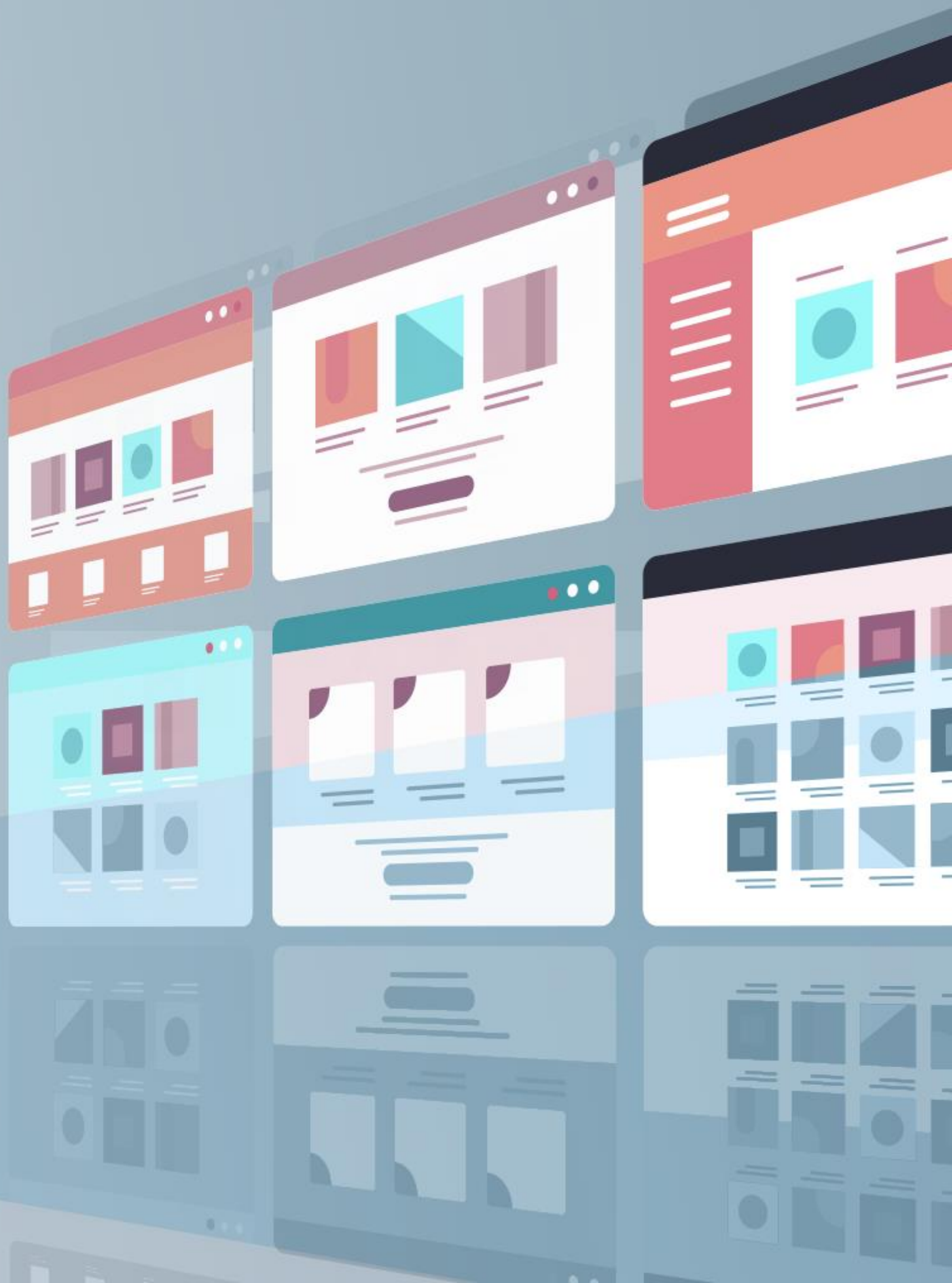
01

데이터분석방법론(1)

