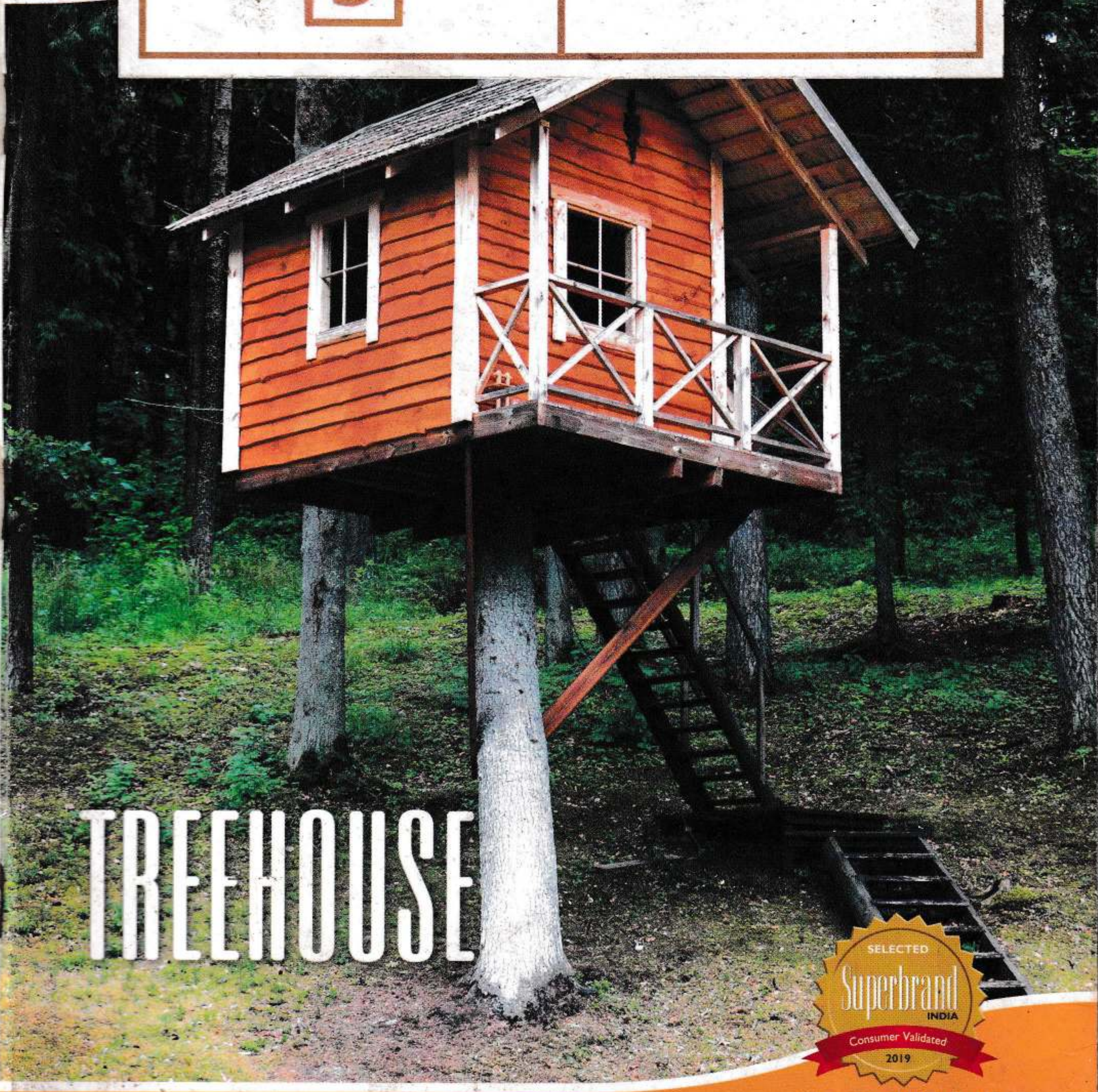
Introduction to Linear Algebra
- Gilbert Strang

5

Vector Spaces & Subspaces

↓

Orthogonality



soorajss1729@gmail.com +91-9400635788

INDEX

5

NAME: SOORAJ · S · STD.: ______ SEC.: ______ ROLL NO.: ______ SUB.: ______

S. No.	Date	Title	Page No.	Teacher's Sign / Remarks
		INTRODUCTION TO		
		LINEAR ALGEBRA		
		— Gilbert Strang, MIT		
		(5th Edition)		

$$\dim(\text{outputs}) + \dim(\text{Null space}) = \dim(\text{inputs})$$

For an $m \times n$ matrix of rank r,

$A_{m \times n}$

outputs = column space

$\dim(\text{outputs}) = r = \dim(C(A))$

$\dim(\text{inputs}) = n$

$\dim(\text{null space}) = n - r$

$$r + (n-r) = n$$

45. Inside $\mathbb{R}^n$, suppose $\dim(V)+\dim(W) > n$.
Show that some non-zero vector is in both V and W.

Ans:
$$\dim(V) + \dim(W) = \dim(V \cap W) + \dim(V+W) > n$$

$$\dim(V+W) \leq n$$

$$\Longrightarrow \dim(V \cap W) > 0$$

$V \cap W$ is non-zero

46 Suppose A is 10×10 and $A^2 = 0$.

$$A^2 = AA = A\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_{10}\end{bmatrix} = \begin{bmatrix}A\vec{a}_1 & A\vec{a}_2 & \cdots & A\vec{a}_{10}\end{bmatrix}$$
$$= \begin{bmatrix}\vec{0} & \vec{0} & \cdots & \vec{0}\end{bmatrix}$$

A multiplies each column of 'A' to give the zero vector.

$$A\vec{a}_1 = \vec{0}, \quad A\vec{a}_2 = \vec{0}, \quad \ldots, \quad A\vec{a}_{10} = \vec{0}$$

$$\boxed{A^2 = 0 \implies C(A) \subset N(A)}$$

$\dim(C(A)) = r$

$\dim(N(A)) = n - r = 10 - r$

$C(A) \subset N(A) \longrightarrow r \leq 10 - r$

$2r \leq 10 \longrightarrow \boxed{r \leq 5}$

46 Suppose A is 10×10 and $A^2 = 0$.

$$A^2 = AA = A\left[\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_{10}\right] = \left[A\vec{a}_1\ A\vec{a}_2 \cdots A\vec{a}_{10}\right]$$
$$= \left[\vec{0}\ \vec{0} \cdots \vec{0}\right]$$

A multiplies each column of 'A' to give the zero vector.

$$A\vec{a}_1 = \vec{0},\ A\vec{a}_2 = \vec{0},\ \cdots,\ A\vec{a}_{10} = \vec{0}$$

$$\boxed{A^2 = 0 \implies C(A) \subset N(A)}$$

$\dim(C(A)) = r$

$\dim(N(A)) = n - r = 10 - r$

$C(A) \subset N(A) \longrightarrow r \leq 10 - r$

$2r \leq 10 \longrightarrow \boxed{r \leq 5}$

$\boxed{3\cdot 5}$

2. Find bases & dimensions for the 4 subspace associated with

ⓐ $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} = R_A$

$x+2y+4z=0$

Ans: Row space – Dimension: 1
Basis: $(1,2,4)$

Null space – Dimension: 2
Basis: $(-2,1,0), (-4,0,1)$

Column space – Dimension: 1
Basis: $(1,2)$

$A^T y = 0$
$y^T A = 0^T$

Left null space – Dimension: 1
Basis: $(-2,1)$

$A^T = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \Longrightarrow y_1 + 2y_2 = 0$

$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

(6) $B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 0 \end{bmatrix}$

Ans:

$\rightarrow \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \end{bmatrix} = R_B$

$x + 4z = 0$
$y = 0$

Row space Basis : $(1,2,4), (2,5,8)$

Column space Basis : $(1,2), (2,5)$

Null space Basis : $(-4,0,1)$

$m - r = 2 - 2 = 0$.

left null space basis is empty

only $y = 0$ is the solution to $A^T y = 0$

$\therefore$ the rows of B are independent

3. Find a basis for each of the 4 subspaces associated with A:

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$A = LU$.

Ans:

Rowspace basis: $(0,1,2,3,4), (0,0,0,1,2)$

row operations does not change row space

Column span basis: Pivot columns of A not U.

$(1,1,0), (3,4,1)$

$3d = 3(-2e) = -6e.$

$\cancel{\ }\; b + 2c + 3d + 4e = 0$

$d + 2e = 0.$

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + e \begin{bmatrix} 0 \\ +2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

Nullspace basis: $(1,0,0,0,0), (0,2,-1,0,0),$

$(0,2,0,-2,1)$

$$A^T = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 0 \\ 3 & 4 & 1 \\ 4 & 6 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$x + y = 0$

$y + z = 0$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

4. Construct a matrix with the required property (or)

(a) Column space contains $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, row space contains $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}$

Ans :- $$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b)

(c)

Ans:-

(d)

Ans:-

(e)

Ans

~~(B) Dimension~~

(c) ~~D(A~~ $\dim(N(A^T)) = 1 + \dim(N(A^T))$

Ans: $n - \cancel{r} = 1 + m - \cancel{r} \Rightarrow \underline{n - m = 1}$

$$\begin{bmatrix} 1 & 2 \end{bmatrix}$$

(d) $N(A)$ contains $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, column space contains $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

Ans: A is 2×2

$$\begin{bmatrix} -9 & 3 \\ -3 & 1 \end{bmatrix}$$

(e) Row space = colum space, $N(A) \neq N(A^T)$

Ans: $m = n \longrightarrow n - r = m - r$

$\dim(N(A)) = \dim(N(A^T))$

N.P.

