이차 형식과 정부호 행렬

Quadratic Forms and Definite Matrices

Keon M. Lee

- 이차형식
- 이차형식의 기하학적 형태
- 대칭행렬의 분류 : 정부호 성질
- 이차형식의 최적화 문제
- 정부호 행렬과 eigenvalue
- 정부호 행렬의 역행렬
- 공분산행렬의 정부호 특성
- Gram 행렬의 정부호 특성
- Cholesky 분해

이차 형식

❖ **이차 형식**(Quadratic form)

- R^n에서 정의된 함수로 $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$형태로 표현될 수 있는 것

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$$

$$\mathbf{x}^T A\mathbf{x} = [\,x_1 \quad x_2\,]\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,]\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2$$

$$A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= [\,x_1 \quad x_2\,]\begin{bmatrix} 3 & \mathbf{-2} \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,]\begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\ &= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2 \end{aligned}$$

A는 대칭행렬로 표현

2차형식 함수의 대칭행렬 표현

❖ **각 항(term)의 계수(coefficient)를 대응하는 원소에 배치**

- 혼합항의 경우, 계수의 ½값을 분할하여 배치

$$Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$$

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = [\,x_1 \quad x_2 \quad x_3\,] \begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

2차 형식의 변수변환(change of variable)

❖ **가역 행렬 P를 사용한 변수 변환** $\mathbf{x} \Leftrightarrow \mathbf{y}$

$$\mathbf{x} = P\mathbf{y} \qquad \mathbf{y} = P^{-1}\mathbf{x}$$

$$\mathbf{x}^\top A\mathbf{x} = (P\mathbf{y})^\top A(P\mathbf{y}) = \mathbf{y}^\top P^\top AP\mathbf{y} = \mathbf{y}^\top(P^\top AP)\mathbf{y}$$
$$= \mathbf{y}^\top D\mathbf{y}$$

A가 **대칭행렬**이기 때문에, **직교 대각화 가능** P 존재

변수 변환을 통해 혼합항이 없는 2차형식으로 표현

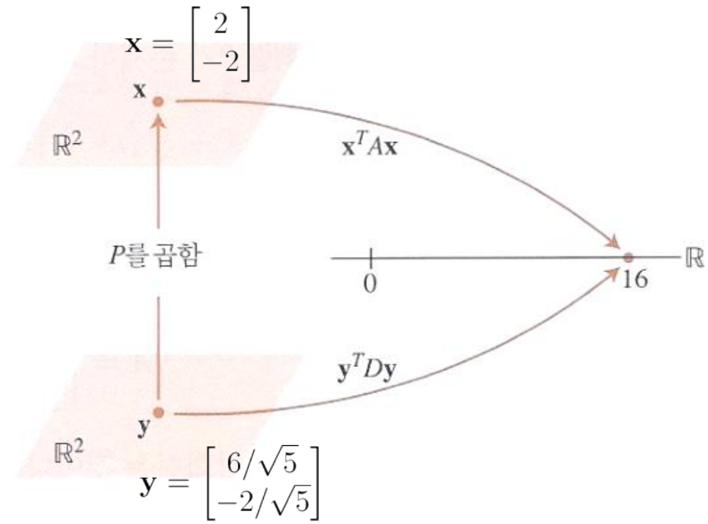
❖ A가 **대칭행렬**이면, 2차형식 $\mathbf{x}^T A\mathbf{x}$ 를 **혼합항이 없는 2차형식** $\mathbf{y}^T A\mathbf{y}$ 로 변환시키는 변수의 **직교변환** $\mathbf{x} = P\mathbf{y}$가 존재

2차 형식의 변수변환(change of variable)

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

$$\lambda = 3: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$



$$D = P^T A P$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\mathbf{x} = P\mathbf{y}$$