Matrix Algebra Basics

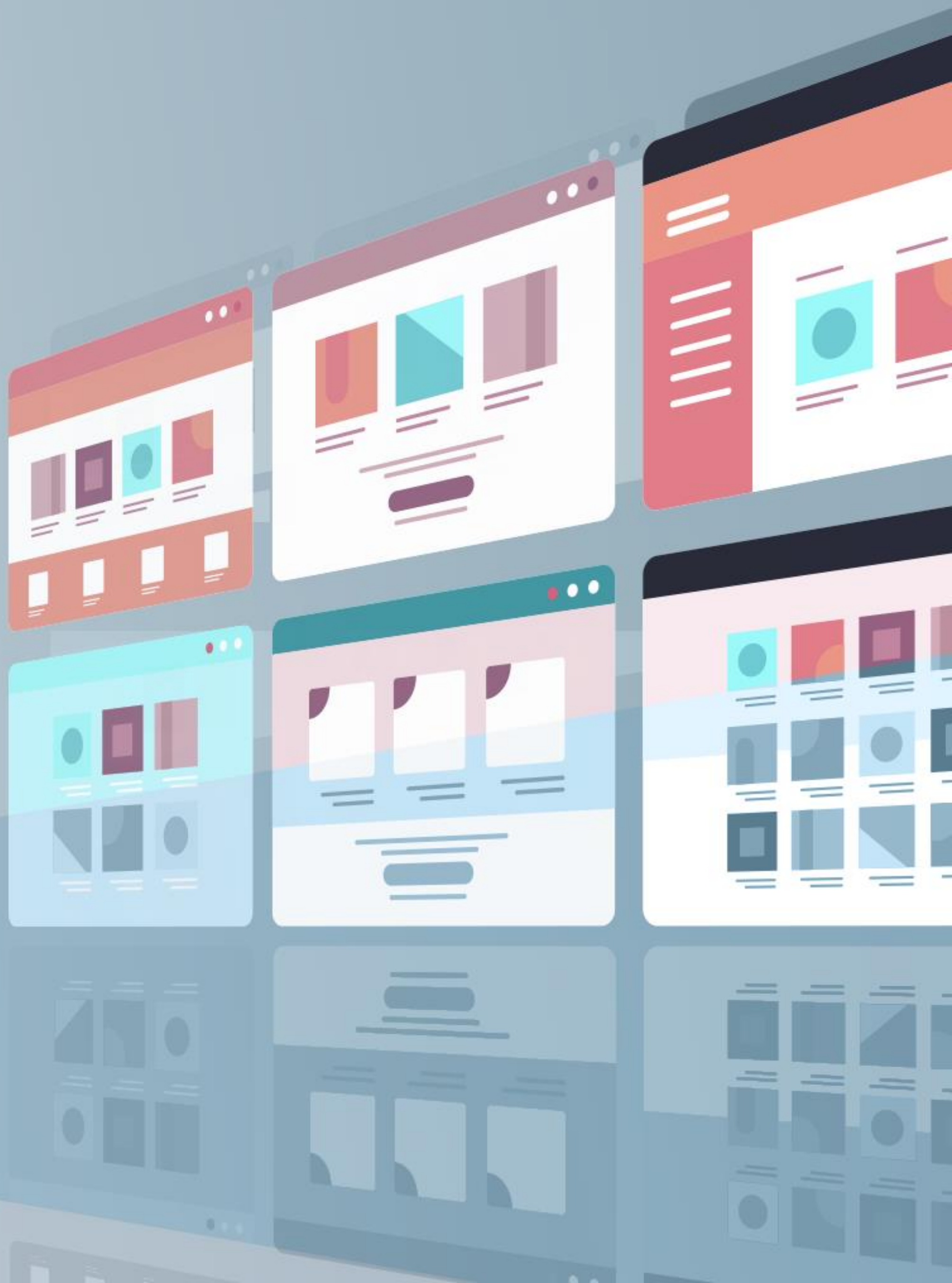
통계·데이터과학과 장영재 교수



학습목차

01

Basic Concepts



1. Scalar

A scalar a is a single number.

