Lecture Notes for Theory of Linear Models

(Ch 2 in Rencher et al.)

- Review of Matrix Algebra
- Generalized Inverse

Longhai Li

Department of Mathematics and Statistics
University of Saskatchewan

Review of Matrix Theory.

Eigenvalues & Eigen Vector

$$\underset{n\times n}{A}x = \lambda x \;, \quad Ax - \lambda I_n \cdot x = 0$$

λ — eigen value

x — eigen vector.

$$|\lambda I_n - A| = 0, \text{ a polynomial of } \lambda.$$

In R, SVD(A) will give λ & x.

Singular value decomposition

Spectral Decomposition for Symmetric Matrices

A is a symmetric matrix : $n \times n$
(all eigen values are real)

$$A = \underbrace{(q_1, q_2, \cdots, q_n)}_{Q} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \lambda_n \end{bmatrix} \underbrace{\begin{bmatrix} q_1' \\ q_2' \\ \vdots \\ q_n' \end{bmatrix}}_{Q'}$$

$$= (q_1, \cdots, q_n) \begin{bmatrix} \lambda_1 q_1' \\ \vdots \\ \lambda_n q_n' \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \, q_i q_i' = \underline{Q \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q'}$$

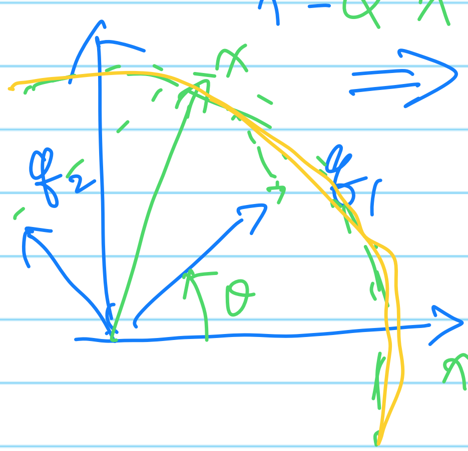
$q_i \perp q_j \quad (i \neq j), \quad \|q_i\|^2 = 1$

$Q = (q_1, \cdots, q_n)$ is an orthogonal matrix, $Q'Q = Q \cdot Q' = I_n$

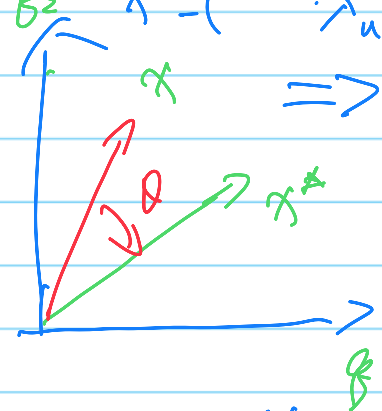
$$Ax = Q\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} Q' x$$

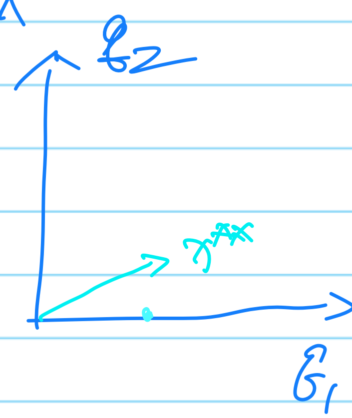
$$= Q\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} \cdot \begin{pmatrix} q_1' x \\ \vdots \\ q_n' x\end{pmatrix} = Q \cdot \begin{pmatrix}\lambda_1 q_1' x \\ \vdots \\ \lambda_n q_n' x\end{pmatrix}$$

$x^* = Q' x \quad \Longrightarrow \quad x^{**} = \begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} \cdot x^* \quad \Longrightarrow$



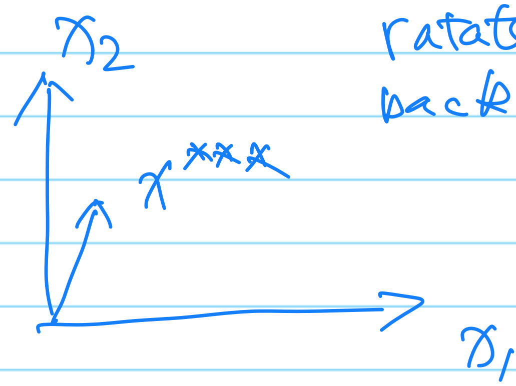
rotation





$x^{***} = (q_1, \cdots, q_n)\, x^{**}$

rotate back



Some facts about symm. matrices $\quad tr(AB) = tr(BA)$

1) $tr(A) = \sum_{i=1}^{n} \lambda_i$, $|A| = \prod_{i=1}^{n} \lambda_i$, $tr(A) = \sum_{i=1}^{n} a_{ii}$

