LINEAR ALGEBRA

Kuangyu Wen

Huazhong University of Science and Technology

July 21, 2021

Contents

Pseudoinverse of matrix

Generalized inverse(pseudoinverse) Moore-Penrose

Matrix inversion is defined for square matrix with full rank. For a general matrix A:$m \times n$, the matrix A^+ is called pseudoinverse of A if the following conditions hold.

1. $\mathrm{A}\mathrm{A}^+\mathrm{A} = \mathrm{A}$
2. $\mathrm{A}^+\mathrm{A}\mathrm{A}^+ = \mathrm{A}^+$
3. $(\mathrm{A}\mathrm{A}^+)' = \mathrm{A}\mathrm{A}^+$
4. $(\mathrm{A}^+\mathrm{A})' = \mathrm{A}^+\mathrm{A}$

Note that A^+ must be $n \times m$.

Pseudoinverse of matrix

Theorem 1

Let A *be a matrix. If its pseudoinverse exists, then it is unique.*

Proof.

Suppose B and C are two pseudoinverse of A. Then we have

$$\begin{aligned}\mathrm{AB} &= (\mathrm{ACA})\mathrm{B} = (\mathrm{AC})(\mathrm{AB}) = (\mathrm{AC})^{'}(\mathrm{AB})^{'} = \mathrm{C}^{'}\mathrm{A}^{'}\mathrm{B}^{'}\mathrm{A}^{'} \\ &= \mathrm{C}^{'}(\mathrm{ABA})^{'} = \mathrm{C}^{'}\mathrm{A}^{'} = (\mathrm{AC})^{'} = \mathrm{AC}\end{aligned}$$

By the same way, we have

$$\begin{aligned}\mathrm{BA} &= \mathrm{B}(\mathrm{ACA}) = (\mathrm{BA})(\mathrm{CA}) = (\mathrm{BA})^{'}(\mathrm{CA})^{'} = \mathrm{A}^{'}\mathrm{B}^{'}\mathrm{A}^{'}\mathrm{C}^{'} \\ &= (\mathrm{ABA})^{'}\mathrm{C}^{'} = \mathrm{A}^{'}\mathrm{C}^{'} = (\mathrm{CA})^{'} = \mathrm{CA}\end{aligned}$$

Thus, $\mathrm{B} = \mathrm{BAB} = (\mathrm{BA})\mathrm{B} = (\mathrm{CA})\mathrm{B} = \mathrm{C}(\mathrm{AB}) = \mathrm{CAC} = \mathrm{C}$.
That is $\mathrm{B} = \mathrm{C}$. □

Pseudoinverse of matrix

Example 2

Suppose A is a square matrix with full rank, then A^+=?

$$A^+ = A^{-1}$$

Proof.

$AA^+ = AA^{-1} = I$
$A^+A = A^{-1}A = I$
$AA^+A = AA^{-1}A = A$
$A^+AA^+ = A^{-1}AA^{-1} = A^{-1} = A^+$ □

Pseudoinverse of matrix

Example 3

Let A= $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ be an $n \times 1$ column vector, then A^+=?

$$\mathrm{A}^+ = \lambda \mathrm{A}', \text{ where } \lambda = \frac{1}{a_1^2 + a_2^2 + \cdots + a_n^2}$$

Proof.

$\mathrm{A}\mathrm{A}^+ = \lambda \mathrm{A}\mathrm{A}' = \lambda [a_i a_j]_{n \times n}$

$\mathrm{A}^+\mathrm{A} = \lambda \mathrm{A}'\mathrm{A} = \lambda (a_1^2 + a_2^2 + \cdots + a_n^2) = 1$

$\mathrm{A}\mathrm{A}^+\mathrm{A} = \mathrm{A}(\mathrm{A}^+\mathrm{A}) = \mathrm{A}$

$\mathrm{A}^+\mathrm{A}\mathrm{A}^+ = (\mathrm{A}^+\mathrm{A})\mathrm{A}^+ = \mathrm{A}^+$ □

Pseudoinverse of matrix

Example 4

Suppose an $m \times n$ matrix $\mathrm{A} = \begin{bmatrix} \mathrm{B} & 0 \\ 0 & 0 \end{bmatrix}$, where B is an $r \times r$ non-singular matrix. Then

$$\mathrm{A}^+ = \begin{bmatrix} \mathrm{B}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