Next chapter

(7) Suppose the 3x3 matrix A is invertible. Write down bases for the 4 subspaces for A, and also for the 3x6 matrix $B = [A\ A]$.
(The basis for Z is empty)

Ans. ~~Row space basis : (1,0,0), (0,1,0), (0,0,1)~~

(A) $n = m = r = 3$

$$C(A) = \mathbb{R}^3 = C(A^T)$$

Column & Row space basis : $(1,0,0), (0,1,0), (0,0,1)$

$$N(A) = N(A^T) = Z$$

Nullspace & left null space basis are empty

(B)

C

C

$BX =$

check
29 (3.2)

(B) Rowspace basis : $(1,0,0,1,0,0)$, $(0,1,0,0,1,0)$
$(0,0,1,0,0,1)$

$C(B) = \mathbb{R}^3$.

column space basis : $(1,0,0), (0,1,0), (0,0,1)$

$$BX = \begin{bmatrix} A & A \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = AX_1 + AX_2 = 0$$

$$\Longrightarrow X_1 = -X_2$$

check 29(3.2)

Nullspace basis : $(1,0,0,1,0,0)$, $(0,1,0,0,1,0)$,
$(0,0,1,0,0,1)$

$n - r = 3 - 3 = 0$

Nullspace basis is empty

8. What are the dimensions of the 4 subspaces for A, B, C, if I is the 3x3 identity matrix and O is the 3x2 zero matrix

$I_{3\times 3}$, $O_{3\times 2}$

Ans: ⓐ $A = [I \; O]$

Ans: $\dim(C(A^T)) = 3$; $\dim(C(A)) = 3$;

$\dim(N(A)) = 5 - 3 = 2$; $\dim(N(A^T)) = 0$

ⓑ $B = \begin{bmatrix} I & I \\ O^T & O^T \end{bmatrix}$

Ans: $\dim(C(B')) = \dim(C(B)) = 3$

$\dim(N(A)) = 6 - 3 = 3$

$\dim(N(A^T)) = 5 - 3 = 2$

ⓒ $C = [O] = [$

Ans: $\dim(C(C^T)) = \dim(C(C)) = 0$

$C = [O] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

$\dim(N(C)) = 2 - 0 = 2$

$\dim(N(C^T)) = 3 - 0 = 3$

9. Which subspaces are the same for these matrices of different sizes?

Prove that all 3 of those matrices have the same rank?

Ans:

(a) $[A]$ and $\begin{bmatrix} A \\ A \end{bmatrix}$

Ans: same row space & null space

(b) $\begin{bmatrix} A \\ A \end{bmatrix}$ and $\begin{bmatrix} A & A \\ A & A \end{bmatrix}$.

Ans: $\begin{bmatrix} A^T & A^T \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = A^T X + A^T Y = 0 \Rightarrow X + Y = 0$

$$\begin{bmatrix} A^T & A^T \\ A^T & A^T \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} A^T X + A^T Y \\ A^T X + A^T Y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow X + Y = 0.$$

Same column space & left null space.

10. If the entries of a 3x3 matrix are chosen randomly b/w 0 and 1, what are the most likely dimensions of the 4 subspaces?
What if the random matrix is 3x5?

Stark 15/9/2020

Ans:

Suppose,

each column of the 3x3 matrix is drawn independently from the uniform distribution on the cube $[0,1]^3$.

The probability that each column lies in the region 'R' of the cube is equal to the volume of R.

ie 1^{st} column is drawn.

After drawing the 1^{st} column, you need to draw the 2^{nd} column so that it does not lie on the subspace spanned by the 1^{st} column. This subspace is a line inside the cube $[0,1]^3$, that has zero volume.

$\therefore$ The probability that the 2^{nd} column lies in this subspace is zero.

ie., the probability that the 2^{nd} column is linearly independent of the 1^{st} column is 1.

After drawing the 1^{st} & 2^{nd} columns (where the 2^{nd} column is linearly independent of the 1^{st} column), you need the 3^{rd} column to lie outside the subspace spanned by the 1st two columns. This subspace is a plane inside the cube, and this plane also has zero volume.

So again, the probability that the 3^{rd} column lies in this subspace is zero.

ie.,
the probability that the 3rd column lies
in this subspace is zero.

ie., the probability that the 3^{rd} column is linearly independent of the 1^{st} two columns is 1.

⇒ For all distributions of $[0,1]^3$ that has a density function,

the prob. of lying in a 1-D (or) 2-D subspace is zero.

So,
the probability that a given column is linearly independent of the previous columns is 1.

For $A_{3\times 5}$, rank is 3.

11. A is an $m \times n$ matrix of rank r. Suppose there are right sides b for which $Ax = b$ has no solution.

(a) What are all inequalities that must be true between m, n & r?

Ans: No solution $\Longrightarrow$ rref$(A|b)$ has a row $[0, 0, \ldots, 0 | 1]$

$$\therefore \underline{\underline{r < m}}.$$

& $r \leq n$

Can't compare m & n.

(b) How do you know that $A^T y = 0$ has solutions other than $y = 0$?

Ans. $m > r \Longrightarrow \underline{\underline{m - r > 0}}$

$\therefore$ left null space must contain a non-zero vector.

12. Construct a matrix with (1,0,1) and (1,2,0) as a basis for its row space and its column space. Why can't this be a basis for the row space and null space?

Ans:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$n = 3 = (n-r) + r$$

But $2+2=4$

13. True / False

(c) If A & B share the same 4 subspaces then A is a multiple of B.

Ans: Consider 2 matrices A & B of same size & invertible

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

~~column space (A) = column space~~

$$C(A) = C(B) = C(A^T) = C(B^T) = \mathbb{R}^3$$

$$N(A) = N(B) = Z$$

$$N(A^T) = N(B^T) = Z$$

14. Without computing A, find bases for its 4 fundamental subspaces

$$\underset{3\times 4}{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 9 & 8 & 1 \end{bmatrix}}_{3\times 3} \underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}}_{3\times 4} = BC$$

free column

Ans: ~~Row~~ Each row of A is a linear combination of the rows of C.

$\Rightarrow$ $\boxed{\text{Row space basis : } (1,2,3,4),\ (0,1,2,3),\ (0,0,1,2)}$

$$\left.\begin{aligned} c+2d &= 0 \\ b+2c+3d &= 0 \\ a+2b+3c+4d &= 0 \end{aligned}\right\} \quad \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = d\begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

$b = -3d-2c$
$= -3d-2[-2d] = -3d+4d = d$

$a = -2d+6d-4d = 0$

$\boxed{\text{Null space basis is same as that of } C \ : \ (0,1,-2,1)}$

Each column of A is a linear combination of the columns of B.

$C(A) = C(B) = \mathbb{R}^3 \longleftarrow B_{3\times 3}$ is invertible

column space basis : $(1,0,0), (0,1,0), (0,0,1)$

13. True / False

(c) If A & B share the same 4 subspaces then A is a multiple of B.

Ans: Consider 2 matrices A & B of same size & invertible

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

~~column space (A) = column space~~

$$C(A) = C(B) = C(A^T) = C(B^T) = \mathbb{R}^3$$

$$N(A) = N(B) = Z$$

$$N(A^T) = N(B^T) = Z$$

14. Without computing A, find bases for its 4 fundamental subspaces

$$\underset{3\times 4}{A} = \underset{3\times 3}{\begin{bmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 9 & 8 & 1 \end{bmatrix}} \underset{3\times 4}{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}} = BC$$

free column

Ans: ~~Row~~ Each row of A is a linear combination of the rows of C.

$\Longrightarrow$ $\boxed{\text{Row space basis : } (1,2,3,4), (0,1,2,3), (0,0,1,2)}$

$$\left.\begin{aligned} c+2d &= 0 \\ b+2c+3d &= 0 \\ a+2b+3c+4d &= 0 \end{aligned}\right\} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = d \begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

$b = -3d - 2c$
$= -3d - 2[-2d] = -3d + 4d = d$

$a = -2d + 6d - 4d = 0$

$\boxed{\text{Null space basis is : } (0,1,-2,1) \text{ same as that of } C}$

Each column of A is a linear combination of the columns of B.

$C(A) = C(B) = \mathbb{R}^3 \longleftarrow B_{3\times 3}$ is invertible

column space basis : $(1,0,0), (0,1,0), (0,0,1)$

Left nullspace has empty basis since B is invertible.

16. Explain why $v = (1, 0, -1)$ can not be a row of A & also in the nullspace

Ans.

$$Av = 0 \implies \begin{bmatrix} \vdots \\ v^T \\ \vdots \end{bmatrix} v = 0 = \begin{bmatrix} \vdots \\ v^T v \\ \vdots \end{bmatrix}$$

$$v^T v = v \cdot v = 1 - 1 = 0 = |v|$$

$$\implies v = 0$$

$$\boxed{\text{Not true for } v = (1, 0, -1)}$$

ertible.

18.

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & b_1 \\ 4 & 5 & 6 & b_2 \\ 7 & 8 & 9 & b_3 \end{array}\right] \rightarrow \begin{bmatrix} 1 & 2 & 3 & b_1 \\ 0 & -3 & -6 & b_2-4b_1 \\ 0 & 0 & 0 & b_3-2b_2+b_1 \end{bmatrix}$$

What combination of the rows of 'A' has produced the zero row?

Which vectors are in $N(A^T)$ & which vectors are in $N(A)$?

Ans: ~~is~~ $b_3 - 2b_2 + b_1 = 0$

row 3 − 2(row 2) + row 1 = 0

$A^T y = 0 \longrightarrow y^T A = 0^T$

Linear combination of rows of $A = 0$

$N(A^T) = \underline{\underline{c(1, -2, 1)}}$.

$N(A) = c(1, -2, 1)$

19- reduce A to echelon form & look at zero rows. The b column tells which combinations you have taken of the rows:

From the b column after elimination, read off $(m-r)$ basis vectors in the left nullspace.
Those y's are combinations of rows that give zero rows in the echelon form.