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) \\ &= \mathbf{y}^T P^T A P\mathbf{y} = \mathbf{y}^T D\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned}$$

$$\mathbf{y} = P^{-1}\mathbf{x} = P^T\mathbf{x}$$

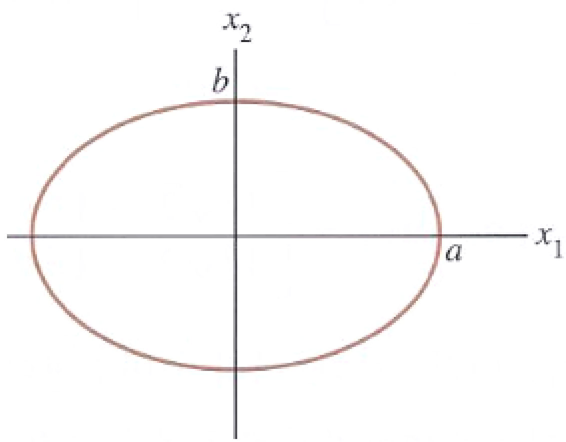
$$\mathbf{y} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5) \\ &= 80/5 = 16 \end{aligned}$$

2차형식의 기하학적 형태

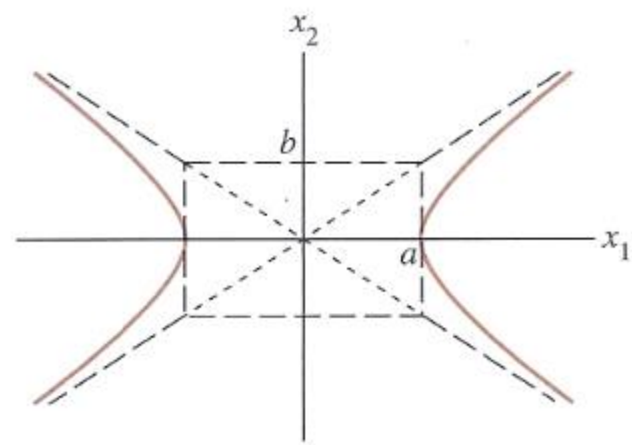
❖ 혼합항이 없는 2차형식은 **주축(principal axes)**에 대칭

