Matrix Algebra

- vector (向量)
  - An array \( \mathbf{x} \) of \( n \) real numbers \( x_1, x_2, \ldots, x_n \) is called a vector, and it is written as
    \[
    \mathbf{x} = \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
    \end{bmatrix}
    \]
  - Two basic operations:
    1. Scalar multiplication: \( c \mathbf{x} \) is the vector obtained by multiplying each element of \( \mathbf{x} \) by \( c \).
    2. Addition: \( \mathbf{x} + \mathbf{y} = \begin{bmatrix}
    x_1 + y_1 \\
    x_2 + y_2 \\
    \vdots \\
    x_n + y_n
    \end{bmatrix} \]
  - A vector has both length and direction:
    - \( L_{\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \|\mathbf{x}\| \)
    - \( L_{c\mathbf{x}} = |c|L_{\mathbf{x}} \)
  - Multiplication by \( c \) does not change the direction of \( \mathbf{x} \).
  - Unit vectors on the direction of \( \mathbf{x} \):
    \[ L_{\mathbf{x}}^{-1}\mathbf{x} \]
    Length of \( L_{\mathbf{x}}^2\mathbf{x} = 1 \).

- Vector space

**Definition 2A.4.** The space of all real \( m \)-tuples, with scalar multiplication and vector addition as just defined, is called a vector space.

**Definition 2A.5.** The vector \( \mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k \) is a linear combination of the vectors \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \). The set of all linear combinations of \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \), is called their linear span.

**Definition 2A.6.** A set of vectors \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \) is said to be linearly dependent if there exist \( k \) numbers \( (a_1, a_2, \ldots, a_k) \), not all zero, such that
\[
a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k = \mathbf{0}
\]
Otherwise, the set of vectors is said to be linearly independent.

**Definition 2A.7.** Any set of \( m \) linearly independent vectors is called a basis for the vector space of all \( m \)-tuples of real numbers.

**Result 2A.1.** Every vector can be expressed as a unique linear combination of a fixed basis.

- Inner product (內積)
  \[ \mathbf{x}' \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n \]
  \[ = L_{\mathbf{x}}L_{\mathbf{y}} \cos(\theta) \]
  - Length: \( L_{\mathbf{x}} = \text{length of } \mathbf{x} = \sqrt{x' \mathbf{x}} \)
angle between 2 vectors:
\[
\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_x L_y} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}} \sqrt{\mathbf{y}'\mathbf{y}}}
\]

projection of \( \mathbf{x} \) on \( \mathbf{y} \):
\[
\text{Projection of } \mathbf{x} \text{ on } \mathbf{y} = \left( \frac{\mathbf{x}'\mathbf{y}}{\mathbf{y}'\mathbf{y}} \right) \mathbf{y} = \left( \frac{\mathbf{x}'\mathbf{y}}{L_y} \right) \mathbf{y} = \mathbf{L}_x |\cos(\theta)|
\]

Length of projection
\[
\frac{|\mathbf{x}'\mathbf{y}|}{L_y} = \frac{\mathbf{x}'\mathbf{y}}{L_x L_y} = \mathbf{L}_x |\cos(\theta)|
\]

• matrix

Definition 2A.13. An \( m \times k \) matrix, generally denoted by a boldface uppercase letter such as \( \mathbf{A}, \mathbf{R}, \Sigma \), and so forth, is a rectangular array of elements having \( m \) rows and \( k \) columns.

\[
\mathbf{A} = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1k} \\
  a_{21} & a_{22} & \cdots & a_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mk}
\end{bmatrix}
\]

\[
\mathbf{A}_{(m \times k)} = \{a_{ij}\}
\]

\[
\mathbf{A} \times \mathbf{X} = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n \text{ where } a_i \text{ is the } i \text{th column of } \mathbf{A}
\]

Definition 2A.19. Consider the \( m \times k \) matrix \( \mathbf{A} \) with arbitrary elements \( a_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, k \). The transpose of the matrix \( \mathbf{A} \), denoted by \( \mathbf{A}' \), is the \( k \times m \) matrix with elements \( a_{ij}, j = 1, 2, \ldots, k, i = 1, 2, \ldots, m \). That is, the transpose of the matrix \( \mathbf{A} \) is obtained from \( \mathbf{A} \) by interchanging the rows and columns.