Proof.

$$\mathrm{AA}^+ = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathrm{A}^+\mathrm{A} = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}$$
$$\mathrm{AA}^+\mathrm{A} = \mathrm{A}, \qquad \mathrm{A}^+\mathrm{AA}^+ = \mathrm{A}^+$$

□

Pseudoinverse of matrix

Some basic properties of pseudoinversion

1. $(\mathrm{A}^{+})^{+} = \mathrm{A}$
2. $(\mathrm{A}^{'})^{+} = (\mathrm{A}^{+})^{'}$
3. $rank(\mathrm{A}^{+}) = rank(\mathrm{A})$

Proof.

$$rank(\mathrm{A}) = rank(\mathrm{A}\mathrm{A}^{+}\mathrm{A}) \leq rank(\mathrm{A}\mathrm{A}^{+}) \leq rank(\mathrm{A}^{+})$$
$$rank(\mathrm{A}^{+}) = rank(\mathrm{A}^{+}\mathrm{A}\mathrm{A}^{+}) \leq rank(\mathrm{A}\mathrm{A}^{+}) \leq rank(\mathrm{A})$$

So, we have $rank(\mathrm{A}) = rank(\mathrm{A}^{+})$ □

Question

How to find a formula for pseudoinverse?

Pseudoinverse of matrix

Theorem 5

Suppose that an $m \times n$ *matrix* A *has full column rank, that is* $rank(\mathrm{A}) = n$*, then* A *has a pseudoinverse.*

$$\mathrm{A}^{+} = (\mathrm{A}^{'}\mathrm{A})^{-1}\mathrm{A}^{'}$$

where A^{+} *is an* $n \times m$ *matrix.*

Proof.

First, we need to show $\mathrm{A}^{'}\mathrm{A}$ is non-singular.
Suppose $rank(\mathrm{A}^{+}\mathrm{A}) < n$, then $\mathrm{A}^{'}\mathrm{A}x = 0$ has a non-zero solution x. Then we have

$$x^{'}\mathrm{A}^{'}\mathrm{A}x = \mathrm{o}$$
$$(\mathrm{A}x)^{'}\mathrm{A}x = \mathrm{o}$$

Finally, we can get $\mathrm{A}x = 0$ has non-zero solution, A is singular. This is a contradiction. □

Proof.

Second, we have

$$\mathrm{AA}^{+} = \mathrm{A}(\mathrm{A}'\mathrm{A})^{-1}\mathrm{A}'$$
$$(\mathrm{AA}^{+})' = (\mathrm{A}(\mathrm{A}'\mathrm{A})^{-1}\mathrm{A}')' = \mathrm{A}(\mathrm{A}'\mathrm{A})^{-1}\mathrm{A}' = \mathrm{AA}^{+}$$
$$\mathrm{A}^{+}\mathrm{A} = (\mathrm{A}'\mathrm{A})^{-1}\mathrm{A}'\mathrm{A} = \mathrm{I}$$
$$\mathrm{AA}^{+}\mathrm{A} = \mathrm{AI} = \mathrm{A} \qquad \mathrm{A}^{+}\mathrm{AA}^{+} = \mathrm{IA}^{+} = \mathrm{A}^{+}$$

□

When we regress Y on X, we have $\hat{\beta} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{Y}$, where X is an $n \times k$ matrix and $rank(\mathrm{X}) = k$. Then the $\hat{\beta}$ can be also written as $\mathrm{X}^{+}\mathrm{Y}$.

Pseudoinverse of matrix

Example 6

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ -1 & 3 \end{bmatrix}$$

Solution

We have $rank(A) = 2$, then

$$A'A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 13 \end{bmatrix}$$

$$A^+ = (A'A)^{-1}A' = \begin{bmatrix} 3 & -1 \\ -1 & 13 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 3 \end{bmatrix}$$

$$= \frac{1}{38} \begin{bmatrix} 13 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 3 \end{bmatrix} = \frac{1}{38} \begin{bmatrix} 13 & 15 & -10 \\ 1 & 7 & 8 \end{bmatrix}$$