(a) $A = \begin{bmatrix} 1 & 2 & b_1 \\ 3 & 4 & b_2 \\ 4 & 6 & b_3 \end{bmatrix}$

Ans: $\longrightarrow \begin{bmatrix} 1 & 2 & b_1 \\ 0 & -2 & b_2-3b_1 \\ 0 & -2 & b_3-4b_1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & b_1 \\ 0 & 2 & 3b_1-b_2 \\ 0 & 0 & b_1 - b_2 + b_3 \end{bmatrix}$

$-b_1 - b_2 + b_3 = 0 \implies (-1, -1, 1)$,
$-(\text{row } 1) - (\text{row } 2) + \text{row } 3 = 0$

$\therefore (-1,-1,1)$ is in $N(A^T)$

(b) $\begin{bmatrix} 1 & 2 & b_1 \\ 2 & 3 & b_2 \\ 2 & 4 & b_3 \\ 2 & 5 & b_4 \end{bmatrix}$

Ans: $\longrightarrow \begin{bmatrix} 1 & 2 & b_1 \\ 0 & -1 & b_2-2b_1 \\ 0 & 0 & b_3-2b_1 \\ 0 & 1 & b_4-2b_1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & b_1 \\ 0 & 1 & 2b_1-b_2 \\ 0 & 0 & b_3-2b_1 \\ 0 & 0 & -4b_1+b_2+b_4 \end{bmatrix}$

$$\therefore \left.\begin{array}{r} -2b_1+b_3=0 \\ -4b_1+b_2+b_4=0 \end{array}\right\} \begin{array}{l} -2(\text{row } 1)+(\text{row } 3)=0 \\ -4(\text{row } 1)+(\text{row } 2)+(\text{row } 4)=0 \end{array}$$

$(-2,0,1,0), (-4,1,0,1)$ are in $N(B^T)$

22. Construct $A = uv^T + wz^T$ whose column space has basis $(1, 2, 4)$, $(2, 2, 1)$ and whose row space has basis $(1, 0)$, $(1, 1)$.
Write A as (3×2) times (2×2)

Ans.
$$A = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 4 & 2 \\ 5 & 1 \end{bmatrix}$$

$$= uv^T + wz^T = \begin{bmatrix} u & w \end{bmatrix} \begin{bmatrix} v^T \\ z^T \end{bmatrix}$$

25. (c) If $C(A^T) = C(A)$ then $A^T = A$

Ans. $\boxed{\text{False}}$

~~$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$~~

Ex :- $A = \begin{bmatrix} 1 & 2 \\ -2 & 4 \end{bmatrix}$

$$A^T = \begin{bmatrix} 1 & -2 \\ 2 & 4 \end{bmatrix} \neq A.$$

$$C(A) = R(A^T) = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right)$$

---

27. Find the ranks of the 8×8 checkerboard matrix B & the chess matrix C:

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ \vdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \vdots \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} r & n & b & q & k & b & n & r \\ p & p & p & p & p & p & p & p \\ \cdot & \cdot & \cdot & 4 & \text{rows} & - & - & - \\ p & p & p & p & p & p & p & p \\ r & n & b & q & k & b & n & r \end{bmatrix}$$

The #'s r, n, b, q, k, p are all different. Find bases for the row space & left null space of B & C.

Find a basis for the nullspace of C.

Ans: $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \Longrightarrow \text{rank}(B) = 2$

$\begin{vmatrix} r & n \\ p & p \end{vmatrix} = p(r-n) \neq 0$ if $p \neq 0$

rank (C) = 2 if $p \neq 0$

row 1 & row 2 are a basis for the row space of C.

$m - r = 8 - 2 = 6 = n - r$

$(1, 0, -1, 0, 1, 0, 0, 0), (1, 1, -1, -1, 1, 1, -1, -1)$

$a + c + e + g = 0$

$b + d + f + h = 0$

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = c \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + e \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + f \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + g \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + h \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

are the Bases of $\mathcal{N}(B)$ & $\mathcal{N}(B^T)$

$$\boxed{r x_1 + n x_2} + b x_3 + q x_4 + k x_5 + b x_6 + n x_7 + r x_8 = 0$$

$$p(x_1 + x_2 \cdots x_8) = 0$$

$$C^T = \begin{bmatrix} r & p & 0 & \cdots & 0 & p & r \\ n & p & 0 & \cdots & 0 & p & n \\ b & p & 0 & \cdots & 0 & p & b \\ q & p & 0 & \cdots & 0 & p & q \\ k & p & 0 & \cdots & 0 & p & k \\ b & p & 0 & \cdots & 0 & p & b \\ n & p & 0 & \cdots & 0 & p & n \\ r & p & 0 & \cdots & 0 & p & r \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} r & p & 0 & \cdots & 0 & p & r \\ 1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 0 & & 0 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

$\cancel{r x_1 + p x_2 + p x_7 + p x_8} = 0$

$\cancel{x_1 + x_8} = 0$

Basis of $N(C^T)$:
$(-1,0,0,0,0,0,0,1), (0,-1,0,0,0,0,1,0),$
$(0,0,-1,0,0,1,0,0), (0,0,0,-1,+1,0,0,0)$
$(0,0,1,0,0,1,0,0), (0,0,0,1,1,0,0,0)$

Basis of $N(C)$: $(-1,0,0,0,0,0,0,1), (0,-1,0,0,0,0,1,0),$
$(0,0,-1,0,0,1,0,0), ($

?

29. If $A = uv^T$ is a 2×2 matrix of rank 1,
draw figure to show the 4 fundamental subspaces. If B produces those same 4 subspaces, what is the exact relation of B to A?

Ans: $A_{2\times 2} = uv^T$

$C(A)$ & $C(A^T)$ are lines in $\mathbb{R}^2$

$uv^T x = uv^T x = u(v \cdot x) = 0$

$N(A) \perp$ ~~row~~ v in the row space

$n - r = 2 - 1 = 1$

$R(A)$ & $N(A)$: v, v^T

$C(A)$, & $N(A^T)$: u, u^T

If B produces those same 4 subspaces then $B = cA$ with $c \neq 0$

30) M is the space of 3×3 matrices.
Multiply every matrix X in M by

$$A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} . \quad A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

(a) Which matrices X lead to $AX = 0$?

$$X = \begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix}$$

Ans: $AX = 0$ if each column of X is a multiple of $(1,1,1)$.

$$\Longrightarrow \underline{\dim(N(A)) = 3}$$

(b) Which matrices have the forme AX for some matrix X?

Ans: $\underset{3\times3}{A}\underset{3\times3}{X} = \underset{3\times3}{B} \Longrightarrow \left[\begin{array}{ccc|c|c|c} 1 & 0 & -1 & b_1 & c_1 & d_1 \\ 0 & 1 & -1 & b_2+b_1 & c_2+c_1 & d_2+d_1 \\ 0 & 0 & 0 & b_1+b_3+b_2 & c_1+c_2+c_3 & d_1+d_2+d_3 \end{array}\right]$

all columns of B add to zeros

$AX = B$ iff $B = \begin{bmatrix} a & b & c \\ d & e & f \\ -a-d & -b-e & -c-f \end{bmatrix}$.

$Ax = b \Rightarrow \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$, $x + y + z = 0$

, $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$\dim(C(A)) = 6$

$3 + 6 = 9$, which is the dimension of the space of inputs of this matrix

31. Suppose the $m \times n$ matrices A & B have the same 4 subspaces. If they are both in rref, prove that $F = G$,

$$A = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} I & G \\ 0 & 0 \end{bmatrix}$$

Ans: Rowspace (A) = Rowspace (B)

Row 1 of A = combination of the rows of B

$\Rightarrow$ Only possible combination is

$$1 \cdot [I \;\; F] = 1 \, [I \;\; G]$$

$$\Rightarrow F = G$$

$\boxed{4}$

ORTHOGONALITY

Orthogonal vectors : $v^T w = v \cdot w = 0$ & $|v|^2 + |w|^2 = |v+w|^2$

* Two subspaces V and W of a vector space are orthogonal if every vector v in V is ⊥r to every vector w in W.

$$\boxed{\text{Orthogonal subspaces} : V^T W = 0 \text{ for all } v \in V \text{ and all } w \in W}$$

Test

Put bases for the subspaces V and W into the columns of V and W.

Orthogonal subspaces V & W : $V^T W = 0$

Ex:1 The floor of your room (extended to ∞) is a subspace V. The line where 2 walls meet is a subspace W (1D). Those subspaces are orthogonal. Every vector up the meeting line of the walls is $\perp$ to every vector in the floor.

Ex:2 2 walls look $\perp$ but those 2 subspaces are not orthogonal.

The meeting line is in both V and W. & this line is not $\perp$ to itself.
Two planes (dimensions 2 and 2 in $\mathbb{R}^3$) can't be orthogonal subspaces.

$\Rightarrow$ When a vector is in 2 orthogonal subspaces, it must be zero. It is $\perp$ to itself.

- Zero is the only point where the nullspace meets the row space.

* The nullspace $N(A)$ and the row space $C(A^T)$ are orthogonal subspaces of $\mathbb{R}^n$.

Every vector 'x' in the nullspace is $\perp$ to every row of A; because $Ax = 0$.

Proof: 1

$$Ax = \begin{bmatrix} \text{row } 1 \\ \vdots \\ \vdots \\ \text{row } m \end{bmatrix} x = \begin{bmatrix} (\text{row } 1) \cdot x \\ \vdots \\ (\text{row } m) \cdot x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

→ x is $\perp$ to every combination of the rows.

∴ The whole row space $C(A^T)$ is orthogonal to $N(A)$

Proof 2:

Vectors in the row space are combinations of the rows. Take the dot product of $A^T y$ with any x in the nullspace, i.e., $Ax = 0$

$$x \cdot (A^T y) = x^T (A^T y) = (Ax)^T y = 0^T y = 0$$

* The left nullspace $N(A^T)$ and the column space $C(A)$ are orthogonal in $\mathbb{R}^m$.

Proof 1:

$$A^T y = \begin{bmatrix} (\text{column } 1)^T \\ \vdots \\ (\text{column } n)^T \end{bmatrix} y = \begin{bmatrix} (\text{column } 1) \cdot y \\ \vdots \\ (\text{column } n) \cdot y \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Proof 2:

Vectors in $C(A)$ are combinations $A\cancel{x}$ of the columns.