$tr(Q\Lambda Q') = tr(\Lambda Q'Q) = tr(\Lambda)$, $|Q\Lambda Q'| = |\Lambda|$

2) A is singular if $\exists\, \lambda_i = 0$

3) $A^{-1} = Q \cdot \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix} Q'$, if $\lambda_i \neq 0$ for all i

4) $A^{\frac{1}{2}} \equiv Q \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot Q'$, if $\lambda_i \geq 0$ for all i

$(Q\Lambda Q')^{-1} = Q \cdot \Lambda^{-1} Q'$

square root $\quad A^{\frac{1}{2}} \cdot A^{\frac{1}{2}} = A$, $\quad A^{\frac{1}{2}}$ is symmetric

5)

Quadric form of symmetric +

$A: n\times n$, $\quad x: n\times 1$

$$x'Ax = \sum_{i,j} x_i a_{ij} x_j \qquad (x_1, \cdots, x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x' \cdot Q \cdot \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} Q'x, \quad \text{let } y = Q'x$$

$$= \sum_{i=1}^{n} \lambda_i (q_i' x)^2, \quad \text{where } Q = (q_1, \cdots, q_n)$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2 \qquad y_i = q_i' x$$

Projection Matrix

(1) $P = P'$

(2) $P^2 = P \Rightarrow \lambda_i = 0$ or 1

$$P = Q \cdot \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \cdot Q' = \sum_{i=1}^{r} q_i q_i'$$

some dimensions are re-scaled to 0 if $\lambda_i = 0$
or unchange if $\lambda_i = 1$
why $\lambda_i = 0$ or 1. ?

$x \in C(P)$, $Px = x = 1 \cdot x$

$x \perp C(P)$, $Px = 0 \cdot x$

Another proof:

$$Px = \lambda x, \quad P^2 = P$$

$$\Rightarrow \left.\begin{array}{l} P^2 x = \lambda P x = \lambda^2 x \\ P^2 x = Px = \lambda x \end{array}\right\} \Rightarrow \lambda^2 x = \lambda x, \ |x| \neq 0$$

$$\Rightarrow \lambda^2 = \lambda \Rightarrow \lambda = 0/1$$

Example

$$P = \frac{1}{n} \hat{J}_n \cdot \hat{J}_n'$$

$$= \frac{1}{n} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [1, 1, \ldots, 1]$$

$$= \frac{1}{n} \begin{bmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{bmatrix}$$

$$\text{rank}(P) = tr(P)$$

$$= tr\left(\frac{1}{n} \hat{J}_n \cdot \hat{J}_n'\right)$$

$$= \frac{1}{n} \cdot tr(\hat{J}_n' \hat{J}_n)$$

$$= \frac{1}{n} \cdot tr([n]) = \frac{1}{n} \cdot n$$

$$= 1$$

Positive Definite (p.d.) and Positive semi-definite (p.s.d) Matrices

A is symmetric

A is p.d. iff $x' A x > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0$ $[A > 0]$

other notation $(A > 0)$

A is p.s.d. iff $x' A x \geq 0 \quad \forall x \in \mathbb{R}^n$ $[A \geq 0]$

other notation $(A \geq 0)$

Examples:

1) The matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$$

is positive definite.

Q: Why?

A: Because the associated quadratic form is

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= 2x_1^2 - 2x_1x_2 + 3x_2^2 = 2(x_1 - \frac{1}{2}x_2)^2 + \frac{5}{2}x_2^2,$$

2)

The matrix

$$\mathbf{B} = \begin{pmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}$$

is positive semidefinite because its associated quadratic form is

$$\mathbf{x}^T \mathbf{B} \mathbf{x} = (2x_1 - x_2)^2 + (3x_1 - x_3)^2 + (3x_2 - 2x_3)^2,$$

which is always non-negative, but does equal 0 for $\mathbf{x} = (1, 2, 3)^T$ (or any multiple of $(1, 2, 3)^T$).

Some facts about p.d. & p.s.d.

Let $A = (a_{ij})_{n \times n}$

1) A is p.d. $\Rightarrow a_{ii} > 0$

A is p.s.d. $\Rightarrow a_{ii} \geq 0$

Pf: let $x = (0, \cdots 0, \underset{\uparrow \atop ith}{1}, 0, \cdots, 0)'$, $x'Ax = \underline{a_{ii}}$

2)

A is p.d. $\Leftrightarrow$ $\underline{\lambda_i > 0}$ (SVD(A))

A is p.s.d. $\Leftrightarrow$ $\lambda_i \geq 0$

pf: $A = Q \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q'$, where $Q = (q_1, \cdots, q_n)$

$$\underline{x'Ax} = \underline{\sum_{i=1}^{n} \lambda_i (q_i' x)^2} = \underline{\sum_{i=1}^{n} \underline{\underline{\lambda_i}} \cdot y_i^2}$$

where $y = Q'x$,

$x'Ax = y_1^2 - y_2^2$ isn't always ≥ 0

$X'X$

3) Let B : $n \times p$ matrix

a) if $\text{rank}(B) = p$, then $\underset{p\times p}{B'B}$ is p.d.

b) if $\text{rank}(B) < p$, then $B'B$ is p.s.d.

