Department of Economics
Seoul National University

Matrix Algebra: A Review

1 Basic Definitions and Axioms.

- *Matrix*:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})$$

  is an $m \times n$ matrix(m rows, n columns) with a_{ij} as its element in the i^{th} row and j^{th} columns for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

- $c \cdot \boldsymbol{A} = (c \cdot a_{ij})$, if c is a scalar.

- $\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})$ if $\boldsymbol{B} = (b_{ij})$ is $m \times n$.

- $\boldsymbol{AB} = (\sum_{k=1}^{n} a_{ik} b_{kj})$ if $\boldsymbol{B}$ has n rows.

- *Associative law*: $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) = \boldsymbol{ABC}$.
  But in general, matrices don't commute, i.e., $\boldsymbol{AB} \neq \boldsymbol{BA}$ in general.

- *Distributive law*: $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$.

- $\boldsymbol{A}' = (a_{ji})$ is the *transpose* of $\boldsymbol{A}$.
  - $(\boldsymbol{A}')' = \boldsymbol{A}$.
  - If $\boldsymbol{C} = (\boldsymbol{A} + \boldsymbol{B})$, $\boldsymbol{C}' = (\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$.
  - $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$, $(\boldsymbol{ABC})' = \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'$.
  - $\boldsymbol{I}' = \boldsymbol{I}$.
  - If c is a scalar, $c' = c$.
  - If c is a scalar, $(c\boldsymbol{A})' = c\boldsymbol{A}'$.
  - If $\boldsymbol{A}$ is a $n \times n$ matrix and $\boldsymbol{A} = \boldsymbol{A}'$, $\boldsymbol{A}$ is *symmetric*.

2 Sum of values.

- Denote by $\boldsymbol{i}$ a vector that contains a columns of ones. Then,

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n = \boldsymbol{i'x}.$$

- If all elements in $\boldsymbol{x}$ are equal to the same constant a, then $\boldsymbol{x} = a\boldsymbol{i}$ and

$$\sum_{i=1}^{n} x_i = \boldsymbol{i}'(a\boldsymbol{i}) = a(\boldsymbol{i'i}) = na.$$

- For any constant a and vector $\boldsymbol{x}$, $\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i = a\boldsymbol{i'x}$.
- If $a = \frac{1}{n}$, then we obtain the arithmetic mean, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\boldsymbol{i'x}$, from which it follows that $\sum_{i=1}^{n} x_i = \boldsymbol{i'x} = n\bar{x}$.
- The sum of squares of the elements in vector $\boldsymbol{x}$ is $\sum_{i=1}^{n} x_i^2 = \boldsymbol{x'x}$; while the sum of the products of the n elements in vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\sum_{i=1}^{n} x_i y_i = \boldsymbol{x'y}$.
- If $\boldsymbol{X}$ is $n \times k$, then

$$\boldsymbol{X'X} = \sum_{i=1}^{n} \boldsymbol{x_i x_i'}.$$

3 Geometry of Matrices.

- Let $\boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n \Rightarrow$ coordinates of a point in $\mathbb{R}^n$.
  Here, $\mathbb{R}^n$ is an example of a vector space.

- V is a *vector space* iff $\forall \boldsymbol{x}, \boldsymbol{y} \in V$ and $\alpha, \beta \in \mathbb{R}$, we have $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in V$, i.e., it is closed under scalar multiplication and addition.

- *Basis Vector*: A set of vectors in a vector space is a *basis* for that vector space if any vector in the vector space can be written as a linear combination of that set of vectors.

- A set of vectors $\boldsymbol{v_1}, \cdots, \boldsymbol{v_m}$ in a vector space V is said to *span* V iff $\forall \boldsymbol{v} \in V$ there exists $\alpha_1, \cdots, \alpha_m \in \mathbb{R}$ such that $\boldsymbol{v} = \sum_{i=1}^n \alpha_i v_i$.

- dim(V) = dimension of V = the smallest number of vectors that span V.

- A set of vectors is *linearly dependent* iff any one of the vectors in the set can be expressed as a linear combination of the others.

- A set of vectors $\{\boldsymbol{v_1}, \cdots, \boldsymbol{v_m}\}$ is *linearly independent* iff the only solution to $\alpha_1\boldsymbol{v_1} + \cdots + \alpha_m\boldsymbol{v_m} = 0$ is $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$.

- *Basis* of a vector space V: A set of vectors that span V and that are linearly independent.