Pseudoinverse of matrix

Theorem 7

Suppose A *is an* $m \times n$ *matrix,* $rank(\mathrm{A}) = r$. *Then there exists an* $m \times r$ *matrix* F *and an* $r \times n$ *matrix* G *such that* $\mathrm{A} = \mathrm{FG}$ *and* $rank(\mathrm{F}) = rank(\mathrm{G}) = r$ *(Full rank Factorization)*

Consider any r linearly independent columns of A. Let F be the submatrix of A formed by these columns.
For each column $\mathrm{A_k}$(k-th column of A), $\mathrm{A_k} = \mathrm{FG_k}$, where $\mathrm{G_k}$ is a vector of dimension r.

$$\mathrm{G} = \begin{bmatrix} \mathrm{G}_1 & \mathrm{G}_2 & \cdots \mathrm{G}_n \end{bmatrix} \qquad r \times n$$

$$\begin{aligned} \mathrm{FG} &= \mathrm{F} \begin{bmatrix} \mathrm{G}_1 & \mathrm{G}_2 & \cdots \mathrm{G}_n \end{bmatrix} = \begin{bmatrix} \mathrm{FG}_1 & \mathrm{FG}_2 & \cdots \mathrm{FG}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 & \cdots \mathrm{A}_n \end{bmatrix} = \mathrm{A} \end{aligned}$$

Pseudoinverse of matrix

Example 8

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad rank(A) = 2$$

$$F = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix}$$

$$FG_1 = A_1 \rightarrow \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} G_1 = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} \rightarrow G_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$FG_2 = A_2 \rightarrow \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} G_2 = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} \rightarrow G_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Pseudoinverse of matrix

Example 9

$$\mathrm{FG}_3 = \mathrm{A}_3 \rightarrow \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \mathrm{G}_3 = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \rightarrow \mathrm{G}_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\mathrm{G} = \begin{bmatrix} \mathrm{G}_1 & \mathrm{G}_2 & \mathrm{G}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

Theorem 10

For an arbitrary $m \times n$ *matrix* A*, its pseudoinverse exists. If* $\mathrm{A} = \mathrm{FG}$ *is a full rank decomposition of* A*, then*

$$\mathrm{A}^{+} = \mathrm{G}'(\mathrm{G}\mathrm{G}')^{-1}(\mathrm{F}'\mathrm{F})^{-1}\mathrm{F}'.$$

Pseudoinverse of matrix

Proof.

$$\begin{aligned}
AA^+ &= FGG'(GG')^{-1}(F'F)^{-1}F' = F(F'F)^{-1}F' \\
A^+A &= G'(GG')^{-1}(F'F)^{-1}F'FG = G'(GG')^{-1}G \\
AA^+A &= F(F'F)^{-1}F'FG = FG = A \\
A^+AA^+ &= G'(GG')^{-1}GG'(GG')^{-1}(F'F)^{-1}F' \\
&= G'(GG')^{-1}(F'F)^{-1}F' = A^+
\end{aligned}$$

□

Determinant

Definition 11

Suppose A is a 2×2 matrix

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}, \end{bmatrix}$$

then the determinant of A is

$$det(\mathrm{A}) = |\mathrm{A}| = a_{11}a_{22} - a_{12}a_{21}$$

Example 12

$$\mathrm{A} = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$$

$$det(\mathrm{A}) = -2 \times 3 - 0 \times 0 = -6$$

Determinant

Some properties

Let A is a matrix which can be written as

$$\mathrm{A} = \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix}$$

- $det \begin{bmatrix} \mathrm{A}_1 + \mathrm{A}_1' & \mathrm{A}_2 \end{bmatrix} = det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix} + det \begin{bmatrix} \mathrm{A}_1' & \mathrm{A}_2 \end{bmatrix}$
- $det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 + \mathrm{A}_2' \end{bmatrix} = det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix} + det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2' \end{bmatrix}$
- $det \begin{bmatrix} c \cdot \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix} = c \cdot det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix}$
- $det \begin{bmatrix} \mathrm{A}_1 & c \cdot \mathrm{A}_2 \end{bmatrix} = c \cdot det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix}$
- $det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_2 \end{bmatrix} = -det \begin{bmatrix} \mathrm{A}_2 & \mathrm{A}_1 \end{bmatrix}$
- $det \begin{bmatrix} \mathrm{A}_1 & \mathrm{A}_1 \end{bmatrix} = 0$
- $det(\mathrm{I}_2) = 1$