Pf: $\text{rank}(B) = p$, then $Bx \neq 0 \ \forall x \neq 0$

so $x'B'Bx = \|Bx\|^2 > 0$

If $\text{rank}(B) < p$, then Bx may be 0

for some $x \neq 0$. But we always have

$$x'Ax = \|Bx\|^2 \geq 0.$$

4) A is p.d. $\Rightarrow$ A^{-1} exists. (non-singular)

5) A is p.d. $\Rightarrow$ A^{-1} is p.d.

$A = Q\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_p\end{pmatrix}Q'$, $A^{-1} = Q\begin{pmatrix}\lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_p^{-1}\end{pmatrix}Q'$

Cholesky Decomposition

If A is p.d. . $\exists$ B s.t. $A = B'B$

where B is an upper triangle matrix

The factorization is unique.

$$\underbrace{\begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{12} & b_{22} & \cdots & 0 \\ \vdots & & & \vdots \\ b_{1n} & b_{2n} & \cdots & b_{nn} \end{pmatrix}}_{B'} \cdot \underbrace{\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ & \cdots & \cdots & \\ 0 & \cdots & \cdots & b_{nn} \end{pmatrix}}_{B}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = A$$

$b_{11}^2 = a_{11} \Rightarrow b_{11} = \sqrt{a_{11}}$ $\quad a_{ii} > 0$

$b_{11} \cdot b_{1j} = a_{1j} \Rightarrow b_{1j} = \dfrac{a_{1j}}{b_{11}}$

$b_{12}^2 + b_{22}^2 = a_{22} \Rightarrow b_{22} = \sqrt{a_{22} - b_{12}^2}$

Why Cholesky? B^{-1} can be obtained easily.

$X'X$

Singular value decomposition

X : $n \times p$, suppose $\mathrm{rank}(X) = r \le \min(n,p)$

Then X can be written as:

$$X = \underset{n\times r}{(u_1, \cdots, u_r)} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_r \end{pmatrix} \underset{r\times p}{\begin{pmatrix} v_1' \\ \vdots \\ v_r' \end{pmatrix}}$$

$$= \underset{n\times n}{U} \cdot \begin{pmatrix} \underset{r\times r}{\Lambda} & \underset{r\times(p-r)}{O_{12}} \\ \underset{(n-r)\times r}{O_{21}} & \underset{(n-r)\times(p-r)}{O_{22}} \end{pmatrix} \underset{p\times p}{V'},$$

where, O_{12}, O_{21}, O_{22} are 0 matrix.

$u_i \perp u_j$ for $i \ne j$, $\|u_i\| = 1$, $i = 1, \cdots, r$

$v_i \perp v_j$ for $i \ne j$, $\|v_i\| = 1$, $i = 1, \cdots, r$

$U'U = U \cdot U' = I_n$

$V'V = V \cdot V' = I_p$

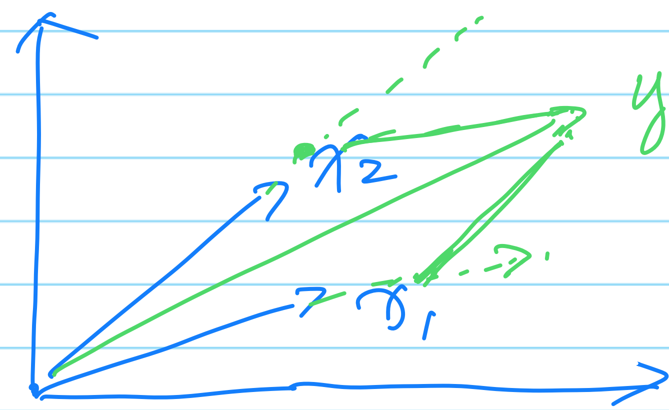
$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_r \end{pmatrix} = \mathrm{diag}(\lambda_1, \cdots, \lambda_r)$$