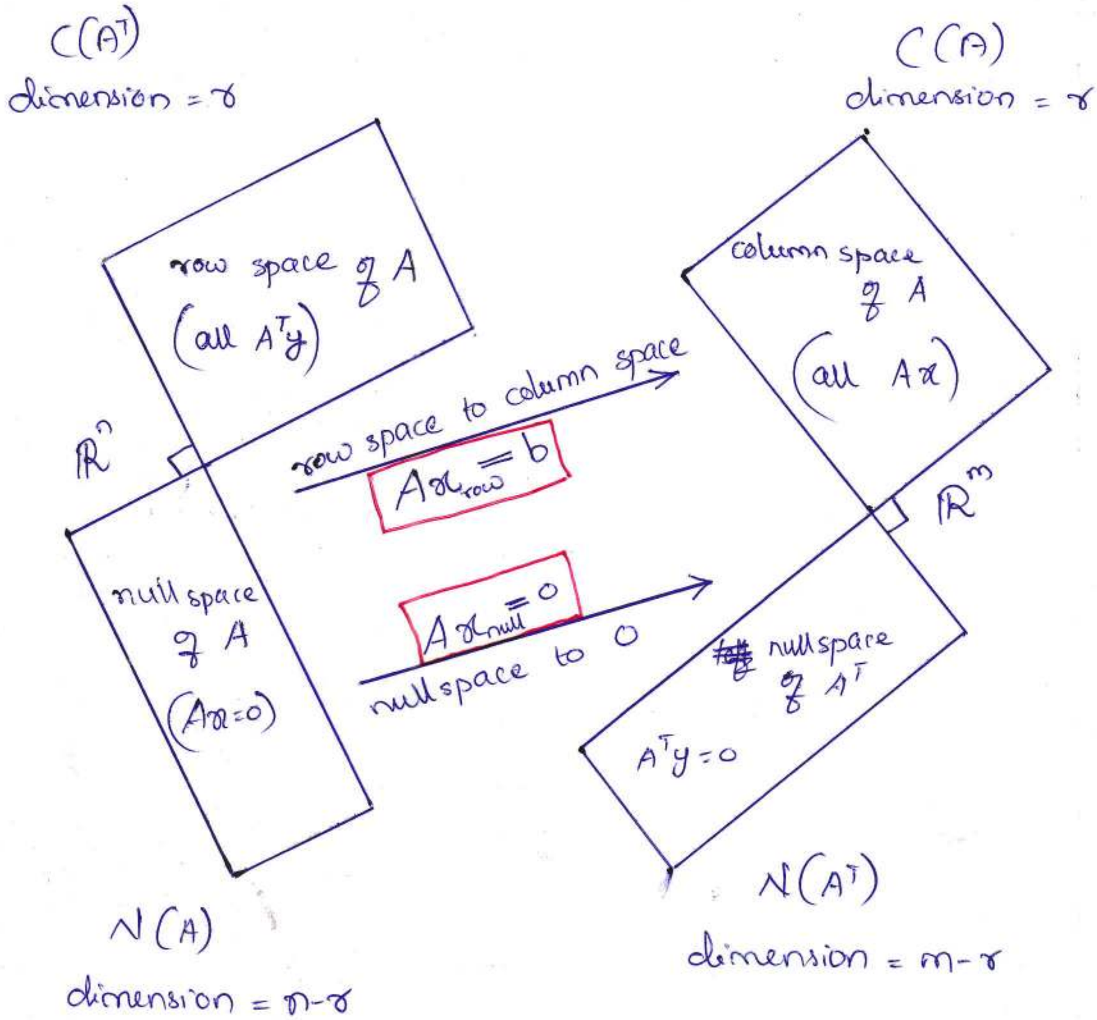
$y \in N(A^T) \longrightarrow A^T y = 0 \iff y^T A = 0^T$

$$y \cdot (Ax) = y^T (Ax) = (A^T y)^T x = 0^T x = 0$$

$C(A^T) \perp N(A)$	in $\mathbb{R}^n$
$C(A) \perp N(A^T)$	in $\mathbb{R}^m$

$C(A^T) \oplus N(A) = \mathbb{R}^n$

$C(A) \oplus N(A^T) = \mathbb{R}^m$



$C(A^T)$
dimension = r
row space of A
(all A^Ty)
$\mathbb{R}^n$
null space of A
($Ax=0$)
$N(A)$
dimension = $n-r$
row space to column space
$Ax_{row} = b$
$Ax_{null} = 0$
nullspace to 0
$C(A)$
dimension = r
column space of A
(all Ax)
$\mathbb{R}^m$
nullspace of A^T
$A^Ty=0$
$N(A^T)$
dimension = $m-r$

Orthogonal Complements

$\mathbb{R}^n$: $r + (n-r) = n$

Two lines could be $\perp$ in $\mathbb{R}^3$, but those lines could not be the row space and null space of a 3×3 matrix.

* The orthogonal complement of a subspace V contains every vector that is $\perp$ to V.
This orthogonal subspace is denoted by $V^\perp$ ("V perp").

Every 'x' that is $\perp$ to the rows satisfies $Ax=0$, and lies in the null space.

If v is orthogonal to the nullspace, it must be in the row space. Otherwise we could add this v as an extra row of the matrix, without changing its null space. The rowspace could grow, which breaks the law $r+(n-r)=n$.

$$\Longrightarrow N(A)^{\perp} = C(A^T)$$

* If a subspace S is contained in a subspace V, $S^{\perp}$ contains $V^{\perp}$.

[4.1] (23) $S \subset V \Rightarrow V^{\perp} \subset S^{\perp}$

Fundamental Theorem of Linear Algebra — Part 2

$N(A)$ is the orthogonal complement of the row space $C(A^T)$ in $\mathbb{R}^n$

$N(A^T)$ is the orthogonal complement of the column space $C(A)$ in $\mathbb{R}^m$.

"complements" $\Rightarrow$ every 'x' can be split into a row space component x_r & a nullspace component x_n.

$$\boxed{\begin{aligned} &Ax = A(x_r + x_n) = Ax_r + Ax_n \\ &Ax_n = 0 \text{ : The nullspace component goes to zero} \\ &Ax_r = Ax \text{ : The rowspace component goes to the column space} \end{aligned}}$$

$\Rightarrow$ Every vector goes to the column space. Multiplying by A can not do anything else

- Every vector b in the column space comes from one & only one vector x_r in the row space.

Proof

If $Ax_r = Ax_r' \implies A(x_r - x_r') = 0$

$x_r - x_r'$ is in the $N(A)$

It is also in the row space, since x_r, x_r' are in the row space of A.

$N(A) \perp C(A^T) \longrightarrow$ $x_r - x_r'$ must be the zero vector.

$\therefore$ $x_r = x_r'$

- There is an $r \times r$ invertible matrix hiding inside A, if we throw away the 2 null spaces ($N(A)$ & $N(A^T)$).

From the row space to the column space, A is invertible.

The "pseudoinverse" will invert that part of A. in section 7.4.

Ex: 4 Every matrix of rank r has an $r \times r$ invertible submatrix:

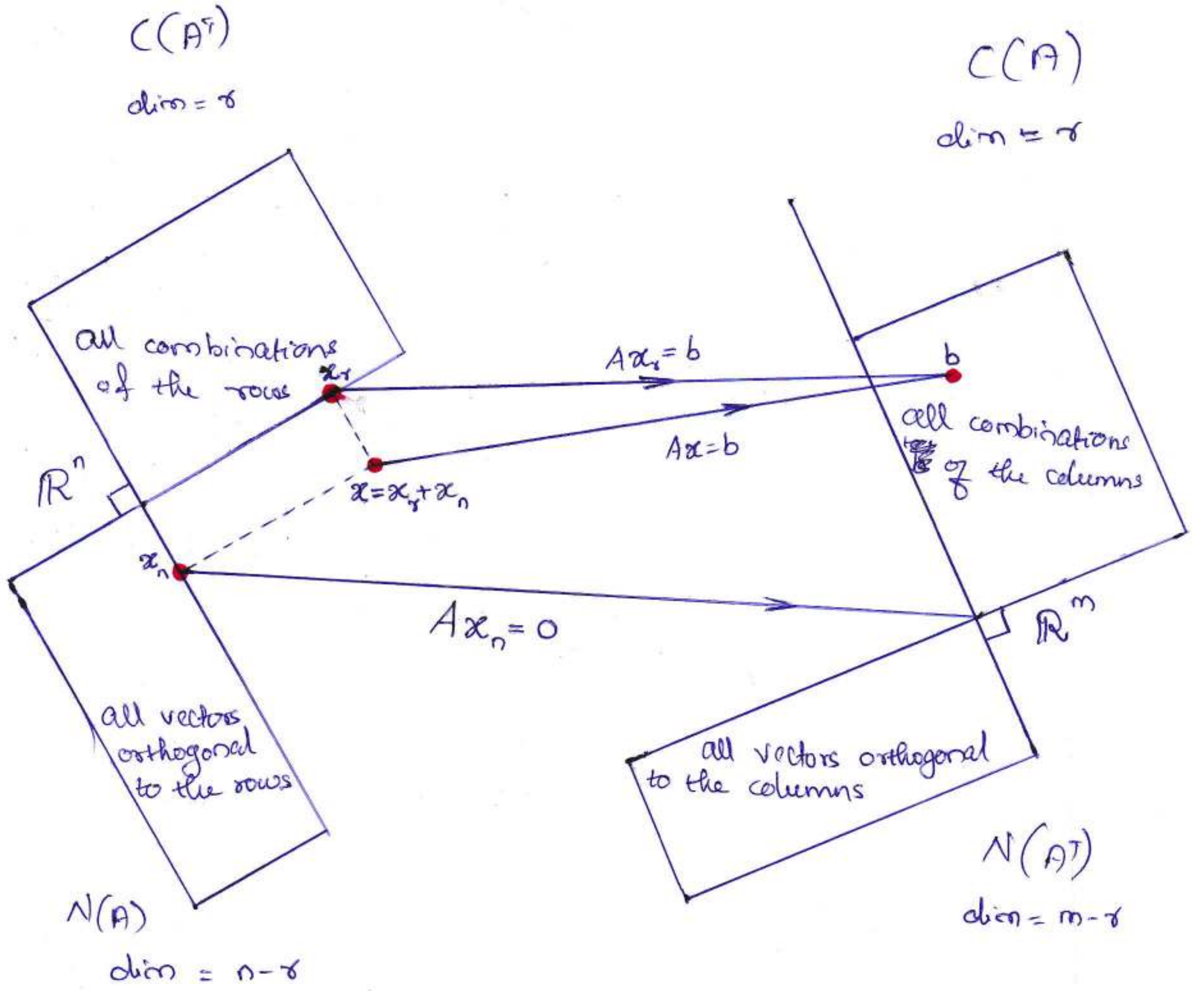
$$A = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ contains the submatrix $\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$.

The other 11 zeros are responsible for the nullspaces. The rank of B is also $\cancel{2} = 2$.

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & 6 \\ 1 & 2 & 4 & 5 & 6 \end{bmatrix}$$ contains $\begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$ in the pivot rows & columns.

Every matrix can be diagonalized, when we choose the right bases for $\mathbb{R}^n$ and $\mathbb{R}^m$.

