STAT5030 Linear Models

Yuanyuan Lin[a] (Instructor)[1]

January 16, 2023

[1]Correspondence to: Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T, Hong Kong, China (E-mail: ylin@sta.cuhk.edu.hk)

"A journey of a thousand miles begins with a single step." – *Lao Tzu (Laozi).*

Chapter 1. Generalized inverse

We review some advanced linear algebra concepts and theorems.

-Full Rank Factorization

full row rank.

THEOREM 1. *$A_{p\times q}$ of rank r can always be factorized as $A = K_{p\times r}L_{r\times q}$, where K and L have full column and full row rank respectively.*

full columnrank

Proof. There exist nonsingular matrices P and Q such that

P & Q are primary operation on rows & columns, respectively.

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

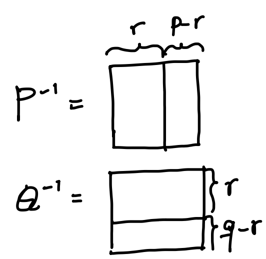
$$\Longrightarrow A = P^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q^{-1}.$$

Partition P^{-1} and Q^{-1} as

Simply partition.

$$P^{-1} = [K_{p\times r} W_{p\times(p-r)}]$$

$$Q^{-1} = \begin{bmatrix} L_{r\times q} \\ Z_{(q-r)\times q} \end{bmatrix}.$$



$$\begin{aligned} A =& [K \quad W]\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} L \\ Z \end{bmatrix} \\ =& [K \quad 0]\begin{bmatrix} L \\ Z \end{bmatrix} \\ =& KL. \end{aligned}$$

□

Remark 1. For any matrix $A_{p\times q}$ of rank r, $r \le \min(p, q)$.

Remark 2. A matrix $A \in \mathbb{R}^{m\times n}$ is

(i) full column rank iff $A^\top A$ is invertible; rank $(A^TA) = n$

(ii) full row rank iff $AA^\top$ is invertible. rank $(AA^T) = m$

Remark 3. If matrix A has dimension $n \times m$ and is full rank, then we use the *left inverse* if $n > m$ and the *right inverse* if $n < m$:

$r = \min\{n, m\}$.

make sure that A^TA is invertable

$(A^TA)^{-1}A^T$

$n<m$ $A^T(AA^T)^{-1}$

put the inverse on the left side of A.

(i) Left inverse: $A^{-1}_{left} = (A^\top A)^{-1} A^\top$, that is, $A^{-1}_{left} A = I_m$, where I_m is an $m \times m$ identity matrix.

(ii) Right inverse: $A^{-1}_{right} = A^\top (AA^\top)^{-1}$, that is, $AA^{-1}_{right} = I_n$, where I_n is an $n \times n$ identity matrix.

$y = X\beta + \varepsilon$ $\quad X \in \mathbb{R}^{n\times p}$ $\quad \beta \in \mathbb{R}^p$ $\quad \mathbb{E}\varepsilon_i = 0$ $\quad \mathrm{Var}\,\varepsilon_i = \sigma^2 < \infty$

$\hat\beta = (X^\top X)^{-1} X^\top y$ $\quad \hat y = X\hat\beta = \underbrace{\boxed{X(X^\top X)^{-1}X^\top}}_{H.}\, y.$

full rank

nonfull rank.

-Idempotent Matrices ($A^2 = A$) $A \in \mathbb{R}^{n \times n}$

Of course, I is an idempotent matrix.

THEOREM 2. *All idempotent matrices(except I) are singular.*

Proof. Since $A^2 = A$, if A is nonsingular,

Is there any other idempotent matrix?

$$A = A^{-1}AA = A^{-1}A = I.$$

□

THEOREM 3. *For any idempotent matrix A, $r(A) = tr(A)$.* $tr(A) = \Sigma_i A_{ii}$

Proof. Consider the full rank factorization, let

$$A = BC \text{ and } A^2 = BCBC = BC.$$

But B has left inverse L and C has a right inverse R, then

$A = BC$ full rank.

then $CB = I_r$

$L = (B^T B)^{-1} B^T$

$R = C^T (CC^T)^{-1}$

$$\overset{I}{\underline{LB}}CB\overset{I}{\underline{CR}} = \overset{I}{\underline{LB}}\,\overset{I}{\underline{CR}}$$
$$\Longrightarrow CB = I_{r \times r}.$$

Thus,

$$\begin{aligned} tr(A) &= tr(BC) \leftarrow \text{full rank factorization} \\ &= tr(CB) \leftarrow \text{exchangeable.} \\ &= tr(I_{r \times r}) \\ &= r \\ &= r(A). \end{aligned}$$

□

THEOREM 4. *Eigenvalues of idempotent matrices are either 0 or 1.*

Proof. Let λ, x be a pair of eigenvalue and eigenvector. Then,

$$Ax = \lambda x$$
$$\Longrightarrow A^2x = A(Ax) = A(\lambda x) = \lambda Ax = \lambda^2 x.$$

However,

$A^2x = A\lambda x = \lambda^2 x = \lambda x = Ax$

$\Rightarrow (\lambda^2 - \lambda)x = 0$

$\lambda = 0$ or 1.

$$A^2x = Ax = \lambda x$$
$$\lambda^2 x = \lambda x \Rightarrow \lambda(\lambda - 1)x = 0.$$

In view of the definition of eigenvector,

$$x \neq 0 \Rightarrow \lambda = 0 \text{ or } 1.$$

Thm 4 A idempotent $\Rightarrow$ $\lambda = \{0, 1\}$

Thm 5 $\lambda = \{0, 1\}$ & A symmetric $\Rightarrow$ A idempotent.