A vector a is a $n \times 1$ list of numbers, typically arranged in a column. We write this as

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

2. Vector

Equivalently, a vector a is an element of Euclidean k space, hence $a \in R^n$.
If $n = 1$ then a is a scalar.

A matrix A is a $n \times k$ rectangular array of numbers, written as

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

3. Vector to Matrix

By convention a_{ij} refers to the i'th row and j'th column of A. If $n = 1$ or $k = 1$ then A is a vector. If $n = k = 1$, then A is a scalar.

A matrix can be written as a set of column vectors or as a set of row vectors. That is,

$$\mathrm{A} = [a_1 \quad a_2 \quad \cdots \quad a_k] = \begin{bmatrix} \alpha'_1 \\ \alpha'_2 \\ \vdots \\ \alpha'_n \end{bmatrix}$$

3. Vector to Matrix

where

$$a_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

are column vectors and

$$\alpha'_j = [a_{j1} \quad a_{j2} \quad \cdots \quad a_{jk}]$$

are row vectors.

4. The Transpose of a Matrix

The transpose of a matrix, denoted $B = A'$, is obtained by flipping the matrix on its diagonal.

$$B = A' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{nk} \end{bmatrix}$$

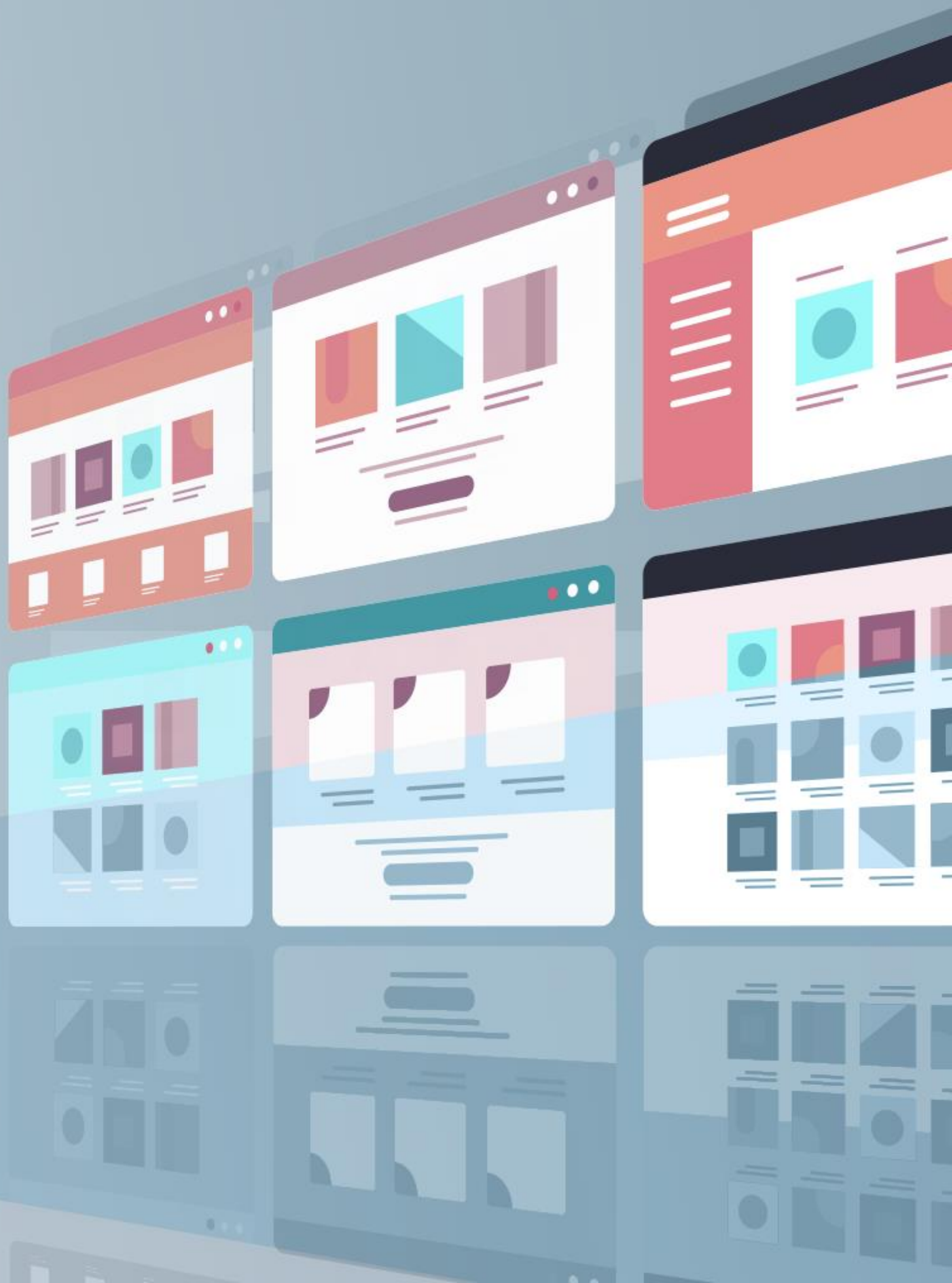
Thus $b_{ij} = a_{ji}$ for all i and j. Note that if A is $\mathrm{n} \times k$, then A′ is $k \times n$. If a is a $\mathrm{n} \times 1$ vector, then a' is a $1 \times n$ row vector.