$U = (u_1, \cdots, u_r, u_{r+1}, \cdots, u_n)$
$= (U_1, U_2)$

$V = (v_1, \cdots, v_r, v_{r+1}, \cdots, v_p)$
$= (V_1, V_2)$

Generalized Inverses

Motivation



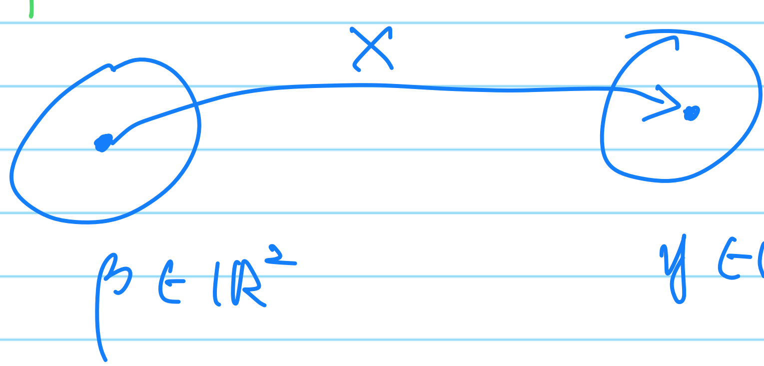
$\mathbb{R}^2$

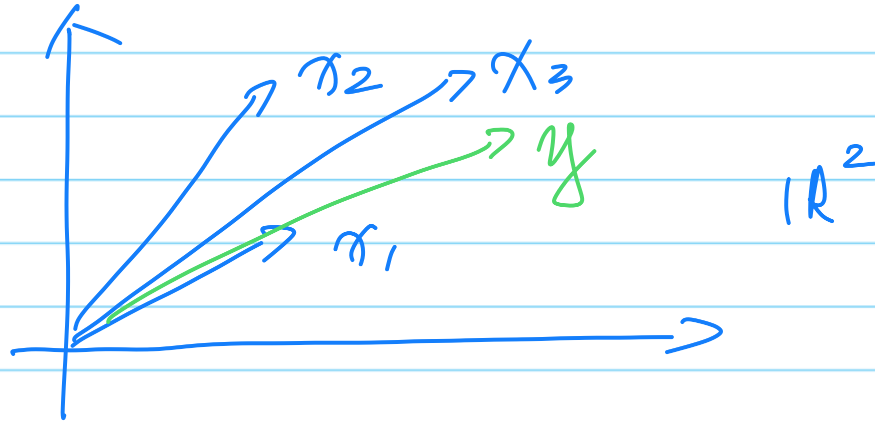
$X = [\,x_1,\; x_2\,]$ invertible

$X\beta = y \quad \forall\, y \in \mathbb{R}^2$

$[\,x_1,\; x_2\,] \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$.

$\beta = X^{-1} y$, $\quad X \cdot (X^{-1} y) = y$, for all $y \in \mathbb{R}^2$

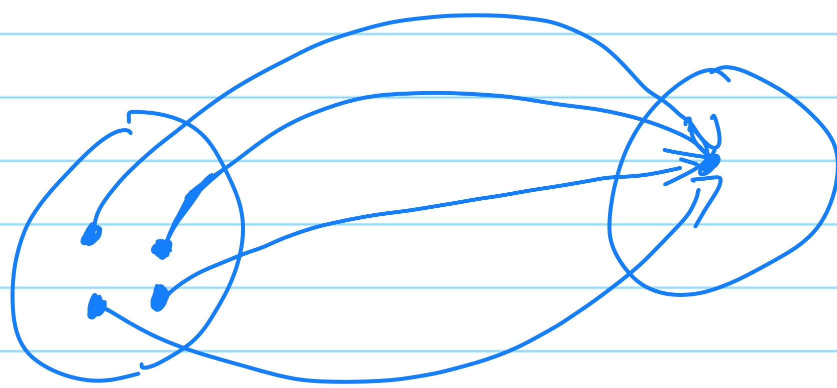




$x_i \in \mathbb{R}^2$

$X = [x_1, x_2, x_3]$

$X\beta = y$ doesn't have a unique sol.



$\beta \in \mathbb{R}^3$ $\quad$ $y \in \mathbb{R}^2$

$\beta = X^- y$ should be a solution to

$X\beta = y$, for $y \in C(X)$

What X^- should be ?

suppose $X = [x_1, \dots, x_p]$

$X \cdot (X^- x_j) = x_j$ for each x_j (*)

$\beta_j = X^- x_j$ should be a solution to $X\beta = x_j$

then $XX^- y = y$, for each $y \in C(X)$

Writing (*) in matrix form:

$X \cdot X^- \cdot [x_1, \dots, x_p] = [x_1, \dots, x_p]$

$X \cdot X^- X = X$

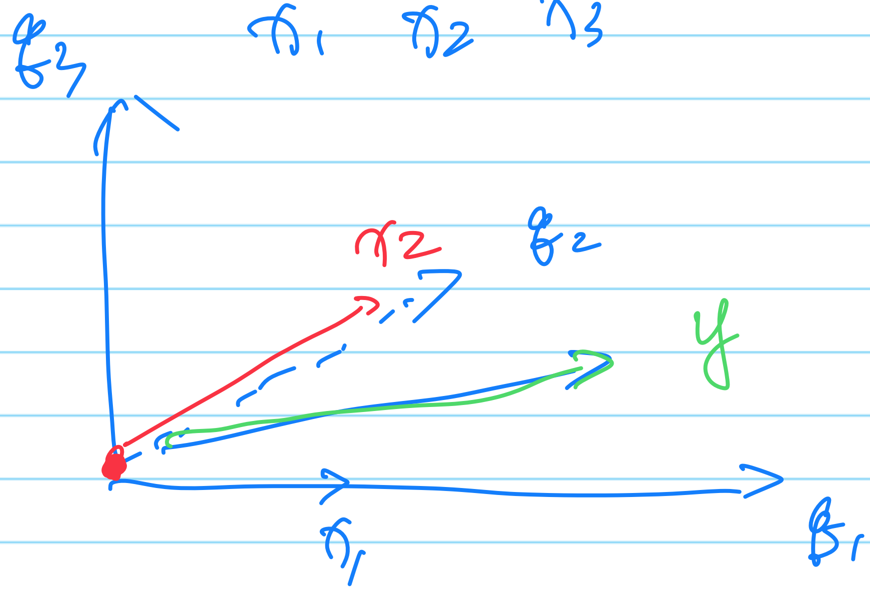
Generalized Inverse (Def):

Let X be an $n \times p$ matrix. X^- is a matrix of $p \times n$ and satisfies $X \cdot X^- X = X$. X^- is called a generalized inverse of X

a version of

Example:

$$X = \begin{array}{ccc} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ x_1 & x_2 & x_3 \end{array} \end{array} \qquad X^{-1} \text{ Not exist}$$



$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = y, \qquad y \in L(\beta_1, \beta_2)$$

$$X^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\underline{X \cdot X^- \cdot X} = \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \boxed{\begin{matrix} 0 \\ 0 \end{matrix}} \\ \boxed{\begin{matrix} 0 & 0 \end{matrix}} & \boxed{0} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= X$$

Example 1:

$$X = (\overset{x_1}{1}, \overset{x_2}{2}, \overset{x_3}{3}),$$

$$X^- = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad X \cdot X^- X = 1 \cdot X = X$$

$$X^- = \begin{pmatrix} 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \quad X \cdot X^- X = 1 \cdot X = X$$

Example

$$\mathbf{A} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 1 \\ 3 & 2 & 4 \end{pmatrix} \end{array}$$

rank$(A) = 2$, since

$x_3 = x_1 + \frac{1}{2} x_2$

$$\mathbf{A}_1^- = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{A}_2^- = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

A version of X^-.

$$X = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}_{n \times p} \quad R_{11}^{-1} \text{ exists.} \quad \dim(R_{11}) = r$$

$$X^- = \begin{pmatrix} R_{11}^{-1} & O_{21} \\ O_{12} & O_{22} \end{pmatrix}_{p \times n}$$

$\text{shape}(O_{12}) = \text{shape}(R_{12}')$
$\text{shape}(O_{21}) = \text{shape}(R_{21}')$
$\text{shape}(O_{22}) = \text{shape}(R_{22}')$

O_{12}, O_{21}, O_{22} are all 0 matrice.

$$X \cdot X^- X = \begin{pmatrix} I_r & 0 \\ R_{21}R_{11}^{-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