$\square$

THEOREM 5. *For a symmetric matrix A, if all its eigenvalues are 1 or 0, then A is idempotent.*

Proof. Since A is symmetric, there exists an orthogonal matrix P such that

$$P^\top AP = D$$ eigenvalue decomposition

where D is a diagonal matrix with eigenvalues of A on the diagonal.

Hence, $P^\top APP^\top AP = D^2$.

But, $P^\top APP^\top AP = P^\top AAP$.

Nevertheless, if all eigenvalues are 1 or 0,

$\Rightarrow D = D^2$

$\Rightarrow P^\top AP = P^\top AAP$

$\Rightarrow A = AA$

$\Rightarrow A$ is idempotent. $\square$

$\lambda \in \{1, 0\} \Rightarrow A^2 = A \Rightarrow$ A is idempotent.

$\Rightarrow D^2 = D$

$\Rightarrow P^\top AAP = P^\top AP$

row & column vectors are all norm-1.

Remark 4. Orthogonal matrix is a square matrix with real entries whose columns and rows are orthogonal unit vectors (i.e orthonormal vectors), that is

orthogonal & unit $\Leftrightarrow$ orthonormal

$$Q^\top Q = QQ^\top = I,$$

implying $\boxed{Q^\top = Q^{-1}.}$

潘罗斯.

-Moore- Penrose Inverse

Definition: Let A be an $m \times n$ matrix.

If a matrix A^+ exists that satisfies

It may not exists for some special A? No any matrix A has Penrose-Inverse A^+ by Thm6.

$$\left.\begin{array}{ll} (1) & AA^+ \text{ is symmetric.} \\ (2) & A^+A \text{ is symmetric.} \\ (3) & AA^+A = A. \\ (4) & A^+AA^+ = A^+. \end{array}\right\} (*)$$

A^+ is defined as a Moore-Penrose inverse of A.

THEOREM 6. *Each matrix(A) has an A^+.*

Proof. If $A = 0$, $A^+ = 0$.

If $A \neq 0$, A can be factored by the full-rank factorization

$$A = A_L A_R = BC,$$

where B is $m \times r$ of rank r and C is $r \times n$ of rank r.

Hence, $B^\top B$ and $CC^\top$ are both nonsingular.

Define

$$A^+ = C^\top (CC^\top)^{-1} (B^\top B)^{-1} B^\top,$$

and it can be shown that A^+ satisfies the 4 conditions in $(*)$. □

$A = BC$

$\Rightarrow A^+ = (BC)^+ \boxed{= C^+ B^+}$ just for memorizing.

$= \underbrace{C^\top (CC^\top)^{-1}}_{\text{right.}} \underbrace{(B^\top B)^{-1} B^\top}_{\text{left}}$

Properties:

1. The Moore-Penrose inverse is unique.

2. $(A^\top)^+ = (A^+)^\top$.

3. $r(A^+) = r(A)$.

4. If $A = A^\top$, then $A^+ = (A^+)^\top$.

5. If A is nonsingular, $A^{-1} = A^+$.

6. If A is symmetric idempotent, $A^+ = A$.

7. If $r(A_{m\times n}) = m$, then $A^+ = A^\top (AA^\top)^{-1}, AA^+ = I$,
   If $r(A_{m\times n}) = n$, then $A^+ = (A^\top A)^{-1} A^\top, A^+A = I$.

8. The matrices AA^+, A^+A, $I - AA^+$ and $I - A^+A$ are all symmetric idempotent.

For the computation of Moore-Penrose inverse, see for example"Matrixes with application in Statistics" By Graybill p.118-129.

-Generalized Inverse

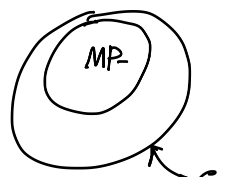
Definition: Let A be an $m \times n$ matrix and the generalized inverse A^- satisfies

$$AA^-A = A.$$

$A^-AA^- = A^-$? not care!

Properties:

A^+ is also a generalized inverse



G-Inverse is not unique.

1. The M-P inverse is also a generalized inverse.

2. G-inverse may not be unique.

3. Let X be $m \times n$, $r(X) = k > 0$.

   (a) $r(X^-) \geq k$.

   ✓ (b) X^-X and XX^- are idempotent.

   (c) $r(X^-X) = r(XX^-) = k$.

   (d) $X^-X = I$ if and only if $r(X) = n$.

   (e) $XX^- = I$ if and only if $r(X) = m$.

   (f) $tr(X^-X) = tr(XX^-) = k = r(X)$

   (g) If X^- is any G-inverse of X, then $(X^-)^\top$ is a G-inverse of $X^\top$

generalized inverse if X is not full rank.

4. Let $K = X(X^\top X)^- X^\top$, K is invariant for any G-inverse of $X^\top X$.

5. $X(X^\top X)^- X^\top = XX^+$.

6. For $K = X(X^\top X)^- X^\top$

   (a) $K = K^\top$, $K = K^2$(Symmetric Idempotent).

   (b) rank(K)=rank(X)=r.(rank(K)=$tr(K)$=rank(X))

   (c) $KX = X$; $X^\top K = X^\top$.

   (d) $(X^\top X)^- X^\top$ is a G-inverse of X for any G-inverse of $X^\top X$.

   (e) $X(X^\top X)^-$ is a G-inverse of $X^\top$ for any G-inverse of $X^\top X$.

In linear model

$y = X\beta + \varepsilon$

$\hat{y} = \underbrace{X(X^\top X)^{-1}X^\top}_{K} y$