5. Square Matrix

A matrix is square if $n = k$. A square matrix is symmetric if $A = A'$, which implies $a_{ij} = a_{ji}$. A square matrix is diagonal if the only non-zero elements appear on the diagonal, so that $a_{ij} = 0$ if $i \neq j$. A square matrix is upper (lower) diagonal if all elements below (above) the diagonal equal zero.

02

Matrix Multiplication



1. Matrix Multiplication

An alternative way to write the matrix product is to use matrix partitions. For example,

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

1. Matrix Multiplication

and

$$\mathrm{AB} = [A_1 \quad A_2 \quad \cdots \quad A_k] \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix}$$

$$= A_1B_1 + A_2B_2 + \cdots + A_kB_k$$

$$= \sum_{j=1}^{k} \mathrm{A_r B_r}$$

2. Euclidean Norm

The Euclidean norm of an $m \times 1$ vector a is

$$|a| = (a'a)^{1/2} = \left(\sum_{i=1}^{m} a_i^2\right)^{1/2}.$$

For $A = \begin{bmatrix} a & \mathrm{b} \\ c & \mathrm{d} \end{bmatrix}$, $vec(A) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$

If A is a $m \times n$ matrix, then its Euclidean norm is

$$|A| = tr(A'A)^{1/2} = (vec(A)'vec(A))^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}.$$

3. Identity Matrix

An important diagonal matrix is identity matrix, which has ones on the diagonal. A $n \times n$ identity matrix is denoted as

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Important properties are that if A is $n \times k$, then $AI_k = A$ and $I_nA = A$.

4. Orthogonal

We say that two vectors $a\ and\ b$ are orthogonal if $a'b = 0$. The columns of a $\mathrm{n} \times k$ matrix A, $k \leq n$, are said to be orthogonal if $\mathrm{A'A} = I_k$. A square matrix A is called orthogonal if $\mathrm{A'A} = I_n$.