Determinant

Definition 13

Let $\mathrm{A} = \left[a_{ij}\right]_{n \times m}$. The minor M_{ij} of A is the determinant of the matrix formed from A by removing the i-th row and j-th column. The cofactor $\mathrm{A}_{ij} = (-1)^{i+j}\mathrm{M}_{ij}$, then

$$det(\mathrm{A}) = \sum_{j=1}^{n} a_{ij}\mathrm{A}_{ij} \quad \text{or} \quad det(\mathrm{A}) = \sum_{i=1}^{n} a_{ij}\mathrm{A}_{ij}$$

Example 14

$$\mathrm{A} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\begin{aligned} det(\mathrm{A}) &= 5 \times (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 6 \times (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} \\ &= -5 \times (18 - 24) - 6 \times (8 - 14) = 30 + 36 = 66 \end{aligned}$$

Determinant

Example 15

$$\mathrm{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$det(\mathrm{A}) = a_{11}a_{22}\cdots a_{nn}$$

$$det(\mathrm{A}) = a_{11} \cdot (-1)^{1+1} \begin{vmatrix} a_{22} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= a_{11}a_{22} \cdot (-1)^{1+1} \begin{vmatrix} a_{33} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ a_{n3} & \cdots & a_{nn} \end{vmatrix} = \cdots = a_{11}a_{22}\cdots a_{nn}$$

Determinant

Basic Properties of Determinants

- $det[\cdots, \mathrm{A}_i + \mathrm{A}_i^{'}, \cdots] = det[\cdots, \mathrm{A}_i, \cdots] + det[\cdots, \mathrm{A}_i^{'}, \cdots]$
- $det[\cdots, cA_i, \cdots] = c \cdot det[\cdots, \mathrm{A}_i, \cdots]$
- $det[\cdots, \mathrm{A}_i, \cdots, \mathrm{A}_j, \cdots] = -det[\cdots, \mathrm{A}_j, \cdots, \mathrm{A}_i, \cdots]$
- $det(\mathrm{I}_n) = 1$
- $det[\mathrm{A}_1, \cdots, \mathrm{B}, \cdots, \mathrm{B}, \cdots] = 0$
- $det[\cdots, \mathrm{A}_i + c\mathrm{A}_j, \cdots] = det[\mathrm{A}_1, \cdots, \mathrm{A}_n]$
- $det[\cdots, 0, \cdots] = 0$

Determinant

Determinant and elementary operations

If a square matrix A is transformed to another matrix $\bar{\mathrm{A}}$ via an elementary operation e, then

$$det(\bar{\mathrm{A}}) = q \cdot det(\mathrm{A})$$

the number q is

- $q = -1$ if e is a row switching.
- $q = \lambda$ if e is a row multiplication by a number λ.
- $q = 1$ if e is a row replacement.

Determinant

Some Additional Properties

- $det(\mathrm{A}) = det(\mathrm{A}')$
- $det(\alpha\mathrm{A}) = \alpha^n det(\mathrm{A})$
- $det(\mathrm{AB}) = det(\mathrm{A}) \cdot det(\mathrm{B})$
- $det(\mathrm{A}^{-1}) = (det(\mathrm{A}))^{-1}$
- $$det \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{bmatrix} = det(\mathrm{D}) \cdot det(\mathrm{A} - \mathrm{BD}^{-1}\mathrm{C}) \quad \text{if} \quad det(\mathrm{D}) \neq 0$$
  $$= det(A) \cdot det(\mathrm{D} - \mathrm{CA}^{-1}\mathrm{B}) \quad \text{if} \quad det(\mathrm{A}) \neq 0$$
- $det(\mathrm{A}) \neq 0$ if and only if A is non-singular.
- If A is orthogonal, then $det(\mathrm{A}) = 1$ or -1.

Determinant

Proof.