$\Longrightarrow$ Singular Value Decomposition.


C(A^T)
dim = r
C(A)
dim = r
all combinations of the rows
x_r
Ax_r = b
b
Ax = b
all combinations of the columns
R^n
x = x_r + x_n
x_n
Ax_n = 0
R^m
all vectors orthogonal to the rows
all vectors orthogonal to the columns
N(A^T)
dim = m - r
N(A)
dim = n - r

Combining Bases from Subspaces

* Any n independent vectors in $\mathbb{R}^n$ must span $\mathbb{R}^n$. So they are a basis.

Any n vectors that span $\mathbb{R}^n$ must be independent. So they are a basis.

If the n columns of A are independent, they span $\mathbb{R}^n$. So $Ax=b$ is solvable.

If the n columns span $\mathbb{R}^n$, they are independent. So $Ax=b$ has only one solution.

If there are no free variables, the solution x is unique. There must be n pivot columns. Then the backsubstitution solves $Ax=b$. (the solution exists).

Suppose that $Ax=b$ can be solved for every 'b' (existence of solutions). Then elimination produces no zero rows. There are n pivots & no free variables. The nullspace contains only $x=0$ (uniqueness of solutions).

$$\left.\begin{array}{l}\text{Uniqueness} \Longrightarrow \text{existence} \\ \& \\ \text{Existence} \Longrightarrow \text{uniqueness}\end{array}\right\} A \text{ is invertible.}$$

* Each x is the sum ~~of~~ $x_r + x_n$ of a rowspace vector x_r and a nullspace vector x_n.

$$\boxed{x = x_r + x_n}$$

Ex: 5 For $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ split $x = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ into

$$x_r + x_n = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$x_1 + 2x_2 = 0$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Projection

The prototype of an oblique projection is the shadow cast by objects on the ground (when the sun is not vertical). Mathematically, we have a plane (the ground), a direction (direction of the rays of light, which are approximately parallel, since the sun is very far), and the projection of a person along the light rays is its shadow on the ground.

More generally,
we can have projections ||el to higher dimensional subspaces.

* Given a vector space V, a projection is a linear transformation $P: V \to V$ such that $P^2 = P$.

- In the finite dimensional case, a square matrix P is called a projection matrix if it is equal to its square.

  i.e., $P^2 = P \longrightarrow$ idempotent

- The eigenvalues of a projection matrix must be 0 or 1.

  i.e., $\lambda = 0$ (or) 1

Proof.

$Pv = \lambda v$, $P^2 v = \lambda^2 v = \lambda v = Pv$

$(\lambda^2 - \lambda)v = \lambda(\lambda - 1)v = 0$

$v \neq 0 \implies \lambda(\lambda - 1) = 0 \Rightarrow \lambda = 0$ (or) $\lambda = 1$

* Let V be a finite dimensional vectorspace and $P : V \to V$ be a projection on V.

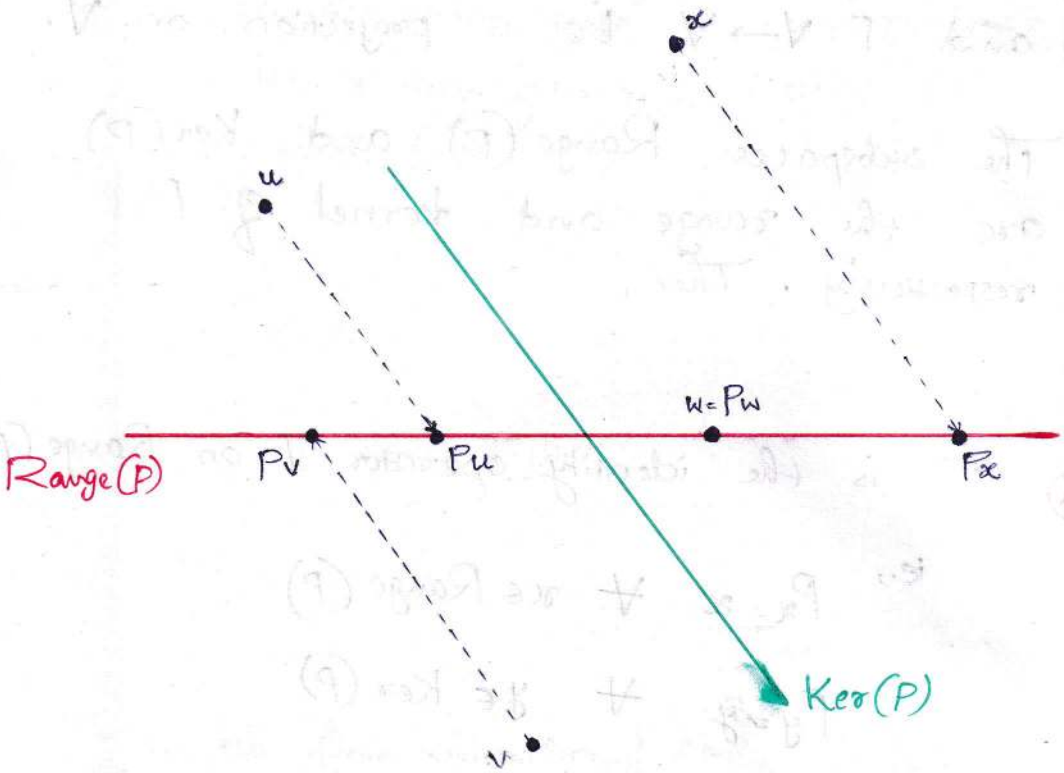
The subspaces Range (P) and Ker (P) are the range and kernel of P, respectively. Then,

(1.) P is the identity operator I on Range (P).

ie., $Px = x \quad \forall \; x \in \text{Range}(P)$

$Py = 0 \quad \forall \; y \in \text{Ker}(P)$

$\Rightarrow$ Range (P) is the subspace on which P projects, and the projection is parallel to (or, along) the subspace Ker (P).


x
u
w=Pw
Range(P)
Pv
Pu
Px
Ker(P)
v

② The vector space V can be written as the direct sum, $\boxed{V = \text{Range}(P) \oplus \text{Ker}(P)}$

⇒ The range and the kernel of a projection P are complementary, as are P and $I-P$. $\boxed{\text{Range}(P) \cap \text{Ker}(P) = \{0\}}$

Every vector $x \in V$ may be decomposed uniquely as, $x = x_R + x_N$ with

$x_R = Px \in \text{Range}(P)$ and

$x_N = (I-P)x = x - Px \in \text{Ker}(P)$.

Proof

$Px_N = P(x - Px) = Px - P^2x = Px - Px = 0 \implies x_N \in \text{Ker}(P)$

⇒ Let U and W be subspaces of a vector space V. A linear transformation $T: V \to V$ is called the projection of V onto the subspace U along W if $V = U \otimes W$ and $Tx = u$ for $x = u + w \in U \oplus W$.

* If $P: V \to V$ is a projection, then

$$V = \text{Range}(P) \oplus \text{Ker}(P)$$

Proof.

Step 1 : $V = \text{Range}(P) + \text{Ker}(P)$

$\text{Ker}(P) \subseteq V$ & $\text{Range}(P) \subseteq V$

$$\implies \text{Range}(P) + \text{Ker}(P) \subseteq V$$

Let $x \in V$,

Let, $y = P(x) \in \text{Range}(P)$ and $z = x - y$ for some $x \in V$.

then $x = y + z$ and

$$P(z) = P(x-y) = P(x) - P(y) = y - y = 0$$

$\Rightarrow z \in \text{Ker}(P)$

$\therefore \quad x = y + z \in \text{Range}(P) + \text{Ker}(P)$

$$\implies V \subseteq \text{Range}(P) + \text{Ker}(P)$$

$$\left.\begin{array}{c} \text{Range}(P) + \text{Ker}(P) \subseteq V \\ \& \\ V \subseteq \text{Range}(P) + \text{Ker}(P) \end{array}\right\} \quad \underline{V = \text{Range}(P) + \text{Ker}(P)}$$

Step 2 : Range $(P) \cap$ Ker$(P) = \{0\}$

Let $x \in$ Range$(P) \cap$ Ker(P),

$$\underset{x \in \text{Range}(P)}{x} = \underset{x \in \text{Ker}(P)}{P(x) = 0} \implies x = 0$$

$$\implies \text{Range}(P) \cap \text{Ker}(P) = \{0\}.$$

$$\left.\begin{array}{l} V = \text{Range}(P) + \text{Ker}(P) \\ \text{Range}(P) \cap \text{Ker}(P) \end{array}\right\} \quad \underline{V = \text{Range}(P) \oplus \text{Ker}(P)}$$

* If V is a vectorspace and $W_1, W_2 \subseteq V$ are subspaces such that $V = W_1 \oplus W_2$,
then there exists a unique projection $P: V \to V$ such that $\text{Range}(P) = W_1$ and $\text{Ker}(P) = W_2$.

Proof $\quad V = W_1 \oplus W_2$

Any $x \in V$ has a unique expression $x = y + z$ where $y \in W_1$ and $z \in W_2$

$\therefore$ The transformation $P: V \to V$ defined by

$$P(x) = y \quad \text{is well defined.}$$

$y \in \text{Range}(P) = W_1$

The zero element is uniquely expressed as $0 = 0 + 0$, so $P(0) = 0$

If $x, x' \in V$, they have unique expressions.

$$x = y + z \quad \& \quad x' = y' + z'$$

where $y, y' \in W_1$ and $z, z' \in W_2$

$\Longrightarrow x + x' = y + y' + z + z'$

~~By uniqueness~~, This must be the expression for $x + x'$ by uniqueness.

$$\therefore \quad P(x + x') = y + y' = P(x) + P(x')$$

By the same reasoning,

$x = y + z \Longrightarrow cx = cy + cz$ for any scalar c.
where $cy \in W_1$ & $cz \in W_2$

$\therefore$

$P(cx) = cy = cP(x)$.

$\Longrightarrow P : V \to V$ is a linear transformation.

If $P(x) = y$, where $y \in W_1$

then y has a unique sum $y = y + 0$

$\therefore$

$P^2(x) = P(P(x)) = P(y) = y = P(x)$

$\Longrightarrow P^2 = P$, P is a projection.

