01강 Matrix Algebra Basics

01 Basic Concepts

1. Scalar

- A scalar *α* is a single number.
- A vector *α* is a *n×1* list of numbers, typically arranged in a column.

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

2. Vector

- Equivalently, a vector *α* is an element of Euclidean k space,
  hence
  $\alpha \in R^n$
  If n=1 then *α* is a scalar
- A matrix A is a n×k rectangular array of numbers, written as

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

3. Vector to Matrix

- By convention,
  α_{ij}
  refers to the i th row and j th column of matrix A.
- A matrix can be written as a set of column vectors or as a set of row vectors.

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

are column vectors and

$$a_j' = \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jk} \end{bmatrix}$$

are row vectors

4. The Transpose of a Matrix(전치행렬)

- 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 *n*×*k*, then A' is *k*×*n*.
- If *α* is n×1 vector, then *α′* is a 1×n row vector.

5. Square Matrix(정방행렬)

- 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 ≠ j.
- A square matrix is upper(lower) diagonal if all elements below(above) the diagonal equal zero.

02 Matrix Multiplication

1. Matrix Multiplication(벡터의 곱)

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

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

- And

$$AB = \begin{bmatrix} A_1 & A_2 & \cdots & A_k \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix}$$
$$= A_1B_1 + A_2B_2 + \cdots + A_kB_k$$
$$= \sum_{j=1}^{k} A_jB_j$$

2. Euclidean Norm(벡터의 크기)

- The Euclidean norm of an m×1 vector α is

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

- if A is m×n matrix, then its Euclidean norm is

$$|a| = tr(A'A)^{1/2} = (vec(A)'vec(A))^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$$
$$For A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, vec(A) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

3. Identity Matrix(항등행렬)

- An important diagonal matrix is identity matrix, which has ones on the diagonal.
- A n×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×k, then
  $AI_k = A \;\; and \;\; I_nA = A$
  .

4. Orthogonal(직교)

- We say that two vectors *a and b* are orthogonal if *a'b=0*.
- The columns of a n×k matrix A, k≤n, are said to be orthogonal if
  $A'A = I_k$
  .
- A square matrix A is called orthogonal if
  $A'A = I_n$

03 Trace, Inverse, Determinant

1. Trace(대각선 원소의 합)

- 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

$$tr(cA) = c \cdot tr(A)$$

$$tr(A') = tr(A)$$

$$tr(A+B) = tr(A) + tr(B)$$

$$tr(I_k) = K$$

$$tr(AB) = tr(BA)$$

2. Inverse(역행렬)

- A *n*×*n* matrix A has full rank, or is nonsingular, if there is no C≠0 such that 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

$$AA^{-1} = A^{-1}A = I_k$$

$$(A^{-1})' = (A')^{-1}$$

$$(AC)^{-1} = C^{-1}A^{-1}$$

$$(A+C)^{-1} = A^{-1}(A^{-1} + C^{-1})^{-1}C^{-1}$$

$$A^{-1} - (A+C)^{-1} = A^{-1}(A^{-1} + C^{-1})A^{-1}$$

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

3. Determinant(행렬식)

- For a general *n*×*n* matrix $A = [a_{ij}]$, we can define the determinant as follows.
    - Let $\pi = (j_1, \cdots, j_n)$ denote a permutation of $(1, \cdots, n)$.
- There are 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

$$DetA = \sum_\pi \varepsilon_\pi a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

- Examples) For a 3×3 matrix $A = [a_{ij}]$,

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

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

$$DetA = \sum_{\pi} \varepsilon_{\pi} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

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

04 Matrix Calculus

1. Matrix Calculus

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

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

and

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

- Some properties are now summarized.
  - $\frac{\partial}{\partial \vec{x}}(a'\vec{x}) = \frac{\partial}{\partial \vec{x}}(\vec{x}'a) = a$
  - $\frac{\partial}{\partial \vec{x}}(A\vec{x}) = A$

