

n 개의 변수 $x_1, x_2, \ldots, x_n$ 으로 이루어진 2차함수를

$$f(x_1, x_2, \ldots, x_n) = \beta_{11} x_1^2 + \ldots + \beta_{nn} x_n^2 + \beta_{12} x_1 x_2 + \ldots + \beta_{n_1, n} x_{n-1} x_n$$

라고 할 때, 이를

$$f(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{x}^T A \mathbf{x} \tag{5.17}$$

로 표현이 가능하며 이는

$$f(x_1, x_2, \ldots, x_n) = \mathbf{x}^T A \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{5.18}$$

이다. 이를 벡터 $\mathbf{x}$ 의 이차형식(quadratic form)이라 한다.

Remark 위의 이차형식에서

$$\beta_{ii} = a_{ii}, \qquad \beta_{ij} = a_{ij} + a_{ji}$$

임을 알 수 있다. 또한 만일 A 가 대칭행렬이라 가정하면

$$\beta_{ii} = a_{ii}, \qquad a_{ij} = a_{ji} = \frac{\beta_{ij}}{2}$$

이므로

$$A = \begin{pmatrix} \beta_{11} & \frac{\beta_{12}}{2} & \cdots & \frac{\beta_{1n}}{2} \\ \frac{\beta_{21}}{2} & \beta_{22} & \cdots & \frac{\beta_{2n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_{n1}}{2} & \frac{\beta_{n2}}{2} & \cdots & \beta_{nn} \end{pmatrix} \tag{5.19}$$

으로 둘 수 있다.

Example

이차다항식

$$f(x_1, x_2, x_3) = 9x_1^2 + 7x_2^2 + 3x_3^2 - 2x_1x_2 + 4x_1x_3 - 6x_2x_3 \tag{5.20}$$

을 벡터 $\mathbf{x} = (x_1, x_2, x_3)^T$ 에 대한 이차형식 $\mathbf{x}^T A \mathbf{x}$ 으로 표현하기 위해 대칭행렬 A 를 구하라.

Theorem

임의의 반대칭행렬(skew-symmetric matrix) A *에 대해 임의의 벡터* $\mathbf{x}$ *에 대한 이차형식은* $\mathbf{x}^T A\mathbf{x} = 0$ *이다.*

Proof.

$\mathbf{x}^T A\mathbf{x}$ 는 스칼라이고 $A^T = -A$ 이므로

$$\mathbf{x}^T A\mathbf{x} = (\mathbf{x}^T A\mathbf{x})^T = \mathbf{x}^T A^T(\mathbf{x}^T)^T = \mathbf{x}^T A^T\mathbf{x} = -\mathbf{x}^T A\mathbf{x}$$

가 성립한다. 따라서

$$2\mathbf{x}^T A\mathbf{x} = 0$$

이므로 $\mathbf{x}^T A\mathbf{x} = 0$ 이다. □

5.10 양정치행렬(Positive Definite Matrix)

Definition

정방행렬 $A_{n\times n}$ 이 $\mathbf{0}_n$ 이 아닌 임의의 벡터 $\mathbf{x} \in \mathbb{R}^n$ 에 대해

$$\mathbf{x}^T A\mathbf{x} > 0, \quad \forall\, \mathbf{x} \neq \mathbf{0}_n$$

을 만족하면, A 를 양정치행렬(positive definite matrix)이라고 한다.

Remark 만일 정방행렬 $A_{n\times n}$ 이 $\mathbf{0}_n$ 이 아닌 임의의 벡터 $\mathbf{x} \in \mathbb{R}^n$ 에 대해

$$\mathbf{x}^T A\mathbf{x} < 0, \quad \forall\, \mathbf{x} \neq \mathbf{0}_n$$

을 만족하면, A 를 음정치행렬(negative definite matrix)이라고 한다.

Remark 만일 정방행렬 $A_{n\times n}$ 이 $\mathbf{0}_n$ 이 아닌 임의의 벡터 $\mathbf{x} \in \mathbb{R}^n$ 에 대해

$$\mathbf{x}^T A\mathbf{x} \geq 0, \quad \forall\, \mathbf{x} \neq \mathbf{0}_n$$

을 만족하면, A 를 반양정치행렬(positive semi-definite matrix)이라고 하고

$$\mathbf{x}^T A\mathbf{x} \leq 0, \quad \forall\, \mathbf{x} \neq \mathbf{0}_n$$

이면, 반음정치행렬(negative semi-definite matrix)이라고 한다.

Example

아래 행렬이 양정치행렬임을 보여라.

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

(풀이) 영이 아닌 임의의 벡터 $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$에 대하여 $\mathbf{x}^T A\mathbf{x} = x_1^2 + x_2^2 > 0$ 이 성립하므로 양정치이다.