By choice, Range $(P) \subseteq W_1$,

if $y \in W_1$, its unique sum is $y = y + 0$

so, $P(y) = y \in$ Range(P).

$\Longrightarrow$ Range$(P) = W_1$

If $z \in W_2$, then

its unique sum is $z = 0 + z$

So, $P(z) = 0 \implies z \in Ker(P)$.

$\therefore \quad W_2 \subseteq Ker(P)$

If $x \in Ker(P)$,

$P(x) = 0$

Let x can be uniquely written as $x = y + z$

where $y \in W_1$ and $z \in W_2$

Then, $P(x) = y = 0 \implies x = z \in W_2$

$\implies \underline{\underline{Ker(P) = W_2}}$

* Let $P, Q : V \to V$ be projections. Then $P+Q : V \to V$ is a projection iff $\text{Range}(P) \subset \text{Ker}(Q)$ and $\text{Range}(Q) \subset \text{Ker}(P)$

Proof

Step 1: Assume that,

$\text{Range}(P) \subset \text{Ker}(Q)$ & $\text{Range}(Q) \subset \text{Ker}(P)$

Let, $z \in \text{Range}(P+Q)$ then $z = x+y$
for some $x \in \text{Range}(P)$ and $y \in \text{Range}(Q)$
$y \in \text{Ker}(P)$ and $x \in \text{Ker}(Q)$

Since $(P+Q)v = P(v) + Q(v)$, $P(v) = x$, $Q(v) = y$

$\therefore$

$$(P+Q)z = P(z) + Q(z) = P(x) + P(y) + Q(x) + Q(y)$$

$$= x + 0 + 0 + y = x + y = z$$

$\therefore$ $P+Q$ is a projection on its image

$\Longrightarrow$ $P+Q$ is a projection.

Step 2: Suppose that there exists $x \in \text{Range}(P)$ that does not belong to $\text{Ker}(Q)$.

i.e., $x \in \text{Range}(P)$ & $x \notin \text{Ker}(Q)$

and $P+Q$ is a projection.

Let, $y = Qx \neq 0$

$$x+y = P(x)+y = P(x)+Q(x)$$
$$= (P+Q)(x) = (P+Q)^2(x)$$
$$= (P+Q)(x+y) = P(x)+P(y)+Q(x)+Q(y)$$
$$= x+P(y)+y+y = x+P(y)+2y$$
$$\implies P(y) = -y$$

$y \neq 0 \implies y \in \text{Range}(P)$ but $P(y) = y$

Contradiction that P is a projection.

Similarly,

* Suppose $P, Q : V \to V$ are projections.
Assume that $P+Q$ is a projection as well.
Then,

$$\text{Range}(P+Q) = \text{Range}(P) \oplus \text{Range}(Q)$$

$$\text{Ker}(P+Q) = \text{Ker}(P) \cap \text{Ker}(Q)$$

$$\implies V = \text{Range}(P) \oplus \text{Range}(Q) \oplus \text{Ker}(P) \cap \text{Ker}(Q)$$

Proof. Assume; $P, Q, P+Q$ are projections

$P, Q, P+Q$ are projections $\implies \text{Range}(P+Q) = \text{Range}(P) + \text{Range}(Q)$

$\implies \text{Range}(P) \subseteq \text{Ker}(Q)$ & $\text{Range}(Q) \subseteq \text{Ker}(P)$

$$\{0\} \subseteq \text{Range}(P) \cap \text{Range}(Q) \subseteq \text{Ker}(Q) \cap \text{Range}(Q) = \{0\}$$

$$\therefore \text{Range}(P+Q) = \text{Range}(P) \oplus \text{Range}(Q)$$

If $z \in \ker(P) \cap \ker(Q)$, then

$$(P+Q)(z) = P(z) + Q(z) = 0 + 0 = 0$$

$$\Rightarrow \ker(P) \cap \ker(Q) \subseteq \ker(P+Q).$$

If $z \in \ker(P+Q)$. Then

$$0 = (P+Q)(z) = P(z) + Q(z)$$

This may happen iff $P(z) = -Q(z)$.

where $P(z) \in \text{Range}(P)$ & $Q(z) \in \text{Range}(Q)$.

$\text{Range}(P) \cap \text{Range}(Q) = \{0\}$.

$$\Longrightarrow P(z) = Q(z) = 0.$$

$$\Longrightarrow z \in \ker(P) \cap \ker(Q)$$

* Let R and N be subspaces of $V = \mathbb{R}^n$ such that $R \oplus N = V$

Let, $\{u_1, \dots, u_r\}$ be a basis for R and $\{v_1, \dots, v_r\}$ be a basis for $N^{\perp}$.

Let, $A = [u_1 \dots u_r]$ and $B = [v_1 \dots v_r]$ (where the vectors are expressed as their coordinates in the standard basis).

Then the matrix (in the standard basis) of the projection on R parallel to N is :

$$P = A(B^T A)^{-1} B^T$$

where, $A = [u_1 \cdots u_r]$ and $B = [v_1 \cdots v_r]$

$R = \text{span}(\{u_i\})$ and $N^{\perp} = \text{span}(\{v_i\})$

& $V = R \oplus N$ where $R = \text{Range}(P)$, $N = \text{Ker}(P)$

Proof Let $x \in V$.

We search for $Px \in R$.

For some scalars $y_1, \ldots, y_r$,

$$Px = \sum_{j=1}^{r} y_j u_j = [u_1 \cdots u_r] \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = Ay$$

$$(I-P)x = x - Px \in N = (N^{\perp})^{\perp} \Longrightarrow x - Px \perp N^{\perp}$$

$$\therefore \quad (v_k, x - Px) = v_k^T (x - Px) = 0 \quad \text{for all } k = 0, 1, \ldots, r.$$

$$\Rightarrow 0 = v_k^T (x - Px) = v_k^T x - v_k^T Px = v_k^T x - v_k^T \sum_{j=1}^{r} y_j u_j$$

$$0 = v_k^T x - \sum_{j=1}^{r} y_j (v_k^T u_j), \quad k = 1, 2, \ldots, r$$

$$\sum_{j=1}^{r} y_j (v_k^T u_j) = v_k^T x$$

$$\Rightarrow \boxed{v_k^T x = \sum_{j=1}^{r} (v_k^T u_j) y_j}$$

$$\sum_{j=1}^{r} (v_k^T u_j) y_j = (v_k^T u_1) y_1 + (v_k^T u_2) y_2 + \cdots + (v_k^T u_r) y_r$$

$$= [v_k^T u_1 \cdots v_k^T u_r] \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = (B^T A y)_k$$

where $v_k^T u_j = (B^T A)_{kj}$

$$V_k^T x = (B^T x)_k$$

$$\Rightarrow \boxed{B^T A y = B^T x}$$

$R \oplus N = V = \mathbb{R}^n$ & $R = \text{Range}(P) = C(P)$ | $N = \text{Ker}(P) = N(P)$

$\text{Range}(P) = R = \text{span}(\{u_i\})$ & $\text{Ker}(P)^{\perp} = N^{\perp} = \text{span}(\{v_i\}) = \text{Range}(P^T)$

$A_{n\times r} = [u_1 \cdots u_r]$ & $B_{n\times r} = [v_1 \cdots v_r]$

Stuck 7/9/2021

If $B^T A$ is invertible iff $N(B^TA) = \{0\}$ ~~(or)~~
$B^T A x = 0$ has only the solution $x = 0$

$$B^T A x = 0 \Rightarrow x \in N(B^T A)$$

$$\searrow$$

$$Ax \in N(B^T)$$

$$Ax = [u_1 \cdots u_r]\begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \vec{u}_1 x_1 + \cdots + \vec{u}_r x_r \, .$$

A has independent columns $\}$ $Ax = 0$ iff $x = 0$

If $x \neq 0$, $Ax \neq 0$ & $Ax \in C(A)$

$A x \in N(B^T)$ & $Ax \in C(A)$
$\in C(B)^{\perp}$

$\Longrightarrow$ $Ax \in C(A) \cap C(B)^{\perp}$

$C(A) = \text{Range}(P) = R$ & $C(B)^{\perp} = \text{Ker}(P) = N$

$R \oplus N = C(A) \oplus C(B)^{\perp} = V = \mathbb{R}^n$

$$\dim\left[C(A) + C(B)^{\perp}\right] = \dim\left[C(A)\right] + \dim\left[C(B)^{\perp}\right] - \dim\left[C(A) \cap C(B)^{\perp}\right]$$

$\text{RHS} \leq r + (n-r) - 1$ [since $Ax \neq 0$

$= n-1 \neq n = \text{LHS}$

$\Longrightarrow$ $x = 0$ is the only solution to $B^T A x = 0$

i.e., $N(B^T A) = \{0\}$.

$\therefore$ $B^T A$ is invertible

$$B^T A y = B^T x \implies y = (B^T A)^{-1} B^T x$$

$\therefore$

$$Px = Ay = A(B^T A)^{-1} B^T x$$

$$\implies P = A(B^T A)^{-1} B^T x$$

Orthogonal Projection

ILA (10)
Pseudo Inverse

* An orthogonal projection is a projection P for which $\text{Range}(P)$ and $\text{Ker}(P)$ are orthogonal subspaces.

$$\text{Range}(P) \perp \text{Ker}(P)$$

P is a projection $\longrightarrow P^2 = P$

Orthogonal projection $\Longrightarrow C(P) \perp N(P)$

$Px \in C(P)$

$(I-P)x = y - Py \in N(P)$

$\forall\, x, y \in \mathbb{R}^n$
$x, y \neq 0.$

$Px \perp y - Py$

$$0 = (Px)^T(y - Py) = x^T P^T (y - Py)$$
$$= x^T P^T (I - P) y$$
$$= x^T (P^T - P^T P) y \quad \forall\, x, y \in \mathbb{R}^n$$

$$\Longrightarrow P^T = P^T P$$

$$\therefore P = (P^T)^T = (P^T P)^T = P^T P = P^T$$

$$\Longrightarrow \underline{P^T = P}$$

⇒ Let V be a finite dimensional inner product space. A linear transformation $P: V \to V$ is an orthogonal projection if and only if $P^2 = P$ and $P = P^T$.