4.$\mathrm{A}\mathrm{A}^{-1} = \mathrm{I} \rightarrow det(\mathrm{A})det(\mathrm{A}^{-1}) = 1 \rightarrow det(\mathrm{A}^{-1}) = (det(\mathrm{A}))^{-1}$

5.

$$\begin{bmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{bmatrix} \rightarrow \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ 0 & \mathrm{D} - \mathrm{C}\mathrm{A}^{-1}\mathrm{B} \end{bmatrix}$$

$$det \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{bmatrix} = det(A) \cdot det(\mathrm{D} - \mathrm{C}\mathrm{A}^{-1}\mathrm{B})$$

7.$\mathrm{A}\mathrm{A}' = \mathrm{I} \rightarrow (det(\mathrm{A}))^2 = 1 \rightarrow det(\mathrm{A}) = 1 \text{or} -1$ □

Idempotent matrices and projection matrix

Definition 16

A square matrix A is idempotent if $\mathrm{AA} = \mathrm{A}$

Definition 17

Let X be an $n \times k$ matrix, $k < n$, two projection matrices are

$$\mathrm{P} = \mathrm{X}(\mathrm{X}^{'}\mathrm{X})^{-1}\mathrm{X}^{'}$$
$$\mathrm{M} = \mathrm{I} - \mathrm{X}(\mathrm{X}^{'}\mathrm{X})^{-1}\mathrm{X}^{'}$$

When we regress Y on X, we have $\hat{\beta} = (\mathrm{X}^{'}\mathrm{X})^{-1}\mathrm{X}^{'}\mathrm{Y}$, then

$$\hat{\mathrm{Y}} = \mathrm{X}\hat{\beta} = \mathrm{X}(\mathrm{X}^{'}\mathrm{X})^{-1}\mathrm{X}^{'}\mathrm{Y} = \mathrm{PY}$$
$$\hat{\mathrm{U}} = \mathrm{Y} - \hat{\mathrm{Y}} = \mathrm{Y} - \mathrm{PY} = \mathrm{MY}$$

Idempotent matrices and projection matrix

Some properties

$\mathrm{P} \cdot \mathrm{P} = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}' = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}' = \mathrm{P}$
$\mathrm{M} \cdot \mathrm{M} = (\mathrm{I} - \mathrm{P})(\mathrm{I} - \mathrm{P}) = \mathrm{I} - 2\mathrm{P} + \mathrm{PP} = \mathrm{I} - 2\mathrm{P} + \mathrm{P} = \mathrm{I} - \mathrm{P} =$
$\mathrm{M}\ \mathrm{M} + \mathrm{P} = \mathrm{I}$
$\mathrm{M} \cdot \mathrm{P} = (\mathrm{I} - \mathrm{P})\mathrm{P} = \mathrm{P} - \mathrm{PP} = \mathrm{P} - \mathrm{P} = 0$

Another useful property of P and M

$tr(\mathrm{P}) = tr(\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}') = tr((\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{X}) = tr(\mathrm{I}_k) = k$
$tr(\mathrm{M}) = tr(\mathrm{I}_n - \mathrm{P}) = tr(\mathrm{I}_n) - tr(\mathrm{P}) = n - k$

Positive Definite Matrices

Definition 18

We say a square matrix A is positive semi-definite if for all non-zero c, $c^{'}\mathrm{A}c \geq 0$. Sometimes, this is written as $\mathrm{A} \geq 0$. We say a square matrix A is positive definite if for all non-zero c, $c^{'}$ A $c > 0$. This is written as $\mathrm{A} > 0$.

Some properties

- If $\mathrm{A} > 0$, then A is non-singular, A^{-1} exists and $\mathrm{A}^{-1} > 0$.
- For an $n \times k$ matrix X which has full column rank, then $\mathrm{XX}^{'}$ is symmetric and positive definite.
- If A is symmetric, then $\mathrm{A} > 0 \iff$ all its characteristic roots are positive.
- If $\mathrm{A} > 0$, we can find B such that $\mathrm{A} = \mathrm{BB}^{'}$. We call B a matrix square root of A. Note that B may not be unique.