03

Trace, Inverse, Determinant

1. Trace

The trace of a n×n square matrix A is the sum of its diagonal elements

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

Some straightforward properties are

$$\begin{aligned} tr(cA) &= ctr(A) \\ tr(A') &= tr(A) \\ tr(A+B) &= tr(A) + tr(B) \\ tr(I_k) &= K \\ tr(AB) &= tr(BA) \end{aligned}$$

1. Trace

The last result follows since

$$tr(AB) = tr\begin{bmatrix} a'_1 b_1 & a'_1 b_2 & \cdots & a'_1 b_n \\ a'_2 b_1 & a'_2 b_2 & \cdots & a'_2 b_n \\ \vdots & \vdots & & \vdots \\ a'_n b_1 & a'_n b_2 & \cdots & a'_n b_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} a_i' b_i$$

$$= \sum_{i=1}^{n} b_i' a_i$$

$$= tr(BA).$$

2. Inverse

A $n \times n$ matrix A has full rank, or is nonsingular, if there is no $\mathrm{C} \neq 0$ such that $\mathrm{AC} = 0$. In this case there exists a unique matrix B such that $AB = BA = I_n$. This matrix is called the inverse of A and is denoted by A^{-1}. Some properties include

$$
\begin{aligned}
AA^{-1} &= A^{-1}A = I_k \\
(A^{-1})' &= (A')^{-1} \\
(AC)^{-1} &= C^{-1}A^{-1} \\
(A+C)^{-1} &= A^{-1}(A^{-1}+C^{-1})^{-1}C^{-1} \\
A^{-1}-(A+C)^{-1} &= A^{-1}(A^{-1}+C^{-1})A^{-1}
\end{aligned}
$$

Also, if A is an orthogonal matrix, then $A^{-1} = A$.

3. Determinant

For a general $n \times n$ matrix $\mathrm{A} = [a_{\mathrm{ij}}]$, we can define the determinant as follows. Let $\pi = (\mathrm{j}_1, \cdots, \mathrm{j}_n)$ denote a permutation of $(1, \cdots, n)$. There are $\mathrm{n}!$ such permutations. There is a unique count of the number of inversions of the indices of such permutations (relative to the natural order $(1, \cdots, n)$), and let $\varepsilon_\pi = +1$ if this count is even and $\varepsilon_\pi = -1$ if the count is odd. Then

$$\mathrm{Det}\ \mathrm{A} = \sum_{\pi} \varepsilon_\pi a_{1\mathrm{j}_1} a_{2\mathrm{j}_2} \cdots a_{n\mathrm{j}_n}$$

3. Determinant

Example)

For a 3×3 matrix $A = [a_{ij}]$,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \Rightarrow$$

permutation	ε_π	*product(to be summed)*
$\{1, 2, 3\}$	0	$+a_{11}a_{22}a_{33}$
$\{1, 3, 2\}$	1	$-a_{11}a_{23}a_{32}$
$\{2, 1, 3\}$	1	$-a_{12}a_{21}a_{33}$
$\{2, 3, 1\}$	2	$+a_{12}a_{23}a_{31}$
$\{3, 1, 2\}$	2	$+a_{13}a_{21}a_{32}$
$\{3, 2, 1\}$	1	$-a_{13}a_{22}a_{31}$

$$\text{Det } A = \sum_{\pi} \varepsilon_\pi a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

3. Determinant

Some properties include

- $\det A = \det A'$
- $\det(\alpha A) = \alpha^n \det A$
- $\det(AB) = (\det A)(\det B)$
- $\det(A^{-1}) = (\det A)^{-1}$
- $\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B)$ if $\det A$ *is invertible.*
- $\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(D)\det(A - BD^{-1}C)$ if $\det D$ *is invertible.*
- $\det A \neq 0$ if and only if A is nonsingular.
- If A is triangular (upper or lower), then $\det A = \prod_{i=1}^{n} a_{ii}$
- If A is orthogonal, then $\det A = \pm 1$

$\det(A)\det\begin{bmatrix} A & B \\ C & D \end{bmatrix}\det\begin{bmatrix} A^{-1} & -A^{-1}B \\ 0 & I \end{bmatrix}$

$\det\begin{bmatrix} I & 0 \\ -D^{-1}C & D^{-1} \end{bmatrix}$

04

Matrix Calculus

1. Matrix Calculus

Let $x = x_1, \cdots, x_n$ be $n \times 1$ and $g(x) = g(x_1, \cdots, x_n)$: $\mathrm{R}^n \to \mathrm{R}$. The vector derivative is

$$\frac{\partial}{\partial x} \mathrm{g}(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \mathrm{g}(x) \\ \vdots \\ \frac{\partial}{\partial x_n} \mathrm{g}(x) \end{pmatrix}$$

and

$$\frac{\partial}{\partial x'} \mathrm{g}(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \mathrm{g}(x) & \cdots & \frac{\partial}{\partial x_n} \mathrm{g}(x) \end{pmatrix}.$$

1. Matrix Calculus

Some properties are now summarized.