* A

is

satis

Prog

P^2

Ran

For

$\langle P$

$\langle$

* A projection P on a Hilbert space V is called an orthogonal projection if it satisfies :
$$\langle Px, y\rangle = \langle x, Py\rangle \quad \forall \; x, y \in V.$$

Proof

$P^2 = P = P^T$ & $V = \text{Range}(P) \oplus \text{Ker}(P)$

$\text{Range}(P) \perp \text{Ker}(P)$

For any $x, y \in V$,

$$Px \in \text{Range}(P) \;\; \& \;\; y - Py \in \text{Ker}(P)$$

$$\langle Px, y - Py\rangle = 0 = \langle P^2x, y - Py\rangle = \langle Px, P(I-P)y\rangle$$
$$= \langle Px, (P - P^2)y\rangle$$

$$\langle Px, y\rangle = \langle Px, Py\rangle = \langle Py, Px\rangle = \langle Py, x\rangle = \langle x, Py\rangle$$

Projection onto a line comes from a rank one matrix. Projection onto a plane comes from a rank-2 matrix :

Projection matrix onto the z-axis : $P_z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Projection matrix onto xy-plane : $P_{xy} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$p_1 = P_1 b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$$

$$p_2 = P_2 b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

The line (z-axis) and plane (xy-plane) are orthogonal complements.

dimensions $1+2=3$.

∴ Every vector b in the whole space is the sum of its parts in the 2 subspaces. The projections p_1 and p_2 are exactly those 2 parts of b:

The vectors give $p_1 + p_2 = b$

The matrices give $P_1 + P_2 = I$

Every subspace of $\mathbb{R}^m$ has its own $m \times m$ projection matrix. To compute P, we absolutely need a good description of the subspace that it projects onto. The best description of a subspace is a basis. We put the basis vectors into the columns of A. Now,

we are projecting onto the $C(A)$.

Our problem is,

to project any b onto the column space of any $m \times n$ matrix.

Projection matrices

- $P^T = P$ & $P^2 = P$
  (symmetric) (idempotent)

- $\lambda = 0$ (or) 1

  Proof.

  $Pv = \lambda v$, $P^2 v = \lambda^2 v = \lambda v$

  $v \neq 0 \Rightarrow \lambda = 0$ (or) 1

- $\text{trace}(P) = \text{rank}(P)$

  Proof.

  $\text{trace}(P) = \sum \lambda_i$ = total # of non-zero eigenvalues that are 1.

  $\dim(C(A^T)) + \dim(N(A)) = r + (n-r) = n$

  $N(A)$ is all vectors of the form, $Av = 0 = 0v$

  $N(A)$ is the eigenspace corresp. to eigenvalue 0.

  $\text{rank}(A) = \dim(C(A^T))$ = total # of non-zero eigenvalues that are 1 $= \text{trace}(P)$

  iff A is diagonalizable

- $P^T = P$ and $P^2 = P$

  $\Rightarrow a_{ii} = |a_i|^2$

<u>**Proof**</u>

$P^T P = P^2 = P$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} |a_1|^2 & & & \\ & |a_2|^2 & & \\ & & \ddots & \\ & & & |a_n|^2 \end{bmatrix} = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

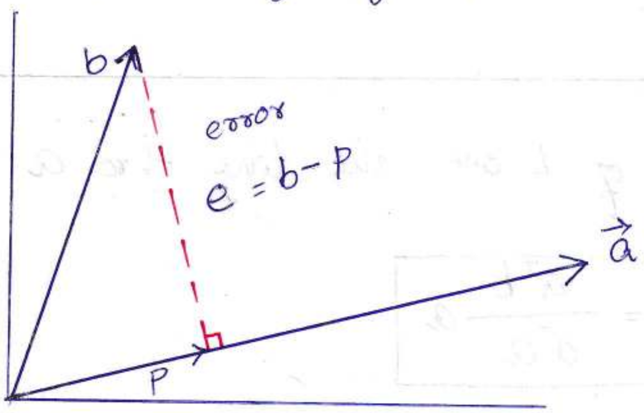
$= P.$

- $P_c A P_R = A$

4.2(30)

Projection Onto a Line

A line goes thro' the origin in the direction of $a = (a_1, \ldots, a_m)$. Along, that line we want a point p closest to $b = (b_1, \ldots, b_m)$. The key to projection is orthogonality: The line from b to p is $\perp$ to the vector a. This is $e = b - p$ for the error on the left side of figure. We now compute p by algebra..



The projection p will be some multiple of a, say $p = \hat{x} a$

find $\hat{x}$, then find the vector p, then find the matrix P.

Projecting b onto a with error, $e = b - p = b - \hat{x}a$

$$a \cdot e = a \cdot (b - \hat{x}a) = a \cdot b - \hat{x}\, a \cdot a = 0$$

$$\longrightarrow \boxed{\hat{x} = \frac{a \cdot b}{a \cdot a} = \frac{a^T b}{a^T a}}$$

The projection of b onto the line thro' a is the vector,

$$\boxed{p = \hat{x}a = \frac{a^T b}{a^T a}\, a}$$

The projection matrix,

$$p = a\hat{x} = a\,\frac{a^T b}{a^T a} = Pb \implies \boxed{P = \frac{aa^T}{a^T a}}$$

$$P = \frac{aa^T}{a^T a} = \frac{a \otimes a}{a \cdot a} = \frac{1}{|a|^2} \begin{bmatrix} \\ \\ \end{bmatrix} \begin{bmatrix} & & \end{bmatrix} \implies \text{rank} = 1$$

∴ • The projection matrix P is $m \times m$, but its rank is one. , $P = \hat{a}\hat{a}^T$

ILA(3)

• We are projecting onto a 1D subspace, the line thro' 'a'. That line is the $C(P)$, column space of P.

- If the vector 'a' is doubled, the matrix P stays the same. It still projects onto the same line

- If the matrix is squared,

  $P^2 = P \iff$ Projecting a 2^{nd} time doesn't change anything.

- When P projects onto one subspace, $I-P$ projects onto the $\perp$ subspace

- For P,

  $$\sum a_{ii} = 1$$

Ex:1 Project $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto $a = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ to find

$p = \hat{x} a$

Ans: $\hat{x} = \dfrac{a^T b}{a^T a}$

$a^T b = \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 5$

$a^T a = 9.$

So,
the projection, $p = \dfrac{5}{9} a = \left(\dfrac{5}{9}, \dfrac{10}{9}, \dfrac{10}{9}\right)$

The error vector b/w b and p is,

$e = b - p = \left(\dfrac{4}{9}, \dfrac{-1}{9}, \dfrac{-1}{9}\right)$

Ex: 2. Find the projection matrix $P = \frac{a a^T}{a^T a}$ onto the line thro, $a = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

Ans: $P = \frac{a a^T}{a^T a} = \cancel{\frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \end{bmatrix}}\ \frac{1}{9}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 2 \end{bmatrix}$

$$= \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}$$

$$p = Pb = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix}$$

Projection onto a subspace

Start with 'n' vectors $a_1, a_2, \dots, a_n$ in $\mathbb{R}^m$.
Assume that these a's are linearly independent.

Problem: Find the combination $p = \hat{x}_1 a_1 + \dots + \hat{x}_n a_n$ closest to a given vector b.

We are projecting each b in $\mathbb{R}^m$ onto the subspace spanned by the a's., i.e., $C(A)$

$A_{m\times n}$

The combinations in $\mathbb{R}^m$ are the vectors Ax in the column space.
The matrix $A_{m\times n}$ has n columns $a_1, \dots, a_n$.
We are looking for $\Rightarrow$ the particular combination $p = A\hat{x}$ (the projection) that is closest to b.

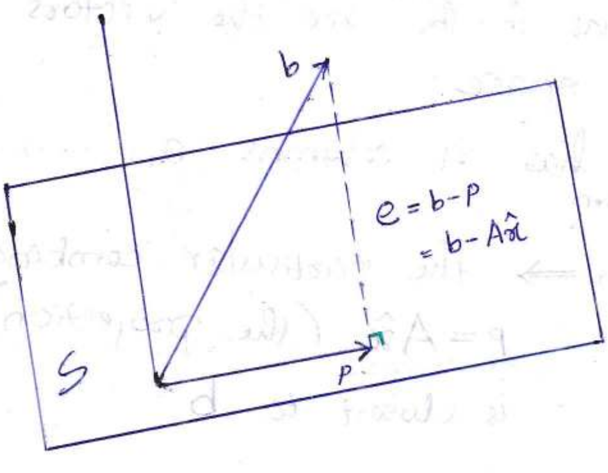
'^' over $\hat{x}$ indicates the best choice $\hat{x}$, to give the closest vector in the column space.

When $n=1$, $\hat{x} = \dfrac{a^T b}{a^T a}$

For $n>1$,

$$\hat{x} = (\hat{x}_1, \cdots, \hat{x}_n) \quad ?$$

Projections onto n-D subspaces : Find the vector $\hat{x}$, find the projection $p = A\hat{x}$, find the projection matrix P.



The projection p of b onto $S = C(A)$

The error vector $e = b - A\hat{x}$ is $\perp$ to the subspace S.

$$e = b - A\hat{x} \quad \perp \quad S = C(A) = x_1 a_1 + \cdots + x_n a_n$$

$\Longrightarrow$ $e = b - A\hat{x}$ makes a right angle with all the vectors $a_1, \ldots, a_n$ in the base.

The 'n' right angles give the 'n' equations for $\hat{x}$:

$$a_1^T (b - A\hat{x}) = 0$$

$$\vdots$$

$$a_n^T (b - A\hat{x}) = 0$$

(or) $$\underbrace{\begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix}}_{A^T} \begin{bmatrix} b - A\hat{x} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

(or)

$$A^T (b - A\hat{x}) = 0$$

$$A^T (b - A\hat{x}) = 0 \Longrightarrow A^T A \hat{x} = A^T b$$

The combination $p = \hat{x}_1 a_1 + \cdots + \hat{x}_n a_n = A\hat{x}$ that is closest to b comes from $\hat{x}$:

$A_{m \times n}$

① Find $\hat{x}_{n \times 1}$

$$A^T(b - A\hat{x}) = 0 \quad \text{(or)} \quad A^T A \hat{x} = A^T b$$

The symmetric matrix $A^T A$ is $n \times n$.

$A^T A$ is invertible if the a's are independent.