Kronecker product and vec operation

Definition 19

Let $\mathrm{A} = \left[a_{ij}\right]$ be an $m \times n$ matrix,

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix},$$

where a_j is the j-th column of A. The vec of A, denoted by $vec(\mathrm{A})$, is the $mn \times 1$ vector.

$$vec(\mathrm{A}) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Kronecker product and vec operation

Definition 20

Let $\mathrm{B}_{s\times t}$ be any matrix. The Kronecker product A and B, denoted by $\mathrm{A}\bigotimes\mathrm{B}$, is the matrix

$$\mathrm{A}\bigotimes\mathrm{B}=\begin{bmatrix} a_{11}\mathrm{B} & a_{12}\mathrm{B} & \cdots & a_{1n}\mathrm{B} \\ a_{21}\mathrm{B} & a_{22}\mathrm{B} & \cdots & a_{2n}\mathrm{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathrm{B} & a_{m2}\mathrm{B} & \cdots & a_{mn}\mathrm{B} \end{bmatrix}$$

Some properties

- $(\mathrm{A}+\mathrm{B})\bigotimes\mathrm{C}=\mathrm{A}\bigotimes\mathrm{C}+\mathrm{B}\bigotimes\mathrm{C}$
- $(\mathrm{A}\bigotimes\mathrm{B})\cdot(\mathrm{C}\bigotimes\mathrm{D})=(\mathrm{AC})\bigotimes(\mathrm{BD})$(Mixed product property)
- $\mathrm{A}\bigotimes(\mathrm{B}\bigotimes\mathrm{C})=(\mathrm{A}\bigotimes\mathrm{B})\bigotimes\mathrm{C}$
- $(\mathrm{A}\bigotimes\mathrm{B})^{'}=\mathrm{A}^{'}\bigotimes\mathrm{B}^{'}$

Kronecker product and vec operation

Some properties

- $tr(\mathrm{A} \bigotimes \mathrm{B}) = tr(\mathrm{A}) \cdot tr(\mathrm{B})$
- If A is $m \times m$, B is $n \times n$, we have
$$det(\mathrm{A} \bigotimes \mathrm{B}) = [det(\mathrm{A})]^n [det(\mathrm{B})]^m$$
- $(\mathrm{A} \bigotimes \mathrm{B})^{-1} = \mathrm{A}^{-1} \bigotimes \mathrm{B}^{-1}$
- If $\mathrm{A} > 0$, $\mathrm{B} > 0$, then $\mathrm{A} \bigotimes \mathrm{B} > 0$
- $$\begin{aligned} vec(\mathrm{ABC}) &= (\mathrm{C}' \bigotimes \mathrm{A}) vec(\mathrm{B}) \\ &= (\mathrm{I} \bigotimes \mathrm{AB}) vec(\mathrm{C}) \\ &= (\mathrm{C}'\mathrm{B}' \bigotimes \mathrm{I}) vec(\mathrm{A}) \end{aligned}$$
$$vec(\mathrm{AB}) = (\mathrm{I} \bigotimes \mathrm{A}) vec(\mathrm{B}) = (\mathrm{B}' \bigotimes \mathrm{I}) vec(\mathrm{A})$$
- $tr(\mathrm{ABCD}) = vec(\mathrm{D}')'(\mathrm{C}' \bigotimes \mathrm{A}) vec(\mathrm{B})$

Kronecker product and vec operation

Example 21

The Kronecker product can be used to present linear equations in which the unknowns are matrices

$$\mathrm{AX} = \mathrm{B},$$

where $\mathrm{X} = \begin{bmatrix} \mathrm{X}_1 & \cdots & \mathrm{X}_k \end{bmatrix}$, $\mathrm{B} = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix}$. The equations above can be reformulated as

$$\begin{gathered} \mathrm{A} \begin{bmatrix} \mathrm{X}_1 & \cdots & \mathrm{X}_k \end{bmatrix} = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix} \\ \downarrow \\ \mathrm{AX}_1 = b_1 \quad \cdots \quad \mathrm{AX}_k = b_k \\ \downarrow \end{gathered}$$

Kronecker product and vec operation

Example 1.38

$$\begin{bmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

$$\downarrow$$

$$(I \bigotimes A) \cdot vec(X) = vec(B)$$

$$\downarrow$$

$$vec(AX) = vec(B)$$

Kronecker product and vec operation

Example 22

$$\mathrm{AX} + \mathrm{XB} = \mathrm{C}$$
$$\downarrow$$
$$vec(\mathrm{AX} + \mathrm{XB}) = vec(\mathrm{C})$$
$$\downarrow$$
$$vec(\mathrm{AX}) + vec(\mathrm{XB}) = vec(\mathrm{C})$$
$$\downarrow$$
$$(\mathrm{I} \bigotimes \mathrm{A}) vec(\mathrm{X}) + (\mathrm{B}' \bigotimes \mathrm{I}) vec(\mathrm{X}) = vec(\mathrm{C})$$
$$\downarrow$$
$$[(\mathrm{I} \bigotimes \mathrm{A}) + (\mathrm{B}' \bigotimes \mathrm{I})] vec(\mathrm{X}) = vec(\mathrm{C})$$