\[\mathbf{A}' = \begin{bmatrix}
  a_1 \\
  a_2 \\
  \vdots \\
  a_m
\end{bmatrix}
\]

addition

Definition 2A.16 (Matrix addition). Let the matrices \( \mathbf{A} \) and \( \mathbf{B} \) both be of dimension \( m \times k \) with arbitrary elements \( a_{ij} \) and \( b_{ij} \), \( i = 1, 2, \ldots, m, j = 1, 2, \ldots, k \), respectively. The sum of the matrices \( \mathbf{A} \) and \( \mathbf{B} \) is an \( m \times k \) matrix \( \mathbf{C} \), written \( \mathbf{C} = \mathbf{A} + \mathbf{B} \), such that the arbitrary element of \( \mathbf{C} \) is given by

\[
c_{ij} = a_{ij} + b_{ij} \quad i = 1, 2, \ldots, m, j = 1, 2, \ldots, k
\]

scalar multiplication

Definition 2A.17 (Scalar multiplication). Let \( c \) be an arbitrary scalar and \( \mathbf{A} = \{a_{ij}\} \).

Then \( \mathbf{cA} = \mathbf{Ac} = \mathbf{B} = \{b_{ij}\}, \) where \( b_{ij} = ca_{ij} = a_{ij}c, \ i = 1, 2, \ldots, m, j = 1, 2, \ldots, k \).

matrix multiplication

Definition 2A.23 (Matrix multiplication). The product \( \mathbf{AB} \) of an \( m \times n \) matrix \( \mathbf{A} = \{a_{ij}\} \) and an \( n \times k \) matrix \( \mathbf{B} = \{b_{ij}\} \) is the \( m \times k \) matrix \( \mathbf{C} \) whose elements are

\[
c_{ij} = \sum_{\ell=1}^{n} a_{i\ell} b_{\ell j} \quad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, k
\]

\[
\mathbf{A} \times \mathbf{B} = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} & a_{14} \\
  a_{21} & a_{22} & a_{23} & a_{24} \\
  a_{m1} & a_{m2} & a_{m3} & a_{m4}
\end{bmatrix}
\begin{bmatrix}
  b_{11} & b_{12} & b_{13} & \cdots & b_{1p} \\
  b_{21} & b_{22} & b_{23} & \cdots & b_{2p} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  b_{n1} & b_{n2} & b_{n3} & \cdots & b_{np}
\end{bmatrix}
= \mathbf{Row}_i \left[ \begin{array}{cccc}
  a_{11}b_{1j} + a_{12}b_{2j} + a_{13}b_{3j} + a_{14}b_{4j} \\
  \vdots \\
  a_{m1}b_{1j} + a_{m2}b_{2j} + a_{m3}b_{3j} + a_{m4}b_{4j}
\end{array} \right]
\]

\[
eq \mathbf{A} \times \mathbf{B} \neq \mathbf{B} \times \mathbf{A}
\]

\[
\text{in general, } \mathbf{AB} \neq \mathbf{BA}
\]
some properties of matrix operations

Result 2A.4. For all matrices \( A, B, \) and \( C \) (of equal dimension) and scalars \( c \) and \( d \), the following hold:

(a) \( (A + B) + C = A + (B + C) \)

(b) \( A + B = B + A \)

(c) \( c(A + B) = cA + cB \)

(d) \( (c + d)A = cA + dA \)

(e) \( (A + B)' = A' + B' \)

(That is, the transpose of the sum is equal to the sum of the transposes.)

(f) \( (cd)A = c(dA) \)

(g) \( (cA)' = cA' \)

Result 2A.5. For all matrices \( A, B, \) and \( C \) (of dimensions such that the indicated products are defined) and a scalar \( c \),

(a) \( c(AB) = (cA)B \)

(b) \( A(BC) = (AB)C \)

(c) \( A(B + C) = AB + AC \)

(d) \( (B + C)A = BA + CA \)

(e) \( (AB)' = B'A' \)

More generally, for any \( x_j \) such that \( Ax_j \) is defined,

\[
(f) \quad \sum_{j=1}^{n} Ax_j = A \sum_{j=1}^{n} x_j
\]

\[
(g) \quad \sum_{j=1}^{n} (Ax_j)(A^t)' = A \left( \sum_{j=1}^{n} x_j x_j^t \right) A'
\]

Definition 2A.25. The row rank of a matrix is the maximum number of linearly independent rows, considered as vectors (that is, row vectors). The column rank of a matrix is the rank of its set of columns, considered as vectors.