Remark 아래를 만족하는 $\mathbf{0}_n$ 이 아닌 벡터 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 이 존재하면 A 를 부정행렬(indefinite matrix)이라고 한다.

$$\mathbf{x}^T A\mathbf{x} < 0 < \mathbf{y}^T A\mathbf{y}.$$

Example

다음 행렬 중 양정치, 반양정치, 부정행렬 등을 구분하라.

$$A = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

5.11 이차형식의 그래프

이차형식으로 표현되는 이차함수 $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ 는 행렬 A 의 특성에 따라 형태가 결정된다.

- A 가 양정치인 경우, $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} > 0$
- A 가 반양정치인 경우, $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \geq 0$
- A 가 음정치인 경우, $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} < 0$
- A 가 반음정치인 경우, $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \leq 0$

$\mathbf{x} \in \mathbb{R}^2$ 이고 $A_{2\times 2}$ 인 경우 이차함수의 그래프는 다음과 같다.

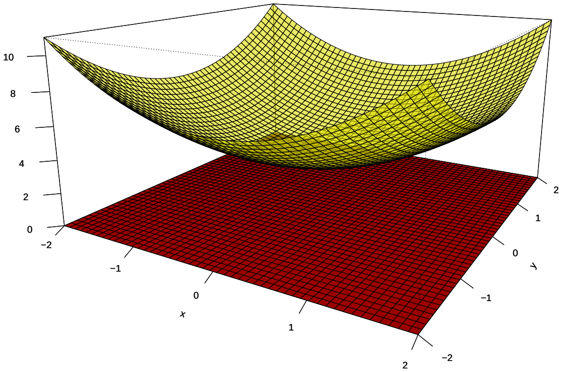

Figure 5.1: 양정치행렬에 대한 이차형식의 예

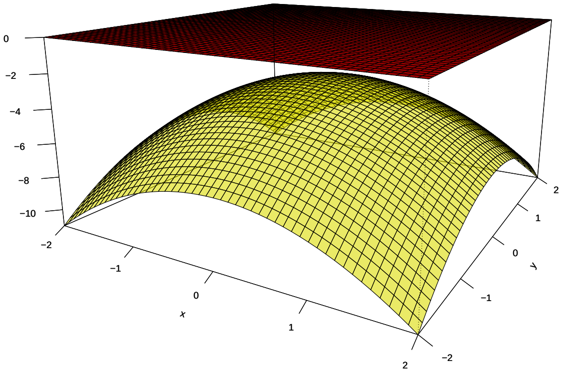

Figure 5.2: 음정치행렬에 대한 이차형식의 예

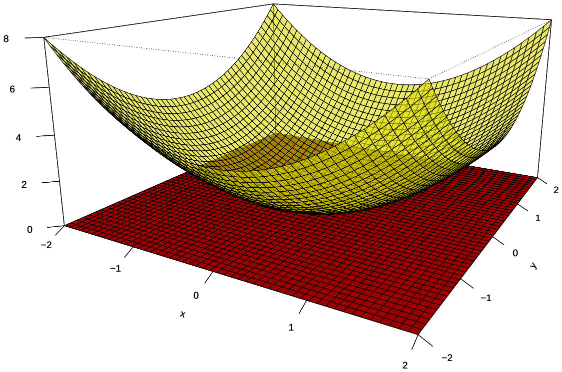

Figure 5.3: 반양정치행렬에 대한 이차형식의 예

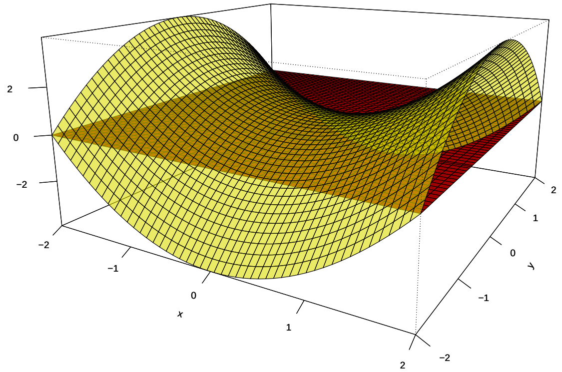

Figure 5.4: 부정행렬에 대한 이차형식의 예

5.12 양정치행렬을 구분하는 법

Definition

행렬 $A_{n\times n} = \{a_{ij}\}_{n\times n}$ 에 대해

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \cdots \quad , \quad \Delta_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

이라 할 때, Δ_k $(k = 1, 2, \ldots, n)$ 를 k 번째 leading principal minor 라 정의한다.

Theorem (정리 5.10)

행렬 $A_{n\times n}$ 에 대해, 모든 leading principal minor 의 행렬식이

$$\det(\Delta_k) > 0, \quad \forall\, k = 1, 2, \ldots, n \tag{5.21}$$

이면 행렬 A 는 양정치행렬이다.