Kronecker product and vec operation

Example 23

$$\mathrm{AXB} = \mathrm{C}$$
$$\downarrow$$
$$vec(\mathrm{AXB}) = vec(\mathrm{C})$$
$$\downarrow$$
$$(\mathrm{B}^{'} \bigotimes \mathrm{A})vec(\mathrm{X}) = vec(\mathrm{C})$$
$$\downarrow$$

If all matrices are invertible, then

$$vec(\mathrm{X}) = [\mathrm{B}^{'} \bigotimes \mathrm{A}]^{-1} vec(\mathrm{C})$$
$$\downarrow$$
$$vec(\mathrm{X}) = [(\mathrm{B}^{'})^{-1} \bigotimes \mathrm{A}^{-1}] vec(\mathrm{C})$$

Matrix Calculus

Definition 24

Let $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}$ be $k \times 1$ matrix and $g(X) = g(X_1, X_2, \cdots, X_k)$: $\mathbb{R}^k \to \mathbb{R}$. The vector derivative is

$$\frac{\partial g(X)}{\partial X} = \begin{bmatrix} \frac{\partial g(X)}{\partial X_1} \\ \frac{\partial g(X)}{\partial X_2} \\ \vdots \\ \frac{\partial g(X)}{\partial X_k} \end{bmatrix}$$

$$\frac{\partial g(X)}{\partial X'} = \begin{bmatrix} \frac{\partial g(X)}{\partial X_1} & \frac{\partial g(X)}{\partial X_2} & \cdots & \frac{\partial g(X)}{\partial X_k} \end{bmatrix}$$

Matrix Calculus

Some properties

- $\frac{\partial}{\partial \mathrm{X}}(a^{'}\mathrm{X}) = \frac{\partial}{\partial \mathrm{X}}(\mathrm{X}^{'}a) = a$
- $\frac{\partial}{\partial \mathrm{X}^{'}}(\mathrm{AX}) = \mathrm{A}$
- $\frac{\partial}{\partial \mathrm{X}}(\mathrm{X}^{'}\mathrm{AX}) = (\mathrm{A} + \mathrm{A}^{'})\mathrm{X}$
- $\frac{\partial^2}{\partial \mathrm{X} \partial \mathrm{X}^{'}}(\mathrm{X}^{'}\mathrm{AX}) = \mathrm{A} + \mathrm{A}^{'}$

Some properties

Let $\mathrm{A} = \left[a_{ij}\right]$ be an $m \times n$ matrix and $g(\mathrm{A})$: $\mathbb{R}^{m \times n} \to \mathbb{R}$. We define

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

- $\frac{\partial}{\partial \mathrm{A}}(\mathrm{X}^{'}\mathrm{AX}) = \mathrm{XX}^{'}$

Matrix Calculus

Some properties

Let $\mathrm{A}_{m\times m}$ be a non-singular matrix whose elements are functions of the scale parameter α

$$\mathrm{A} = \begin{bmatrix} a_{11}(\alpha) & a_{12}(\alpha) & \cdots & a_{1m}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}(\alpha) & a_{m2}(\alpha) & \cdots & a_{mm}(\alpha) \end{bmatrix}$$

then $\frac{\partial \mathrm{A}^{-1}}{\partial \alpha} = -\mathrm{A}^{-1}\frac{\partial \mathrm{A}}{\partial \alpha}\mathrm{A}^{-1}$.

Proof.

$$\begin{gathered} \mathrm{A}\mathrm{A}^{-1} = \mathrm{I} \\ \mathrm{A}\frac{\partial \mathrm{A}^{-1}}{\partial \alpha} + \frac{\partial \mathrm{A}}{\partial \alpha}\mathrm{A}^{-1} = 0 \\ \frac{\partial \mathrm{A}^{-1}}{\partial \alpha} = -\mathrm{A}^{-1}\frac{\partial \mathrm{A}}{\partial \alpha}\mathrm{A}^{-1} \end{gathered}$$

□