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

05 Some Features of Matrix

1. Eigenvalues and Eigenvectors(고유값과 고유벡터)

- For any *n×n* matrix A, the roots of the *n*th degree polynomial equation in λ,
  $det(\lambda I - A) = |\lambda I - A| = 0$
  , denoted by
  $\lambda_1, \lambda_2, \cdots, \lambda_n$
  , are called the ***eigenvalues*** of the matrix A.
- *n×n*인 행렬 A에 대해, 람다(λ)의 n차 다항식
  $det(\lambda I - A) = |\lambda I - A| = 0$
  의 근을
  $\lambda_1, \lambda_2, \cdots, \lambda_n$
  라고 표시하고, 행렬 A의 ***고유값***이라고 부른다.
- Any nonzero *n×1* vector
  $x_i \neq 0$
  such that
  $Ax_i = \lambda_i x_i$
  is called an *eigenvector* of A corresponding to the eigenvalue
  λ_i
  $x_i \neq 0$
- 일 때,
  $Ax_i = \lambda_i x_i$
  를 만족하는 *n×1* vector를 행렬 A의 고유벡터라고 부른다. 고유벡터는 행렬의 A의 고유값에 대응하는 벡터이다.
- 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
  $|A| = 0$
).
- 이것은 A가 비특이(nonsingular)할 때 성립한다.
  - 행렬이 비특이하다는 것은 "행렬의 역행렬이 존재한다" 또는 "행렬의 열들이 선형독립이다" 또는 "행렬의 행렬식(determinant)이 0이 아니다"는 의미이다.
- A가 0과 같은 고유값이 없는 경우에만 A가 비특이하다는 것을 의미이다.
  - 행렬 A의 고유값이 0이면
    $|A| = 0$
    이다.

2. Diagonalization of a Symmetric Matrix(대칭 행렬의 대각화)

- For any *n*×*n* symmetrix matrix A, that is, A'=A, there exists an orthogonal matrix P such that
  $P'AP = \Lambda = diag(\lambda_1, \cdots, \lambda_n)$
  , where the
  λ_1
  are the eigenvalues of A.
- A'=A인 *n*×*n*인 대칭 행렬에 대해
  $P'AP = \Lambda = diag(\lambda_1, \cdots, \lambda_n)$
  를 만족하는 직교행렬 P가 존재하면
  λ_i
  는 행렬 A의 고유값들이다.
  - P가 직교행렬이라는 것은
    $P'P = I$
    를 만족하는 것을 의미한다.
- The corresponding eigenvectors of A are the column vectors of the matrix P.
- 행렬 A의 대응하는 고유벡터는 행렬 P의 열벡터이다.

3. Quadratic Forms(이차형식)

- For an *n*×*n* constant (symmetric) matrix
  $A = a_{ij}$
  , the quadratic function of n variables x, where denoted an *n×1* vector, defined by
  $$Q(\vec{x}) = \vec{x}' A \vec{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$
  , is called a *quadratic form* with matrix A.
- n×n인 상수(대칭) 행렬 A의 경우, n개 변수를 가지는 벡터 x의 이차함수는 행렬 A의 "이차형식"이라고 한다.
  - 행렬 A의 이차형식은
    $$Q(\vec{x}) = \vec{x}' A \vec{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$
    로 표현한다.

4. Idempotent Matrices(멱등 행렬)

- An n×n matrix A is idempotent if
  $A^2 = AA = A$
  .
  $A^2 = AA = A$
- 인 행렬 A는 멱등 행렬이다.
- 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.

- 멱등 행렬이면서 대칭인 행렬은 투영행렬이라고도 합니다.
- The eigenvalues of an idempotent matrix are either zero or one, since
  $\lambda x = Ax \equiv A^2 x = A(Ax) = A(\lambda x) = \lambda^2 x$
  implies that
  $\lambda = \lambda^2$
  so that λ=0 or 1.
- 멱등 행렬의 고유값은 0 또는 1이다.
  - 왜냐하면
    $\lambda x = Ax \equiv A^2 x = A(Ax) = A(\lambda x) = \lambda^2 x$
    이므로,
    $\lambda = \lambda^2$
    이고, 그러면 λ=0 or 1 이기 때문이다.