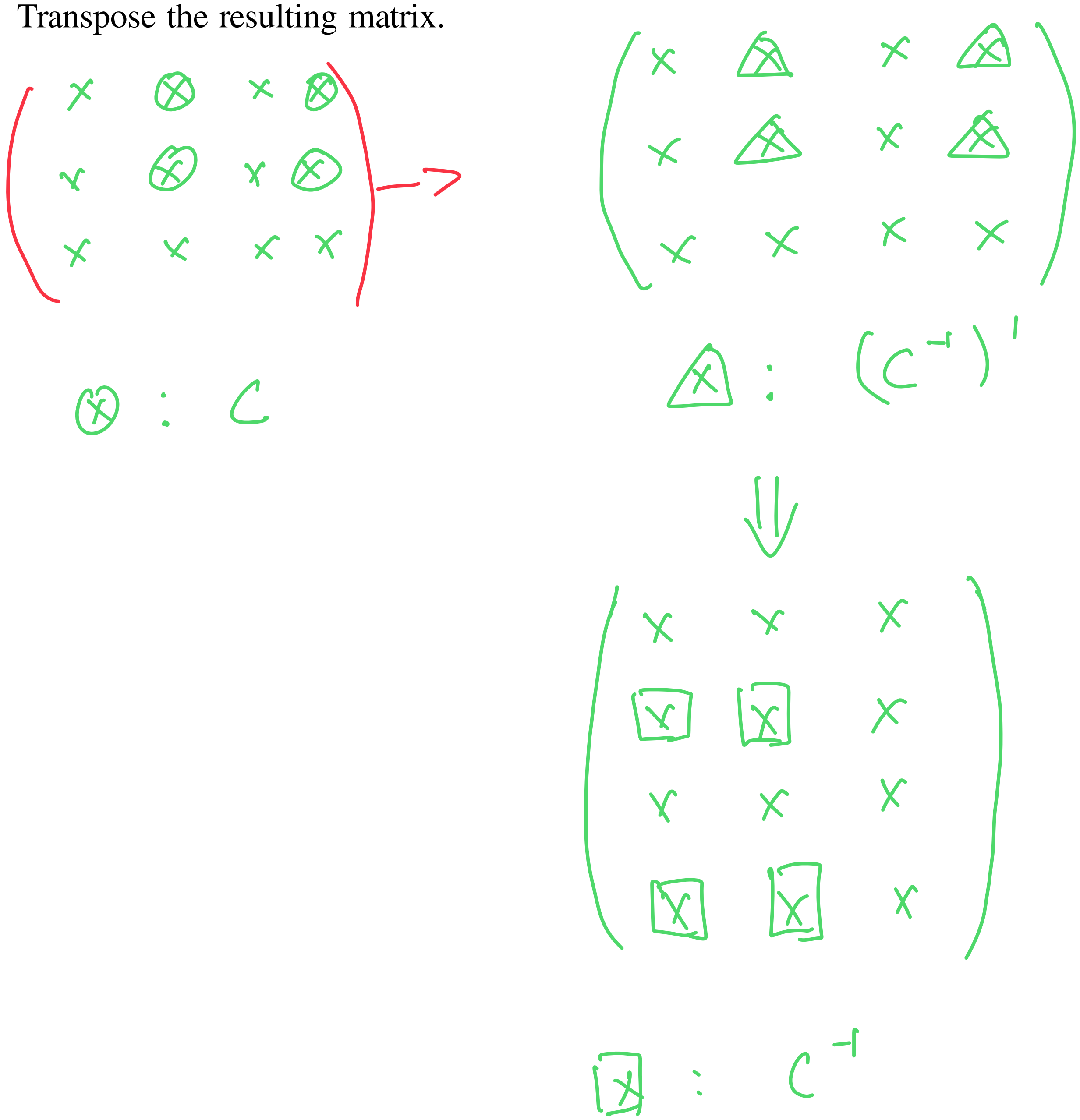
$$= \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = X$$

Note: X being singular implies that

$$R_{22} - R_{21}R_{11}^{-1}R_{12} = 0$$

A Procedure to Find A Version of Generalized Inverse

1. Find any nonsingular $r \times r$ submatrix $\mathbf{C}$. It is not necessary that the elements of $\mathbf{C}$ occupy adjacent rows and columns in $\mathbf{A}$.
2. Find $\mathbf{C}^{-1}$ and $(\mathbf{C}^{-1})'$.
3. Replace the elements of $\mathbf{C}$ by the elements of $(\mathbf{C}^{-1})'$.
4. Replace all other elements in $\mathbf{A}$ by zeros.
5. Transpose the resulting matrix.



Moore - Penrose Inverse

$$X = \underset{n\times n}{U} \cdot \underset{n\times p}{\begin{pmatrix} \Lambda & O_{12} \\ O_{21} & O_{22} \end{pmatrix}} \underset{p\times p}{V'} \quad \text{(SVD)}$$

$$X^{+} = V \begin{pmatrix} \Lambda^{-1} & O_{21}' \\ O_{12}' & O_{22}' \end{pmatrix} U'$$

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix}, \quad \Lambda^{-1} \overset{p\times n}{=} \begin{pmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_r^{-1} \end{pmatrix}$$

O_{12}, O_{21}, O_{22} are all 0 matrix

Checking:

$$X \cdot X^{+} X = X$$

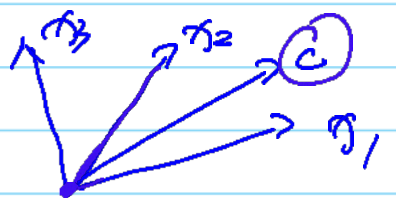
Theorem:

$\beta = X^- c$ is a solution to $X\beta = c$, if it is consistent.

Proof:

First we assume $c \in C(X)$, that is

$X\beta = c$ is consistent



$X = (x_1, x_2, x_3)$

$X\beta \in C(X)$

$\text{rank}([X, c]) = \text{rank}(X)$

$c \in L(x_1, x_2, x_3)$

Given c, suppose $Xb = c$, i.e., b is a solution.