- $\frac{\partial}{\partial x}(a'x) = \frac{\partial}{\partial x}(x'a) = a$
- $\frac{\partial}{\partial x}(Ax) = A$
- $\frac{\partial}{\partial x}(x'Ax) = (A + A')x$
- $\frac{\partial^2}{\partial x \partial x'}(x'Ax) = (A + A')$

1. Matrix Calculus

Some properties are now summarized.

- $\mathrm{A} = [a_{ij}]$ be $m \times n$ and $g(\mathrm{A}): \to R^{mn}\ R$. We define

$$\frac{\partial}{\partial A} g(A) = \left[\frac{\partial}{\partial a_{ij}} g(A)\right]$$

- $\frac{\partial}{\partial \mathrm{A}}(x'\mathrm{A}x) = xx'$
- $\frac{\partial}{\partial \mathrm{A}}\ln(\mathrm{A}) = (\mathrm{A}^{-1})'$
- $\frac{\partial}{\partial \mathrm{A}}\mathrm{tr}(A\mathrm{B}) = \mathrm{B}'$

05

Some Features of Matrix

1. Eigenvalues and Eigenvectors

For any $n \times n$ matrix A, the roots of the nth degree polynomial equation in λ, $|\lambda \mathrm{I} - \mathrm{A}| = 0$, denoted by $\lambda_1, \lambda_2, \cdots, \lambda_n$, are called the *eigenvalues* of the matrix A. The collection $\{\lambda_1, \cdots, \lambda_n\}$ of eigenvalues of A is called the *spectrum* of A.

1. Eigenvalues and Eigenvectors

Any nonzero $n \times 1$ vector $x_i \neq 0$ such that $\mathrm{A}x_i = \lambda_i x_i$ is called an *eigenvector* of A corresponding to the eigenvalue λ_i. It follows that A is nonsingular (i.e., A^{-1} exists) if and only if A has no eigenvalues equal to zero (since a zero eigenvalue would imply that $|\mathrm{A}| = 0$).

2. Diagonalization of a Symmetric Matrix

For any $n \times n$ symmetric matrix A, that is, $\mathrm{A}' = \mathrm{A}$, there exists an orthogonal matrix P such that $\mathrm{P}'\mathrm{AP} = \Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_\mathrm{n})$, where the λ_i are the eigenvalues of A. The corresponding eigenvectors of A are the column vectors of the matrix P.

3. Quadratic Forms

For an $n \times n$ constant (symmetric) matrix $\mathrm{A} = \{a_{ij}\}$, the quadratic function of n variables x, where $x = (x_1, \cdots, x_n)'$ denotes an $n \times 1$ vector, defined by $Q(x) = x'\mathrm{A}x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j$, is called a $quadratic\ form$ with matrix A.

4. Idempotent Matrices

An $n \times n$ matrix A is idempotent if $\mathrm{A}^2 = \mathrm{AA} = \mathrm{A}$. The idempotent matrices that we will consider will also be symmetric matrices, and then the symmetric idempotent matrix A is also referred to as a *projection* matrix. The eigenvalues of an idempotent matrix are either zero or one, since $\lambda x = \mathrm{A}x \equiv \mathrm{A}^2 x = \mathrm{A}(\mathrm{A}x) = \mathrm{A}(\lambda x) = \lambda^2 x$ implies that $\lambda = \lambda^2$ so that $\lambda = 0$ or 1.

다음시간 안내

02

R 개요