$$Q(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x}$$



$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \ a > b > 0$$

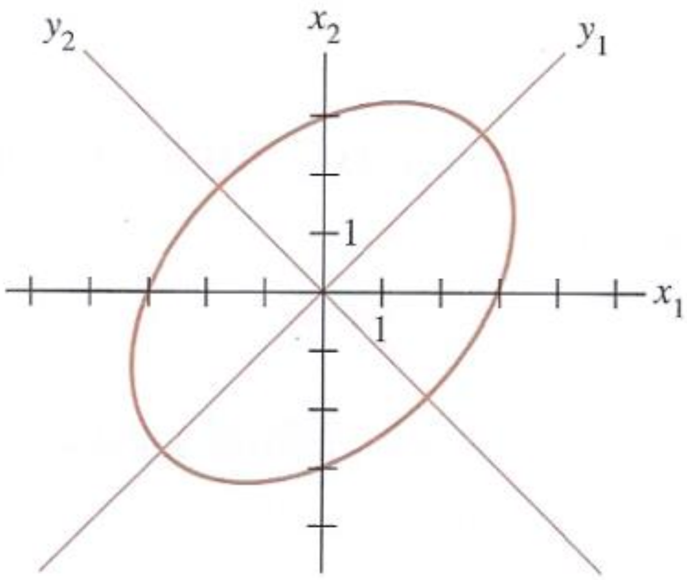
$$\mathbf{x}^\top A\mathbf{x} = c$$



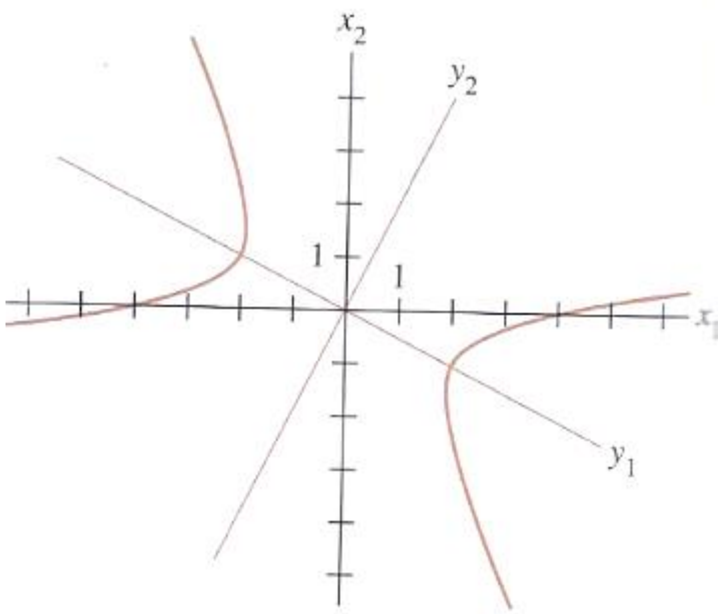
$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \ a > b > 0$$

2차형식의 기하학적 형태

❖ 혼합항이 있는 2차형식



(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$

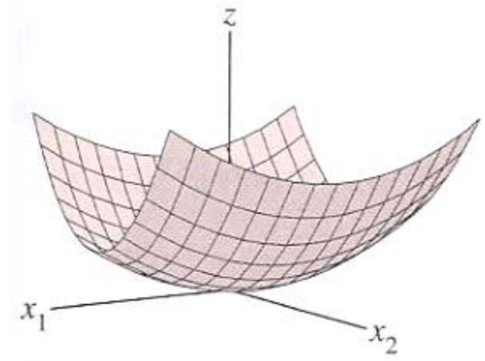


(b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

대칭행렬의 분류

❖ **Positive definite (양의 정부호)**

$$Q(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} > 0, \quad \forall \mathbf{x} \neq \mathbf{0}$$

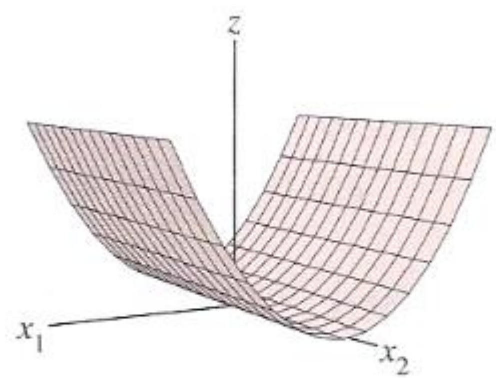


$z = 3x_1^2 + 7x_2^2$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$$

❖ **Positive semi-definte (양의 반정부호)**

$$Q(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} \geq 0, \quad \forall \mathbf{x} \neq \mathbf{0}$$



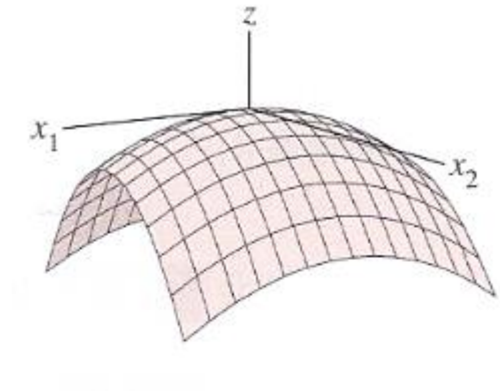
$z = 3x_1^2$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$$

대칭행렬의 분류

❖ **Negative definite (음의 정부호)**

$$Q(\mathbf{x}) = \mathbf{x}^{\top} A \mathbf{x} < 0, \quad \forall \mathbf{x} \neq \mathbf{0}$$



$$A = \begin{bmatrix} -3 & 0 \\ 0 & -7 \end{bmatrix}$$

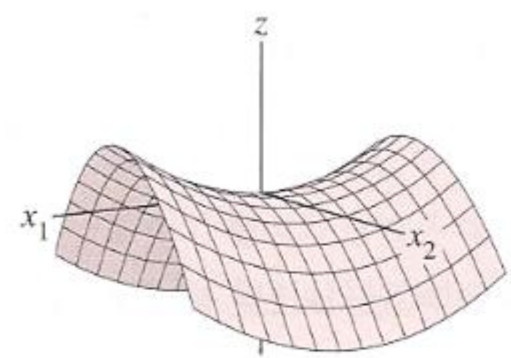
$z = -3x_1^2 - 7x_2^2$

❖ **Negative semi-definte (양의 반정부호)**

$$Q(\mathbf{x}) = \mathbf{x}^{\top} A \mathbf{x} \leq 0, \quad \forall \mathbf{x} \neq \mathbf{0}$$