- If $\boldsymbol{a}$ and $\boldsymbol{b}$ are both $k \times 1$, then their *inner product* is

$$\boldsymbol{a'b} = a_1b_1 + a_2b_2 + \cdots + a_kb_k = \sum_{j=1}^k a_jb_j = \boldsymbol{b'a}.$$

- **Definition.** *Norm* of a vector $\boldsymbol{a}$. $\parallel \boldsymbol{a} \parallel = \sqrt{\boldsymbol{a'a}}$ = length of $\boldsymbol{a}$.

- **Definition.** Vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are *orthogonal* iff $\boldsymbol{a'b} = \boldsymbol{b'a} = 0$.

- **Definition.** *Orthogonal basis* of V: Basis of V that consists of vectors that are orthogonal to each other.

- **Definition.** *Orthonormal basis* of V: Orthogonal basis of V in which the length of the vectors equals one, i.e., $\{\boldsymbol{e_1}, \cdots, \boldsymbol{e_m}\}$ where $\parallel \boldsymbol{e_i} \parallel = 1\ \forall i = 1, \cdots, n$ and $\boldsymbol{e_i'e_j} = 0\ \forall i \neq j$.

4 Trace.

- **Definition.** The *trace* of $k \times k$ square matrix $\boldsymbol{A}$ is the sum of its diagonal elements

$$tr(\boldsymbol{A}) = \sum_{i=1}^{k} a_{ii}.$$

- **Properties.**
  - Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices and c be a scalar.

$$tr(c \cdot \boldsymbol{A}) = c \cdot tr(\boldsymbol{A})$$

$$tr(\boldsymbol{A'}) = tr(\boldsymbol{A})$$

$$tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$$

$$tr(\boldsymbol{I_k}) = k$$

$$\boldsymbol{a'a} = tr(\boldsymbol{a'a}) = tr(\boldsymbol{aa'})$$

$$tr(\boldsymbol{A'A}) = \sum_{i=1}^{k} \boldsymbol{a}_i'\boldsymbol{a}_i = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^2.$$

  - Let $\boldsymbol{A}$ be $n \times k$ and B be $k \times n$.
    Then, we have

$$tr(\boldsymbol{AB}) = tr(\boldsymbol{BA}).$$

5 Inverse.

- **Definition.** For a square $n \times n$ matrix $\boldsymbol{A}$ having full rank, $\boldsymbol{A}^{-1}$ is defined to be a matrix $\boldsymbol{B}$ that satisfies

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I_n}$$

  $\Rightarrow$ exists only if $\boldsymbol{A}$ is full rank or, equivalently, $\boldsymbol{A}$ is nonsingular.
  $\Rightarrow$ If the inverse exists, then it must be unique.

- **Properties.**

  (i) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, $(\boldsymbol{ABC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
  if $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are nonsingular.

  (ii) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.

  (iii) $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$.

  (iv) If $\boldsymbol{A}$ is symmetric, then $\boldsymbol{A}^{-1}$ is symmetric.

6 Rank of a Matrix.

- **Definition.**

$$\begin{aligned} rank(\boldsymbol{A}) &= \text{rank of } \boldsymbol{A} \\ &= \text{maximum number of linearly independent columns} \\ &= \text{maximum number of linearly independent rows} \end{aligned}$$

- **Properties.** Let $\boldsymbol{A}$ be an $m \times n$ matrix.

  (i) $rank(\boldsymbol{A}) = rank(\boldsymbol{A'}) \leq min\{m, n\}$

  (ii) $\boldsymbol{A}$ is said to be *full rank* if $rank(\boldsymbol{A}) = min(m, n)$, i.e., *full rank* matrix is a matrix whose rank is equal to the number of columns it contains.

  (iii) $rank(\boldsymbol{AB}) \leq min\{rank(\boldsymbol{A}), rank(\boldsymbol{B})\}$

  (iv) $rank(\boldsymbol{A}) = rank(\boldsymbol{A'A}) = rank(\boldsymbol{AA'})$

  (v) $rank(\boldsymbol{A}) = rank(\boldsymbol{PA}) = rank(\boldsymbol{AQ}) = rank(\boldsymbol{PAQ})$
  provided $\boldsymbol{A}$: $m \times n,\ \boldsymbol{P} : m \times m,\ \boldsymbol{Q} : n \times n$ and $rank(\boldsymbol{P}) = m,\ rank(\boldsymbol{Q}) = n.$

- A square $k \times k$ matrix $\boldsymbol{A}$ is said to be *nonsingular* if it has full rank, i.e., $rank(\boldsymbol{A}) = k$.
  $\Leftrightarrow$ There is no $k \times 1$ scalar vector $\boldsymbol{c} \neq \boldsymbol{0}$ such that $\boldsymbol{Ac} = \boldsymbol{0}$.