Let X^- be a version of gen. inv. of X.

$$(XX^-)Xb = (X \cdot X^-) \cdot c$$

$$\Rightarrow \quad XX^-Xb = X \cdot (X^- c)$$

$$\Rightarrow \quad Xb = X \cdot (X^- c), \text{ since } XX^-X = X$$

$$\Rightarrow \quad c = X \cdot (X^- c), \text{ since } Xb = c$$

That is, $b_2 = X^- c$ is a solution of $X\beta = c$

may $\neq b$

Example 1:

$X = (1, 2, 3)$

To solve $X\beta = 4$

$$(1,2,3)\cdot\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = 4$$

(1) $X^- = \begin{pmatrix}1/1\\0\\0\end{pmatrix}$, $\beta = X^-\cdot 4 = \begin{pmatrix}1\\0\\0\end{pmatrix}\cdot 4 = \begin{pmatrix}4\\0\\0\end{pmatrix}$

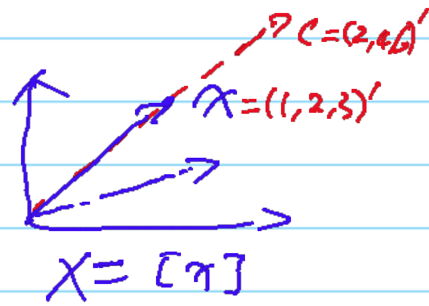
(2) $X^- = \begin{pmatrix}0\\ \frac{1}{2}\\0\end{pmatrix}$, $\beta = X^- 4 = \begin{pmatrix}0\\ \frac{1}{2}\\0\end{pmatrix}\cdot 4 = \begin{pmatrix}0\\2\\0\end{pmatrix}$

(3) $X^- = \begin{pmatrix}0\\0\\ \frac{1}{3}\end{pmatrix}$, $\beta = X^-\cdot 4 = \begin{pmatrix}0\\0\\ \frac{1}{3}\end{pmatrix}\cdot 4 = \begin{pmatrix}0\\0\\ \frac{4}{3}\end{pmatrix}$

Example 2:

$X = \begin{pmatrix}1\\2\\3\end{pmatrix}$

$X\beta = \begin{pmatrix}2\\4\\6\end{pmatrix} = c$



(1) $X^- = (1, 0, 0)$, $\quad \beta = X^- c = 2$

(2) $X^- = (0, \frac{1}{2}, 0)$, $\quad \beta = X^- c = 2$

(3) $X^- = (0, 0, \frac{1}{3})$, $\quad \beta = X^- c = 2$

Thm:

$\hat{\beta} = \underline{(X'X)^- X'y}$ is a solution to

$\underline{(X'X)\beta = X'y}$

$X\cdot(X'X)^{-1}X'y$

Thm:

$\hat{y} = \underline{X(X'X)^- X'y}$ is the projection of y onto $C(X)$.

$X'X\beta = X'y$

$\Leftrightarrow X'(y - X\beta) = 0$

Pf: $\hat{\beta} = (X'X)^- X'y$

is a solution to the normal equation $X'X\beta = X'y$ $\left[\begin{matrix} y - X\hat{\beta} \perp x_i \\ \text{for all } i = 1, \cdots, p \end{matrix} \right]$

$\Rightarrow \hat{y} = X\hat{\beta} = X\cdot(X'X)^- X'y$ is the proj onto $C(X)$ since the projection is unique.

The next pages give a direct proof.

Thm:

$(X^-)'$ is a version of $(X')^-$.

Pf:

$$X'(X^-)'X'$$
$$= (X X^- X)'$$
$$= X'$$

Sometimes we write $(X^-)' = (X')^-$

Theorem: For any version $(X'X)^-$,

$$\left(X(X'X)^- X'\right) X = X$$

$\Updownarrow$ transpose

$$X'X(X'X)^- X' = X'$$

<u>pf</u> 1: using projection

$P_X = X(X'X)^- X'$ is the proj matrix onto $C(X)$

$X = [x_1, \cdots, x_p]$, $P_X x_j = x_j$

so $P_X [x_1, \cdots, x_p] = [x_1, \cdots, x_p]$

that is $P_X X = X$

Pf 2: using direct matrix manipulation

$\forall\ y \in \mathbb{R}^p$, $y = X\hat{\beta} + e$, where

$X\hat{\beta} = \text{proj}(y \mid c(X))$ and $e \perp c(X)$

$$X'X\ (X'X)^{-}\ X' \cdot y$$
$$= X'X\ (X'X)^{-}\ X' \cdot (X\hat{\beta} + e)$$
$$= X'X(X'X)^{-} X'X\ \hat{\beta} + (X'X)\cdot(X'X)^{-}\cdot\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$
$$= X'X\hat{\beta} = X'y \implies X'X(X'X)^{-}X' = X'$$