❖ **Indefinite (부정부호)**

$$Q(\mathbf{x}) = \mathbf{x}^{\top} A \mathbf{x}$$



$$A = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

$z = 3x_1^2 - 7x_2^2$

정부호 행렬과 eigenvalue

❖ **대칭행렬** A에 대해

- 모든 eigenvalue $\lambda_i > 0$이면, A는 양의 정부호 (positive definite)
- 모든 eigenvalue $\lambda_i \geq 0$이면, A는 양의 반정부호 (positive semi-definite)
- 모든 eigenvalue $\lambda_i < 0$이면, A는 음의 정부호 (negative definite)
- 모든 eigenvalue $\lambda_i \leq 0$이면, A는 음의 반정부호 (negative semi-definite)
- 양의 eigenvalue와 음의 eigenvalue를 모두 포함하면,
  A는 부정부호(indefinite)

- 증명

$$\mathbf{x}^\top A\mathbf{x} = \mathbf{x}^\top S\Lambda S^\top \mathbf{x} = \mathbf{y}^\top \Lambda \mathbf{y} = \sum_i \lambda_i y_i^2$$

이차형식의 최적화 문제

❖ 제약된 최적화 문제

- $||\mathbf{x}|| = \mathbf{x}^T\mathbf{x} = 1$일 때, $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$의 **최대값**과 **최소값**

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

$$\mathbf{x}^\top A\mathbf{x} = \lambda\mathbf{x}^\top\mathbf{x}$$

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

$$\mathbf{x}^\top A\mathbf{x} = \lambda \quad \text{when} \quad ||\mathbf{x}|| = 1$$

- 최대값 : 가장 큰 eigenvalue
- 최소값 : 가장 작은 eigenvalue

정부호 행렬의 역행렬

- **양의 정부호(positive definite) 행렬과 음의 정부호(negative definite) 행렬**
  - Eigenvalue가 0이 아니면서 모두 양수이거나, 모두 음수
  - det(A) = $\Pi_i \lambda_i \neq 0$
  - 따라서 역행렬 존재

- **정부호 행렬의 역행렬에 대한 eigenvalue**
  - 역행렬의 eigenvalue = 원래 행렬에 대한 eigenvalue의 역수
  - 양의 정부호 행렬의 역행렬 → 양의 정부호
  - 음의 정부호 행렬의 역행렬 → 음의 정부호

공분산 행렬의 정부호 성질

❖ **공분산 행렬 *C***

$$C = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top$$

- 대칭행렬

$$C^\top = \frac{1}{n}\sum_{i=1}^{n}\left((\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top\right)^\top = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top = C$$

- 양의 반정부호 (positive semi-definite)

$$\begin{aligned}
\mathbf{y}^\top Q\mathbf{y} &= \mathbf{y}^\top\left(\frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top\right)\mathbf{y} \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}^\top(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top\mathbf{y} \\
&= \frac{1}{n}\sum_{i=1}^{n}((\mathbf{x}_i - \bar{\mathbf{x}})^\top\mathbf{y})^\top((\mathbf{x}_i - \bar{\mathbf{x}})^\top\mathbf{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left((\mathbf{x}_i - \bar{\mathbf{x}})^\top\mathbf{y}\right)^2 \geq 0
\end{aligned}$$