7 Determinant.

- Let $\boldsymbol{A}$ be an $n \times n$ square matrix. Then,

$$\begin{aligned} |\boldsymbol{A}| &= \text{determinant of } \boldsymbol{A} \\ &= \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\boldsymbol{M}_{ij} \end{aligned}$$

  where $\boldsymbol{M}_{ij}$ = determinant of a submatrix of $\boldsymbol{A}$ with the i^{th} row and the j^{th} columns deleted.

- A shortcut for the determinant of 3×3 matrix.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{23}a_{32}.$$

- **Properties.**

  (i) $\begin{vmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{vmatrix} = c_1 \cdot c_2 \cdots c_n = \prod_{i=1}^{n} c_i.$

  (ii) $|\boldsymbol{A}| = |\boldsymbol{A'}|$ even if $\boldsymbol{A}$ is not symmetric.

  (iii) $|c \cdot \boldsymbol{A}| = c^n \cdot |\boldsymbol{A}|$, where c is a scalar.

  (iv) $|\boldsymbol{A}| = 0$ if $\boldsymbol{A}$ is not full rank. $\Leftrightarrow$ $\boldsymbol{A}$ is singular.

  (v) $|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$.

  (vi) If $\boldsymbol{A}$ is invertible and $|\boldsymbol{A}| \neq 0$, $|\boldsymbol{A}^{-1}| = \frac{1}{|\boldsymbol{A}|}$.

  (vii) If $\boldsymbol{A}$ is triangular(upper or lower), then $|\boldsymbol{A}| = \prod_{i=1}^{n} a_{ii}$.

8 Eigenvalues and Eigenvectors.

- Consider

$$\underset{n\times n}{\boldsymbol{A}}\ \underset{n\times 1}{\boldsymbol{x}} = \underset{1\times 1}{\boldsymbol{\lambda}}\ \underset{n\times 1}{\boldsymbol{x}}.$$

We have

$$(\boldsymbol{A} - \boldsymbol{\lambda I_n})\boldsymbol{x} = \boldsymbol{0}. \tag{1}$$

Note that the nontrivial solution(s) of $\boldsymbol{x}$(i.e., $x \neq 0$) exists only if

$$|\boldsymbol{A} - \boldsymbol{\lambda I_n}| = 0. \tag{2}$$

Let $\boldsymbol{\lambda}^*$ denote $\boldsymbol{\lambda}$ that satisfies (2). Then, $\boldsymbol{\lambda}^*$= eigenvalue of $\boldsymbol{A}$. Also, corresponding to $\boldsymbol{\lambda}^*$, the values of $\boldsymbol{x}^*$ that satisfies $\boldsymbol{Ax}^* = \boldsymbol{\lambda}^*\boldsymbol{x}^*$ is called the *eigenvector* of $\boldsymbol{A}$.

- **Properties.** Let $\boldsymbol{A}$ be a real *symmetric* $n \times n$ matrix.

(i) Eigenvalues of $\boldsymbol{A}$ are real.

(ii) Eigenvectors corresponding to distinct eigenvalues are pairwise orthogonal.

(iii) *Spectral Decomposition*: $\boldsymbol{A}$ can be diagonalized, i.e.,there exists an *orthogonal matrix* $\boldsymbol{X}$(i.e., $\boldsymbol{X'X} = \boldsymbol{XX'} = \boldsymbol{I_n}$ or equivalently, $\boldsymbol{X'} = \boldsymbol{X}^{-1}$) and a diagonal matrix $\boldsymbol{\Lambda} = diag(\lambda_1, \cdots, \lambda_n)$ such that $\boldsymbol{X'AX} = \boldsymbol{\Lambda}$.

(iv) $|\boldsymbol{A}| = \prod_{i=1}^n \lambda_i$.

(v) $tr(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$.

9 Positive Definite Matrix.

- **Definition.** Let $\boldsymbol{A}$ be an $n \times n$ matrix.
  - $\boldsymbol{A}$ is *positive definite*(p.d.) if $\boldsymbol{x'Ax} > 0 \ \ \forall \boldsymbol{x} \neq 0$.
  - $\boldsymbol{A}$ is *negative definite*(n.d.) if $\boldsymbol{x'Ax} < 0 \ \ \forall \boldsymbol{x} \neq 0$.
  - $\boldsymbol{A}$ is *positive semidefinite*(p.s.d.) if $\boldsymbol{x'Ax} \geq 0 \ \ \forall \boldsymbol{x} \neq 0$.
  - $\boldsymbol{A}$ is *negative semidefinite*(n.s.d.) if $\boldsymbol{x'Ax} \leq 0 \ \ \forall \boldsymbol{x} \neq 0$.