Result 2A.6. The row rank and the column rank of a matrix are equal.

rank of a matrix is either the row rank or the column rank

- square matrix: \# of rows = \# of columns \((m = n)\)
  - a square matrix is said to be symmetric if \( a_{ij} = a_{ji} \) \((A' = A)\)
  - identity matrix \( I \): the square matrix with ones on the diagonal and zero elsewhere

\[
I = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

singularity

Definition 2A.26. A square matrix \( A \) is nonsingular if \( Ax = 0 \) implies that \( x = 0 \). If a matrix fails to be nonsingular, it is called singular. Equivalently, a square matrix is nonsingular if its rank is equal to the number of rows (or columns) it has.
inverse of a square matrix

**Result 2A.7.** Let \( A \) be a nonsingular square matrix of dimension \( k \times k \). Then there is a unique \( k \times k \) matrix \( B \) such that

\[
AB = BA = I
\]

where \( I \) is the \( k \times k \) identity matrix.

**Definition 2A.27.** The \( B \) such that \( AB = BA = I \) is called the inverse of \( A \) and is denoted by \( A^{-1} \). In fact, if \( BA = I \) or \( AB = I \), then \( B = A^{-1} \), and both products must equal \( I \).

**Result 2A.9.** For a square matrix \( A \) of dimension \( k \times k \), the following are equivalent:

(a) \[
A \cdot x = 0 \quad \text{implies} \quad x = 0 \quad (A \text{ is nonsingular}).
\]

(b) \[
|A| \neq 0.
\]

(c) There exists a matrix \( A^{-1} \) such that \( AA^{-1} = A^{-1}A = I \).

**Result 2A.10.** Let \( A \) and \( B \) be square matrices of the same dimension, and let the indicated inverses exist. Then the following hold:

(a) \[
(A^{-1})' = (A')^{-1}
\]

(b) \[
(AB)^{-1} = B^{-1}A^{-1}
\]

orthogonal square matrix

**Definition 2A.29.** A square matrix \( A \) is said to be orthogonal if its rows, considered as vectors, are mutually perpendicular and have unit lengths; that is, \( A'A = I \).

**Result 2A.13.** A matrix \( A \) is orthogonal if and only if \( A^{-1} = A' \). For an orthogonal matrix, \( AA' = A'A = I \), so the columns are also mutually perpendicular and have unit lengths.

**Definition 2A.24.** The determinant of the square \( k \times k \) matrix \( A = \{a_{ij}\} \), denoted by \( |A| \), is the scalar

\[
|A| = a_{11} \quad \text{if} \quad k = 1
\]

\[
|A| = \sum_{j=1}^{k} a_{1j} |A_{1j}| (-1)^{1+j} \quad \text{if} \quad k > 1
\]

where \( A_{1j} \) is the \( (k-1) \times (k-1) \) matrix obtained by deleting the first row and \( j \)th column of \( A \). Also, \( |A| = \sum_{j=1}^{k} a_{ij} |A_{ij}| (-1)^{i+j} \), with the \( i \)th row in place of the first row.
**Result 2A.11.** Let $A$ and $B$ be $k \times k$ square matrices.

(a) $|A| = |A'|$

(b) If each element of a row (column) of $A$ is zero, then $|A| = 0$

(c) If any two rows (columns) of $A$ are identical, then $|A| = 0$

(d) If $A$ is nonsingular, then $|A| = 1/|A^{-1}|$; that is, $|A||A^{-1}| = 1$.

(e) $|AB| = |A||B|$

(f) $|cA| = c^k|A|$, where $c$ is a scalar.

- Eigenvalues and eigenvectors of a square matrix.

**Definition 2A.30.** Let $A$ be a $k \times k$ square matrix and $I$ be the $k \times k$ identity matrix. Then the scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ satisfying the polynomial equation $|A - \lambda I| = 0$ are called the eigenvalues (or characteristic roots) of a matrix $A$. The equation $|A - \lambda I| = 0$ (as a function of $\lambda$) is called the characteristic equation.

\[
A = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}
\]

Three roots: $\lambda_1 = 9$, $\lambda_2 = 9$, and $\lambda_3 = 18$

\[
|A - \lambda I| = \begin{vmatrix} 13 - \lambda & -4 & 2 \\ -4 & 13 - \lambda & -2 \\ 2 & -2 & 10 - \lambda \end{vmatrix} = -\lambda^3 + 36\lambda^2 - 405\lambda + 1458 = 0
\]

- For general $A$, eigenvalues could be real or complex values
- Every eigenvalue of symmetric matrix is real