Example

아래 행렬이 양정치행렬임을 위의 정리를 이용해 증명하라.

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Theorem (정리 5.11)

행렬 $A_{n\times n}$ 에 대해, 모든 고유값(eigenvalue)이

$$\lambda_k > 0, \quad \forall\, k = 1, 2, \ldots, n \tag{5.22}$$

이면 행렬 A 는 양정치행렬이다.

Remark 양정치행렬임을 확인하는 방법을 정리하면 다음과 같다.

1. $\mathbf{0}$이 아닌 임의의 벡터 $\mathbf{x}$ 에 대해, $\mathbf{x}^T A\mathbf{x} > 0$ 이 항상 성립
2. Leading principal minor 의 행렬식이 모두 양수
3. 고유값이 모두 양수

5.13 공분산행렬과 촐레스키분해

Definition

벡터 $\mathbf{y} \in \mathbb{R}^n$ 의 모든 원소가 확률변수(random variable)로 이루어졌을 경우, 이 벡터를 확률벡터(random vector)라고 한다.

Example

만일 확률변수 y_1 이 부모의 소득을 나타내고, 확률변수 y_2 가 자식의 소득을 나타낸다면, 벡터

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2 \tag{5.23}$$

는 부모와 자식의 소득분포를 나타내는 확률벡터이다.

Definition

확률변수 y 가 어떤 분포를 가질 경우, 그 분포상의 위치를 대표하는 값을 기대값(expected value, 혹은 mean)이라고 하고 $E(y)$ 로 표시하며, 그 분포의 퍼짐의 정도를 나타내는 것을 분산(variance)이라고 하며 $\mathrm{Var}(y)$ 라고 표시한다.

Definition

확률벡터 $\mathbf{y} \in \mathbb{R}^n$ 의 기대값은 다음과 같이 정의한다.

$$\boldsymbol{\mu} = E(\mathbf{y}) = \begin{pmatrix} E(y_1) \\ E(y_2) \\ \vdots \\ E(y_n) \end{pmatrix}. \tag{5.24}$$

Definition

확률벡터 $\mathbf{y} \in \mathbb{R}^n$ 의 분산(공분산행렬)은 다음과 같이 정의한다.

$$V = \mathrm{Var}(\mathbf{y}) = \begin{pmatrix} \mathrm{Var}(y_1) & \mathrm{Cov}(y_1, y_2) & \cdots & \mathrm{Cov}(y_1, y_n) \\ \mathrm{Cov}(y_2, y_1) & \mathrm{Var}(y_2) & \cdots & \mathrm{Cov}(y_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(y_n, y_1) & \mathrm{Cov}(y_n, y_2) & \cdots & \mathrm{Var}(y_n) \end{pmatrix}. \tag{5.25}$$

Remark 공분산 $\mathrm{Cov}(x, y)$ 는 두 확률변수 x 와 y 간의 상관성을 나타내며 양의 상관관계에 있는 경우 $\mathrm{Cov}(x, y) > 0$ 이고 음의 상관관계에 있는 경우 $\mathrm{Cov}(x, y) < 0$ 이 된다.

Example

확률변수 x_1 은 평균이 1이고 분산이 2인 정규분포를 따르고, 확률변수 x_2 는 평균이 2이고 분산이 2인 정규분포를 따른다고 하자. 두 확률변수 간의 상관계수가 0.5라고 한다면, 확률벡터 $\mathbf{x} = (x_1, x_2)^T$ 의 평균벡터 $\boldsymbol{\mu}$ 와 공분산행렬 V 를 구하라.

(풀이) 상관계수의 정의에 의해

$$\rho = \frac{\mathrm{Cov}(x_1, x_2)}{\sqrt{\mathrm{Var}(x_1)\mathrm{Var}(x_2)}} = \frac{\mathrm{Cov}(x_1, x_2)}{\sqrt{2}\sqrt{2}} = 0.5 \tag{5.26}$$

이므로 $\mathrm{Cov}(x_1, x_2) = 1 = \mathrm{Cov}(x_2, x_1)$ 이다. 따라서

$$\boldsymbol{\mu} = \begin{pmatrix} E(x_1) \\ E(x_2) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{5.27}$$

이고

$$V = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{5.28}$$

이다.

Definition

상수벡터 $\mathbf{c}$ 와 확률벡터 $\mathbf{y}$ 로 이루어진 확률변수 $\mathbf{c}^T\mathbf{y}$ 의 평균은

$$E(\mathbf{c}^T\mathbf{y}) = \mathbf{c}^T E(\mathbf{y}) \tag{5.29}$$

이고, 분산은

$$\mathrm{Var}(\mathbf{c}^T\mathbf{y}) = \mathbf{c}^T \mathrm{Var}(\mathbf{y})\mathbf{c} \tag{5.30}$$

이다.

