7. Symmetric Matrices and Quadratic Forms

1) Diagonalization of Symmetric Matrices
2) Quadratic Forms
3) Singular Value Decomposition
4) Principal Component Analysis

1. Diagonalization of Symmetric Matrices

1) Symmetric matrix

(1) Definition : Symmetric matrix

A matrix A is symmetric if $A = A^T$

(2) Theorem

If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal

(3) Definition : Orthogonally diagonalizable

An $n \times n$ matrix A is said to be orthogonally diagonalizable if there are and orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that

$$A = PDP^T = PDP^{-1}$$

(4) Theorem

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is symmetric matrix

2) Spectral Decomposition

(1) The Spectral Theorem for Symmetric Matrices

An $n \times n$ symmetric matrix A has the following properties

① A has n real eigenvalues, counting multiplicities.
② The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.
③ The eigenspaces are mutually orthogonal, in the sense that eigenvectors corrsponding to different eigenvalues are orthogonal.
④ A is orthogonally diagonalizable

(2) Spectral Decomposition

Suppose $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ of A and the corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$ are in the diagonal matrix D

$$A = PDP^T = \begin{bmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \cdots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \cdots \\ \boldsymbol{u}_n{}^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \boldsymbol{u}_1 & \cdots & \lambda_n \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \cdots \\ \boldsymbol{u}_n{}^T \end{bmatrix}$$

$$= \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1{}^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n{}^T \qquad \cdots \quad ①$$

: Spectral decomposition of A

Each term in ① is an $n \times n$ matrix of rank 1

Each matrix $\boldsymbol{u}_j \boldsymbol{u}_j{}^T$ is a projection matrix in the sense that for each $\boldsymbol{x}$ in R^n, the vector $(\boldsymbol{u}_j \boldsymbol{u}_j{}^T)\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{x}$ onto the subspace spanned by $\boldsymbol{u}_j$

(Theorem : If $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ is an orthonormal basis for a subspace W of R^n, then

$proj_W \boldsymbol{y} = UU^T \boldsymbol{y}$, where $U = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_p \end{bmatrix}$)

ex) $A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$

2. Quadratic Form

1) Quadratic Form

(1) Definition : Quadratic Form

A quadratic form on R^{n} is a function Q defined on R^{n} whose value at a vector $\boldsymbol{x}$ in R^{n} can be computed by an expression of the form $Q(\boldsymbol{x})=\boldsymbol{x}^{T}A\boldsymbol{x}$, where A is an $n\times n$ symmetric matrix.

The matrix A is called the matrix of the quadratic form

ex) $Q(\boldsymbol{x})=\boldsymbol{x}^{T}I\boldsymbol{x}=\ \|\boldsymbol{x}\|^{2}$

ex) $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$, $A=\begin{bmatrix}4&0\\0&3\end{bmatrix}$, $B=\begin{bmatrix}3&-2\\-2&7\end{bmatrix}$

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

(2) Change of Variable in a Quadratic Form

If $\boldsymbol{x}$ represents a variable vector in R^n, then a change of variable is an equation of the form

$\boldsymbol{x} = P\boldsymbol{y}$, or $\boldsymbol{y} = P^{-1}\boldsymbol{x}$ $\cdots$ ①

where P is an invertible matrix and $\boldsymbol{y}$ is a new variable vector in R^n
$\boldsymbol{y}$: the coordinate vector of $\boldsymbol{x}$ relative to the basis of R^n determined by the columns of P
If the change of variable ① is made in a quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$, then,

$$\boldsymbol{x}^T A\boldsymbol{x} = (P\boldsymbol{y})^T A(P\boldsymbol{y}) = \boldsymbol{y}^T P^T AP\boldsymbol{y} = \boldsymbol{y}^T \left(P^T AP\right)\boldsymbol{y} \ \cdots ②$$

and the new matrix of the quadratic form is $P^T AP$
Since A is symmetric, there is an orthogonal matrix P such that $P^T AP$ is a diagonal matrix D, and the quadratic form in ② becomes $\boldsymbol{y}^T D\boldsymbol{y}$

(3) Theorem : Principal Axes Theorem

Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\boldsymbol{x} = P\boldsymbol{y}$, that transforms the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ into the quadratic form $\boldsymbol{y}^T D\boldsymbol{y}$ with no cross-product term.