- **Properties.**

(i) Let $\boldsymbol{a}$ be an $n \times 1$ vector. Then, $\boldsymbol{A} = \boldsymbol{aa'}$ is always p.s.d.

(ii) If $\boldsymbol{A}$ is p.s.d.(p.d.), then the eigenvalues of $\boldsymbol{A} \geq 0(\boldsymbol{A} > 0)$

(iii) Let $\boldsymbol{A}$ be a real, symmetric p.s.d. $n \times n$ matrix.
Then, there exists a matrix $\boldsymbol{C}$ such that

$$\boldsymbol{A} = \boldsymbol{C}'\boldsymbol{C},$$

$\boldsymbol{C}$ is not unique.

(iv) If $\boldsymbol{A}$ is p.s.d., then $|\boldsymbol{A}| \geq 0$.

(v) If $\boldsymbol{A}$ is p.d., so is $\boldsymbol{A}^{-1}$.

(vi) The identity matrix $\boldsymbol{I}$ is p.d.

(vii) If $\boldsymbol{A}$ is $n \times k$ with full column rank and $n > k$, then $\boldsymbol{A'A}$ is p.d. and $\boldsymbol{AA'}$ is p.s.d.

10 Idempotent Matrix.

- **Definition.** A square $n \times n$ matrix $\boldsymbol{A}$ is *idempotent* iff

$$\boldsymbol{A} \cdot \boldsymbol{A} = \boldsymbol{A}$$

.

- **Properties.**

(i) Eigenvalues of a symmetric idempotent matrix $\boldsymbol{A}$ are 0 or 1.

(ii) Any symmetric idempotent matrix $\boldsymbol{A}$ is p.s.d.

(iii) Let $\boldsymbol{A}$ be symmetric and idempotent. Then, $rank(\boldsymbol{A}) = tr(\boldsymbol{A})$.

(iv) If $\boldsymbol{A}$ is idempotent, then so is $\boldsymbol{I_n} - \boldsymbol{A}$.

11 Partitioned Matrix.

- Let
$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{B_{11}} & \boldsymbol{B_{12}} \\ \boldsymbol{B_{21}} & \boldsymbol{B_{22}} \end{bmatrix}.$$

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

- $\begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{22}} \end{bmatrix}' \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{22}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A'_{11}A_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A'_{22}A_{22}} \end{bmatrix}$

- $\begin{vmatrix} \boldsymbol{A_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{22}} \end{vmatrix} = \left|\boldsymbol{A_{11}}\right| \cdot \left|\boldsymbol{A_{22}}\right|,$

  $\begin{vmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{vmatrix} = \left|\boldsymbol{A_{22}}\right| \cdot \left|\boldsymbol{A_{11}} - \boldsymbol{A_{12}A_{22}^{-1}A_{21}}\right| = \left|\boldsymbol{A_{11}}\right| \cdot \left|\boldsymbol{A_{22}} - \boldsymbol{A_{21}A_{11}^{-1}A_{12}}\right|$

- $\begin{vmatrix} \boldsymbol{A_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{22}} \end{vmatrix}^{-1} = \begin{vmatrix} \boldsymbol{A_{11}^{-1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{22}^{-1}} \end{vmatrix},$

  $$\begin{aligned} \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^{-1} &= \begin{bmatrix} \boldsymbol{C_{11}} & -\boldsymbol{C_{11}A_{12}A_{22}^{-1}} \\ -\boldsymbol{A_{22}^{-1}A_{21}C_{11}} & \boldsymbol{A_{22}^{-1}} + \boldsymbol{A_{22}^{-1}A_{21}C_{11}A_{12}A_{22}^{-1}} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{A_{11}^{-1}} + \boldsymbol{A_{11}^{-1}A_{12}C_{22}A_{21}A_{11}^{-1}} & -\boldsymbol{A_{11}^{-1}A_{12}C_{22}} \\ -\boldsymbol{C_{22}A_{21}A_{11}^{-1}} & \boldsymbol{C_{22}} \end{bmatrix} \end{aligned}$$

  , where $\boldsymbol{C_{11}} = (\boldsymbol{A_{11}} - \boldsymbol{A_{12}A_{22}^{-1}A_{21}})^{-1}, \boldsymbol{C_{22}} = (\boldsymbol{A_{22}} - \boldsymbol{A_{21}A_{11}^{-1}A_{12}})^{-1}.$

12 Kronecker Product.

- Let $\underset{l\times m}{\boldsymbol{A}} = (a_{ij}), \quad \underset{n\times k}{\boldsymbol{B}} = (b_{ij})$.

$$\boldsymbol{A}\otimes\boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1m}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1}\boldsymbol{B} & a_{l2}\boldsymbol{B} & \cdots & a_{lm}\boldsymbol{B} \end{bmatrix} \ : \ ln \times mk$$

- $(\boldsymbol{A}\otimes\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$ if $\boldsymbol{A}$ and $\boldsymbol{B}$ are nonsingular.