Example

확률변수 $y_1 \sim N(1,2)$, $y_2 \sim N(2,4)$ 이고 $\mathrm{Cov}(y_1, y_2) = -1$ 일 때, $y_1 - y_2$ 의 평균과 분산을 구하라.

(풀이) 확률벡터 $\mathbf{y} = (y_1, y_2)^T$ 의 평균과 공분산행렬을 구하면

$$E(\mathbf{y}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathrm{Var}(\mathbf{y}) = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix} \tag{5.31}$$

이다. $y_1 - y_2$ 는 $\mathbf{c} = (1, -1)^T$ 를 이용하여 $y_1 - y_2 = \mathbf{c}^T\mathbf{y}$ 로 표현되므로, 평균은

$$E(\mathbf{c}^T\mathbf{y}) = (1, -1)\begin{pmatrix} 1 \\ 2 \end{pmatrix} = (1)(1) + (-1)(2) = -1 \tag{5.32}$$

이고 분산은

$$\mathrm{Var}(\mathbf{c}^T\mathbf{y}) = (1, -1)\begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = (3, -5)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 8 \tag{5.33}$$

이다.

Example

이전 예제에서 $2y_1 - y_2$ 의 평균과 분산을 구하라.

Theorem (정리 5.12)

확률벡터 $\mathbf{y} \in \mathbb{R}^n$ *의 공분산행렬은 항상 반양정치이다.*

Proof.

임의의 상수벡터 $\mathbf{x} \in \mathbb{R}^n$ 에 대해 $\mathrm{Var}(\mathbf{x}^T\mathbf{y}) = \mathbf{x}^T V\mathbf{x}$ 이다. 모든 확률변수의 분산은 0보다 작을 수 없으므로 임의의 $\mathbf{x}$ 에 대하여

$$\mathrm{Var}(\mathbf{x}^T\mathbf{y}) = \mathbf{x}^T V\mathbf{x} \geq 0$$

이 성립한다. 따라서 V 는 반양정치이다. □

Theorem (정리 5.13 Cholesky Factorization)

행렬 $A_{n\times n}$ *가 양정치행렬이고 대칭행렬이면*

$$A = \Gamma\Gamma^T \tag{5.34}$$

가 성립하는 하삼각행렬 $\Gamma_{n\times n}$ *가 존재한다. 만일 이러한 분해가 여러개 존재하면 대각원소가 양수인 것을 선택한다.*

Example

아래 행렬이 대칭-양정치행렬임을 보이고, 촐레스키 분해를 행하라.

$$A = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$$

(풀이) A 는 대칭행렬이며 양정치 임은 쉽게 보일 수 있다. 하삼각행렬

$$\Gamma = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$$

이 존재하여 $A = \Gamma\Gamma^T$ 를 만족한다고 가정하자. 그러면

$$\Gamma\Gamma^T = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & c^2 \end{pmatrix}$$

이므로 $a^2 = 1$, $ab = -1/2$, $c^2 = 1$ 이다. Γ 의 대각원소를 양수만 취한다면 $a = 1$, $b = -1/2$, $c = 1$ 이므로

$$\Gamma = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \end{pmatrix} \quad \longrightarrow \quad A = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/2 \\ 0 & 1 \end{pmatrix}.$$

Example

선형연립방정식 $A\mathbf{x} = \mathbf{b}$ 가 아래와 같을 때, x_1, x_2, x_3 의 해를 구하는데 있어, A 가 하삼각행렬이라는 사실이 계산에 어떻게 도움이 되는지 생각하라.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 9 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

(풀이) 위의 연립방정식은

$$x_1 = 1, \quad 2x_1 + 9x_2 = 2, \quad x_1 + 2x_3 = 3$$

이므로 $x_1 = 1, x_2 = 0, x_3 = 1$ 이 손쉽게 얻어진다.

Remark 연립방정식 $A\mathbf{x} = \mathbf{b}$ 의 A 가 양정치라고 가정하자. 촐레스키 분해에 의해 $A = \Gamma\Gamma^T$ 를 만족하는 하삼각행렬 Γ 이 존재하므로, $\mathbf{y} = \Gamma^T\mathbf{x}$ 로 두면 위의 연립방정식은 $\Gamma\mathbf{y} = \mathbf{b}$ 가 되며 $\mathbf{y}$ 는 손쉽게 얻어진다. 또한, Γ^T 는 상삼각행렬이므로 $\Gamma^T\mathbf{x} = \mathbf{y}$ 로부터 $\mathbf{x}$ 를 손쉽게 얻을 수 있다.