Gram 행렬의 정부호 성질

❖ Gram 행렬

- 데이터를 열벡터로 갖는 행렬 A = [$\boldsymbol{a}_1 \ldots \boldsymbol{a}_m$] 에 대해
  A^TA, AA^T
- 대칭행렬

$$(A^\top A)^\top = A^\top (A^\top)^\top = A^\top A$$
$$(AA^\top)^\top = (A^\top)^\top A^\top = AA^\top$$

- 양의 반정부호 (positive semi-definite) : A가 가역행렬인 경우

$$\mathbf{x}^\top AA^\top \mathbf{x} = (A^\top \mathbf{x})^\top (A^\top \mathbf{x}) = ||A^\top \mathbf{x}||^2 \geq 0$$
$$\mathbf{x}^\top A^\top A\mathbf{x} = (A\mathbf{x})^\top (A\mathbf{x}) = ||A\mathbf{x}||^2 \geq 0$$

$$\text{If } \mathbf{x}^\top A^\top A\mathbf{x} = 0,\ ||A\mathbf{x}||^2 = 0 \text{ and } ||A\mathbf{x}|| = 0.$$

$$A\mathbf{x} = \mathbf{0}$$
$$\mathbf{x} = \mathbf{0} \text{ because } A \text{ is invertible}$$

$$\text{Except } \mathbf{x} = \mathbf{0},\ \mathbf{x}^\top A^\top A\mathbf{x} > 0.$$

Cholesky 분해

❖ **양의 정부호 (positive definite) 행렬** A ∈ R^{nxn}

- LU 분해에 의해

$$A = LU = LDW$$

$$A = A^\top$$

L : 대각성분이 1인 하 삼각 행렬
D : 대각성분이 U의 대각성분인 대각행렬
W : 대각성분이 1인 상 삼각 행렬

$$A = LDW = (LDW)^\top = W^\top D L^\top$$

$$W = L^\top \qquad \text{(diagnal is fixed, } L \text{ and } U \text{ are unit triangular matrix)}$$

$A = LDL^\top$ 양의 정부호이기 때문에 D의 대각성분은 양수

$$= LD^{1/2}D^{1/2}L^\top = (LD^{1/2})(LD^{1/2})^\top$$

$$A = (LD^{1/2})(LD^{1/2})^\top$$

❖ **Cholesky 분해**

- 양의 정부호 행렬은 하 삼각행렬과 그 전치행렬의 곱으로 분해 가능

$$A = MM^\top \qquad B = LD^{1/2} \quad \text{triangular matrix}$$

Cholesky 분해

❖ Cholesky 분해

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}$$

$$\begin{bmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{bmatrix} \begin{bmatrix} 2 & 6 & -8 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{bmatrix}$$

for k from 1 to n

$$l_{kk} = \left(a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2 \right)^{1/2}$$

for j from $k+1$ to n

$$l_{jk} = \left(a_{jk} - \sum_{s=1}^{k-1} l_{js} l_{ks} \right) / l_{kk}$$

Summary

- ❖ **이차 형식**(Quadratic form)은 R^n에서 정의된 함수로 대칭행렬 A를 사용하여 $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$형태로 표현될 수 있는 것이다.
- ❖ 이차형식은 **직교변환**을 통해 혼합항이 없는 이차형식으로 변환할 수 있다.
- ❖ **대칭행렬**은 이차형식의 값에 따라 양의 정부호(positive definite), 양의 반정부호(positive semi-definite), 음의 정부호(negative definite), 음의 반정부(negative semi-definite), 부정부호(indefinite)로 분류할 수 있다.
- ❖ 대칭행렬의 **정부호 성질**은 **eigenvalue**와 관계가 있다.
- ❖ **이차형식의 최적화 문제**의 **해**가 일정 크기로 제한되는 경우에는 eigenvalue를 통해 쉽게 결정될 수 있다.
- ❖ **정부호 행렬**의 **역행렬**도 동일한 정부호 행렬이다.
- ❖ **공분산행렬**과 **Gram 행렬**은 양의 정부호(positive definite) 행렬이다.
- ❖ **Cholesky 분해**는 양의 정부호 행렬을 삼각행렬과 그 전치행렬의 곱으로 나타내는 것이다.