$(\alpha\boldsymbol{A}\otimes\beta\boldsymbol{B}) = \alpha\beta(\boldsymbol{A}\otimes\boldsymbol{B})$

$(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C} = \boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})$

$(\boldsymbol{A}+\boldsymbol{B})\otimes\boldsymbol{C} = (\boldsymbol{A}\otimes\boldsymbol{C}) + (\boldsymbol{B}\otimes\boldsymbol{C})$

$(\boldsymbol{A}\otimes\boldsymbol{B})' = \boldsymbol{A}'\otimes\boldsymbol{B}'$

$(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) = \boldsymbol{AC}\otimes\boldsymbol{BD}$

$tr(\boldsymbol{A}\otimes\boldsymbol{B}) = tr(\boldsymbol{A})\cdot tr(\boldsymbol{B})$

$|\boldsymbol{A}\otimes\boldsymbol{B}| = |\boldsymbol{A}|^l\cdot|\boldsymbol{B}|^n$ if $l = m$ and $n = k$.

13 Differentiation of Matrix.

- Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = f(x) = f(x_1, \cdots, x_n)$.

  Define

$$\frac{\partial f(x)}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \; : \; n \times 1,$$

$$\frac{\partial f(x)}{\partial \boldsymbol{x'}} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \; : \; 1 \times n.$$

- Let $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}$.

  Define

$$\frac{\partial g(x)}{\partial \boldsymbol{x'}} = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(x)}{\partial x_1} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix} \; : \; m \times n,$$

$$\frac{\partial g'(x)}{\partial \boldsymbol{x}} = \left(\frac{\partial g(x)}{\partial \boldsymbol{x'}} \right)'.$$

- Let $\boldsymbol{y} = \boldsymbol{a'x} = \boldsymbol{x'a} = \sum x_i a_i$, where $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.

  Then,

$$\frac{\partial (\boldsymbol{a'x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial (a'x)}{\partial x_1} \\ \vdots \\ \frac{\partial (a'x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{a}.$$

- $$\boldsymbol{y} = \underset{m\times n}{\boldsymbol{A}} \underset{n\times 1}{\boldsymbol{x}} \Rightarrow \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{a_1'} \\ \vdots \\ \boldsymbol{a_m'} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{a_1' x} \\ \vdots \\ \boldsymbol{a_m' x} \end{bmatrix}$$
$$\Rightarrow y_i = \boldsymbol{a_i' x} \text{ for } i = 1, \cdots, m$$
$$\Rightarrow \frac{\partial y_i}{\partial \boldsymbol{x'}} = \boldsymbol{a_i'}$$

$$\therefore \frac{\partial(\boldsymbol{Ax})}{\partial \boldsymbol{x'}} = \begin{bmatrix} \frac{\partial y_1}{\partial x'} \\ \vdots \\ \frac{\partial y_m}{\partial x'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a_1'} \\ \vdots \\ \boldsymbol{a_m'} \end{bmatrix} = \boldsymbol{A}.$$

Similarly,
$$\frac{\partial(\boldsymbol{x'A'})}{\partial \boldsymbol{x}} = \boldsymbol{A'}.$$

- $$\underset{1\times n}{\boldsymbol{x'}} \underset{n\times n}{\boldsymbol{A}} \underset{n\times 1}{\boldsymbol{y}} = \sum_i \sum_j x_i y_j a_{ij}$$
$$\frac{\partial(\boldsymbol{x'Ay})}{\partial \boldsymbol{x}} = \boldsymbol{Ay}$$

$$\frac{\partial(\boldsymbol{x'Ax})}{\partial \boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{A'})\boldsymbol{x}$$
$$= 2\boldsymbol{Ax} \text{ if } \boldsymbol{A} \text{ is symmetric}$$

$$\frac{\partial(\boldsymbol{x'Ax})}{\partial \boldsymbol{A}} = \boldsymbol{xx'} \text{ since } \frac{\partial(\boldsymbol{x'Ax})}{\partial a_{ij}} = x_i x_j$$