The columns of P are called the principal axes of the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$

The vector $\boldsymbol{y}$ is the coordinate vector of $\boldsymbol{x}$ relative to the orthonormal basis of R^{n} given by these principal axes.

ex) $Q(\boldsymbol{x}) = {x_1}^2 - 8x_1x_2 - 5{x_2}^2$

2) Classifying Quadratic Forms

(1) Classifying Quadratic Forms

When A is an $n \times n$ matrix, the quadratic form $Q(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$ is a real-valued function with domain R^{n}

A quadratic form Q is

① positive definite if $Q(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq 0$
② positive semidefinite if $Q(\boldsymbol{x}) \geqq 0$ for all $\boldsymbol{x} \neq 0$
③ negative definite if $Q(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \neq 0$
④ negative semidefinite if $Q(\boldsymbol{x}) \leqq 0$ for all $\boldsymbol{x} \neq 0$
⑤ indefinite if $Q(\boldsymbol{x})$ assumes both positive and negative values

(2) Theorem : Quadratic Forms and Eigenvalues

Let A be an $n \times n$ symmetric matrix. Then a quadratic form $\boldsymbol{x}^T A \boldsymbol{x}$ is

① positive definite if and only if the eigenvalues of A are all positive
② negative definite if and only if the eigenvalues of A are all negative
③ indefinite if and only if A has both positive and negative eigenvalues

ex) $Q(\boldsymbol{x}) = 3{x_1}^2 + 2{x_2}^2 + {x_3}^2 + 4x_1x_2 + 4x_2x_3$

3. Singular value decomposition

1) Singular Value

(1) Purpose of singular value

(2) Singular value

Let A be $m \times n$ matrix. the singular values of A are the square root of the eigenvalues of A^TA, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order.

$\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \ldots, n$, where λ_i is the ith eigenvalue of A^TA

(3) Theorem

Suppose $\{v_1, \ldots, v_n\}$ is an orthonormal basis of R^{n} consisting of eigenvectors of $A^T A$, arranged so that the corresponding eigenvalues of $A^T A$ satisfy $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_n$, and suppose A has r nonzero singular values. Then $\{Av_1, \ldots, Av_n\}$ is an orthogonal basis for $ColA$, and $rank A = r$

2) Singular Value Decomposition

(1) Singular Value Decomposition

Theorem : Singular Value Decomposition

Let A be $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix Σ such that

$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$, where $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \cdots & \\ 0 & & \lambda_r \end{bmatrix}$. (Diagonal entries of D are the first r singular values of A, $\sigma_1 \geqq \sigma_2 \geqq \cdots \geqq \sigma_r > 0$, and there exists $m \times m$ orthogonal matrix U and $n \times n$ orthogonal matrix V such that

$$A = U \Sigma V^T$$

U and V are not uniquely determined by A, but diagonal entries of Σ are necessarily the singular values of A

The columns of U are called left singular vectors of A, the columns of V are called right singular vectors of A

example

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

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

4. Principal Component Analysis

1) Basic concepts of statistics

(1) 용어정리

data : (x_1, y_1), (x_2, y_2) , $\cdots$, (x_n, y_n)

① mean of data

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

② Sample Variance of data

$$Var(X) = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2, \quad Var(Y) = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{Y}\right)^2$$

③ Covariance of data

$$Cov(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

(2) Covariance matrix

variable : X_1, X_2, $\cdots$, X_p

Covariance matrix of X_1, X_2, $\cdots$, X_p : $\begin{bmatrix} Var(X_1) & Cov(X_1, X_2) \cdots & Cov(X_1, X_p) \\ Cov(X_2, X_1) & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ Cov(X_p, X_1) & \cdots & \cdots & Var(X_p) \end{bmatrix}$

2) Principal Component Analysis

(1) Principal Component Analysis

data : $(x_{11}, x_{12}, \cdots, x_{1p})$, $(x_{21}, x_{22}, \cdots, x_{2p})$, $\cdots$, $(x_{n1}, x_{n2}, \cdots, x_{np})$

$$X = \begin{bmatrix} x_{11}\; x_{12} \cdots x_{1p} \\ x_{21}\; x_{22} \cdots x_{2p} \\ \cdots \;\; \cdots \cdots \cdots \\ x_{n1}\; x_{n2} \cdots x_{np} \end{bmatrix}$$

① scaling

② Covariance matrix

③ Spectral decomposition

④ New axis

⑤ meaning of eigenvalue and eigenvector

⑥ Dimensionality reduction