Note:

$$X \underset{p\times 1}{y} = 0 \quad \forall\ y \in \mathbb{R}^p \iff X = 0$$

$$X = [x_1, \cdots, x_p]$$

$$y_1 = [1, 0, \cdots, 0]', \ x_1 = 0$$

$$\vdots$$

Thm: Let $P = X \cdot (X'X)^- \cdot X'$

(1) $P = P'$ (symmetric) (2) $P^2 = P$ (idempotent)

(3) P is invariant to $(X'X)^-$

Pf:

(1) $P' = X \cdot (X'X)^- X' = P$

Note: $(X^-)' = (X')^-$

(2) $P^2 = \underbrace{X \cdot (X'X)^- X' \cdot X}_{=X} (X'X)^- X'$

$= X \cdot (X'X)^- X'$

$$\begin{pmatrix} x_1'e \\ x_2'e \\ \vdots \\ x_p'e \end{pmatrix} = 0$$

(3)

$e = y - Xb$

$\forall y \in \mathbb{R}^n$, $y = Xb + e$, $e \perp C(X)$, i.e., $X'e = 0$

and $Xb \in C(X)$, i.e., $Xb = \text{proj}(y \mid C(X))$

Then we see that, for any version of $(X'X)^-$

$X(X'X)^- X'y = X \cdot (X'X)^- X'(Xb + e)$

$= \underbrace{X \cdot (X'X)^- X'X}_{=X} b + 0$, since $X'e = 0$

$= Xb$

$= \text{proj}(y \mid C(X))$.

$X(X'X)^-_{(1)} X'y = X \cdot (X'X)^-_{(2)} X'y$

for all $y \in \mathbb{R}^2$

An Explicit Formula of Projection onto Non-full-rank Subspace

(Optional at the moment)

GI in Least Square with rank < p.

$$X = Q\,(R_1, R_2), \quad R_1^{-1} \text{ exists}, \quad k < p.$$

X: $n\times p$, Q: $n\times k$, R_1: $k\times k$, R_2: $k\times(p-k)$

we assume the first k col. of X are LIN.

$$X'X = \begin{bmatrix} R_1' \\ R_2' \end{bmatrix} Q'Q\,(R_1, R_2), \quad X'y = \begin{bmatrix} R_1' \\ R_2' \end{bmatrix} Q'y$$

$$X'X\beta = X'y \Leftarrow$$ note $[\,X'y \in C(X') = C(X'X)\,]$

$$\Leftrightarrow \begin{bmatrix} R_1'R_1 & R_1'R_2 \\ R_2'R_1 & R_2'R_2 \end{bmatrix} \beta = \begin{bmatrix} R_1' \\ R_2' \end{bmatrix} Q'y$$

$$\text{let } (X'X)^- = \begin{bmatrix} (R_1'R_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$ (one version)

$$\hat\beta = (X'X)^- X'y$$

$$= \begin{bmatrix} (R_1'R_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} R_1' \\ R_2' \end{bmatrix} Q'y$$

$$= \begin{bmatrix} \underbrace{(R_1'R_1)^{-1} R_1' Q'y}_{k\times 1} \\ 0 \end{bmatrix} = \begin{bmatrix} R_1^{-1} Q'y \\ 0 \end{bmatrix}$$

$$\hat y = X\hat\beta = Q\,[R_1, R_2] \cdot \hat\beta = Q \cdot R_1 (R_1'R_1)^{-1} R_1' Q'y$$
$$= Q \cdot Q'y$$

Another way to understand:

$$y = Q(R_1, R_2)\beta + \varepsilon$$

Let $\underset{k\times 1}{b} = (\underbrace{R_1}_{k}, \underbrace{R_2}_{p-k}) \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \begin{matrix} \}k \\ \}p-k \end{matrix}$, $\hat{b} \to \hat{\beta}$

$$= R_1\beta_1 + R_2\beta_2$$

$$y = Qb + \varepsilon$$

$$\hat{b} = Q'y = (Q'Q)^{-1}Q'y, \quad \hat{y} = QQ'y$$

Then we find $\hat{\beta}$ s.t.

$$(R_1, R_2)\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \hat{b} = Q'y$$

Sol 1:

We will set $\beta_2 = 0$, and solve

$$R_1\beta_1 = Q'y \Rightarrow \hat{\beta}_1 = R_1^{-1}Q'y$$

Then $\hat{\beta} = \begin{bmatrix} R_1^{-1}Q'y \\ 0 \end{bmatrix}$

Sol 2: using R^-

$$R^- = \begin{bmatrix} R_1^{-1} \\ 0 \end{bmatrix}$$

$$\hat{\beta} = R^- Q'y = \begin{bmatrix} R_1^{-1}Q'y \\ 0 \end{bmatrix}$$

The $\hat{\beta}$ is the same as using $(X'X)^-$ in solving $X'X\beta = X'y$