The solution is,

$$\boxed{\hat{x} = \underbrace{(A^T A)^{-1} A^T}_{A^+} b}$$

rank $(A^T A)$ = rank (A)

3.3(19)

② Find $P_{m \times 1}$

$$\boxed{p = A\hat{x} = A(A^T A)^{-1} A^T b = Pb}$$

③ Find $P_{m\times m}$, the projection matrix that is multiplying b, to project it onto the subspace spanned by the column vectors of A. (or) onto $C(A)$

$$\boxed{P = A(A^TA)^{-1}A^T}$$

- Our subspace is $C(A)$
- The error vector, $e = b - A\hat{x} \perp C(A)$
- $\Downarrow$

  $e = b - A\hat{x}$ is in the null space of A^T (left nullspace of A) $\Rightarrow A^T(b - A\hat{x}) = 0$

  $C(A) \perp N(A^T)$

$\therefore$

The vector b is being split into the projection p and the error $e = b - p$.

$$\Longrightarrow \boxed{b = p + e}$$

- P projects onto $C(A)$, whereas $I - P$ projects onto $N(A^T)$.

- $P^2 = A(A^TA)^{-1}A^T \cdot A(A^TA)^{-1}A^T$

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

$$= A(A^TA)^{-1}A^{-1} = P.$$

Warning :

$$P = A(A^TA)^{-1}A^T = A\,A^{-1}(A^T)^{-1}A^T = I\,I = I \quad !$$

$(A^TA)^{-1} \neq A^{-1}(A^T)^{-1}$ since

The matrix A is rectangular. It has no inverse matrix.

* A^TA is invertible iff A has linearly independent columns.

Proof

Let A be any matrix. If x is in its nullspace. i.e., $x \in N(A)$. Then $Ax = 0$

$\implies A^TAx = 0 \implies x \in N(A^TA)$.

If $x \in N(A^T A)$,

$$A^T A x = 0$$

We can't multiply by $(A^T)^{-1}$, which generally doesn't exist.

$$x^T A^T A x = (Ax)^T Ax = (Ax) \cdot (Ax) = |Ax|^2 = 0$$

$$\Longrightarrow Ax = 0 \longrightarrow x \in N(A)$$

$\rightarrow N(A^T A) = N(A)$

If $A^T A$ has dependent columns, so has A.
If $A^T A$ has independent columns, so has A.

When A has independent columns,
$N(A) = Z \Longrightarrow N(A^T A) = Z$
$\Rightarrow$ $A^T A$ has independent columns.,
$A^T A$ is square, symmetric, & invertible

(OR)

(OR) If $A^T A x = 0$,

$Ax \in N(A^T)$ & $Ax \in C(A)$

$C(A) \perp N(A^T) \implies \cancel{A\mathcal{C}}\ \underline{Ax = 0}$

$\boxed{4.2(A)}$ Project the vector $b = (3,4,4)$ onto the line thro' $a = (2,2,1)$ and then onto the plane that also contains $a^* = (1,0,0)$.

Check the error vector $e = b - p \perp a$? & the 2nd vector $e^* = b - p^*$ is also $\perp a^*$?

Find the 3×3 projection matrix P onto that plane of a and a^*.

Find a vector whose projection onto the plane is the zero vector. Why is it exactly the error e^* ?

Ans:

$$p = Pb = \frac{aa^T}{a^Ta}b = \frac{18}{9}\begin{bmatrix}2\\2\\1\end{bmatrix} = 2(2,2,1) = 2a$$

$$e = b - p = (3,4,4) - (4,4,2) = (-1,0,2)$$

$$a^Te = 0 \implies e \perp a$$

$$A = \begin{bmatrix}2 & 1\\2 & 0\\1 & 0\end{bmatrix} \longrightarrow A^TA = \begin{bmatrix}2 & 2 & 1\\1 & 0 & 0\end{bmatrix}\begin{bmatrix}2 & 1\\2 & 0\\1 & 0\end{bmatrix} = \begin{bmatrix}9 & 2\\2 & 1\end{bmatrix}$$

$$(A^TA)^{-1} = \frac{1}{5}\begin{bmatrix}1 & -2\\-2 & 9\end{bmatrix}$$

$$P = A(A^TA)^{-1}A^T = \frac{1}{5}\begin{bmatrix}2 & 1\\2 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}1 & -2\\-2 & 9\end{bmatrix}\begin{bmatrix}2 & 2 & 1\\1 & 0 & 0\end{bmatrix}$$

$$= \frac{1}{5}\begin{bmatrix}2 & 1\\2 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}0 & 2 & 1\\5 & -4 & -2\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 0.8 & 0.4\\0 & 0.4 & 0.2\end{bmatrix}$$

$p^* = Pb = (3, 4.8, 2.4)$

$e^* = b - p^* = (0, -0.8, 1.6)$. $\perp$ to a & a^*

$\therefore$ e^* is in the

$Pe^* = Pb - P^2 b = Pb - Pb = 0$

$\Longrightarrow e^* \in N(P)$ & its projection is zero.

$$\boxed{P^2 = P = P^T}$$

$\boxed{4.2(B)}$ Suppose,
your pulse is measured at $x = 70$ beats/minute, then at $x = 80$, then at $x = 120$. Those 3 equations in one unknown have $A^T = [1\ 1\ 1]$ and $Ax = b$

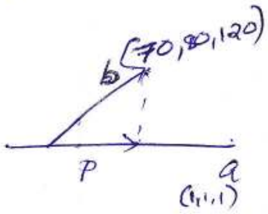
$b = (70, 80, 120)$. The best $\hat{x}$ is the ______ of $70, 80, 120$. Use calculus & projection:

① Minimize $E = (x-70)^2 + (x-80)^2 + (x-120)^2$ by solving $\frac{dE}{dx} = 0$

② Project $b = (70, 80, 120)$ onto $a = (1, 1, 1)$ to find $\hat{x} = \frac{a^T b}{a^T a}$

Ans: The closest horizontal line to the heights $70, 80, 120$ is the average $\hat{x} = 90$.

$b = (70, 80, 120)$; $C(A) = \text{span}(1,1,1)$



(a) $E = \underset{(70,80,120) \to (x,y,z)}{\triangle} = (x-70)^2 + (x-80)^2 + (x-120)^2$.

$a = (x, x, x)$

$$\frac{dE}{dx} = 2(x-70) + 2(x-80) + 2(x-120) = 0.$$

$$\Longrightarrow 3x - 270 = 0 \Longrightarrow \hat{x} = 90$$

(b) $$\hat{x} = \frac{a^T b}{a^T a} = \frac{\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 70 \\ 80 \\ 120 \end{bmatrix}}{\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} = \frac{70+80+120}{3} = 90$$ //

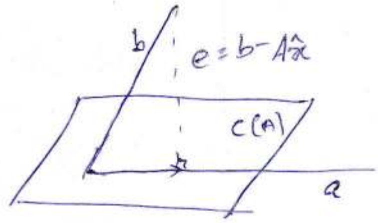
* In recursive least squares,

a 4^{th} measurement 130 changes the average $\hat{x}_{old} = 90$ to $\hat{x}_{new} = 100$.

$$\hat{x}_{new} = \hat{x}_{old} + \frac{1}{4}(130 - \hat{x}_{old})$$

Least Square Approximations

It often happens that $Ax=b$ has no solution. The usual reason is: There are more equations than unknowns. The matrix A has more rows than columns $(m>n)$. The n-columns span a small part of m-D space. Unless all measurements are perfect, b is outside that column space of A. Elimination reaches an impossible equation and stops. But, we can't stop just because measurements include noise!



We can't always get the error $e=b-Ax$ down to zero.

When $e=0$,
x is an exact solution to $Ax=b$

When the length of e is as small as possible,
$\hat{x}$ is a least square solution

* When $Ax=b$ has no solution, multiply by A^T and solve $A^T A \hat{x} = A^T b$

Ex:1. A crucial application of least squares is fitting a straight line to m points. Start with 3 points: Find the closest line to the points $(0,6), (1,0), (2,0)$

Ans: ~~$y = mx + c$~~

No straight line $\underline{y = mx + c}$ goes thro' these 3 points.

What are the 2 numbers m, c that satisfy these 3 points.

$$\begin{aligned} 6 &= m \cdot 0 + c \\ 0 &= m \cdot 1 + c \\ 0 &= m \cdot 2 + c \end{aligned} \Longrightarrow \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$A x = b$$

This 3×2 system has no solution.

$Ax = b$ is not solvable.

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix}$$

$$\hat{x} = (A^T A)^{-1} A^T b = \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{6} \underset{2\times 3}{\begin{bmatrix} -3 & 0 & 3 \\ 5 & 2 & -1 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

$\hat{x} = (m, c) = (-3, 5)$

These numbers are the best m & c, so

$y = -3x + 5$ will be the best line for the 3 points.

Minimizing the error

By geometry: The nearest point ~~closer~~ to b is the projection p.

$\therefore$ The best choice for $A\hat{x}$ is $p \in C(A)$

The smallest possible error is, $e = b - p$, $\perp$ to the columns.

By algebra: Every vector b splits into 2 parts. The part in $C(A)$ is p. The $\perp$ part (part in $N(A^T)$) is e.

$\boxed{Ax = b = p + e} \rightarrow$ is impossible to solve

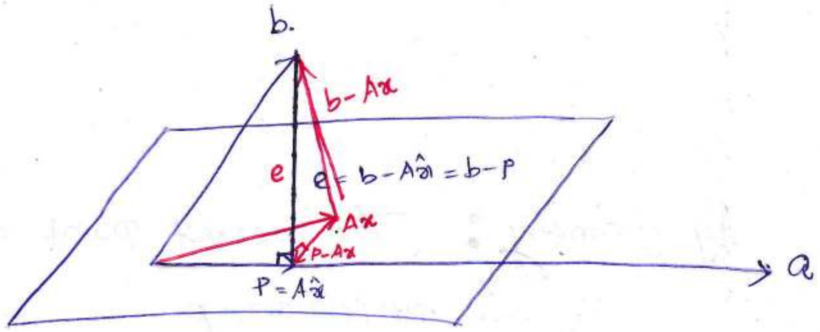
$A\hat{x} = p$ is solvable

$$\Longrightarrow A^T A \hat{x} = A^T p$$

where $\boxed{p = Pb}$

$$\Longrightarrow \hat{x} = (A^T A)^{-1} A^T p$$

→ The solution to $A\hat{x}=p$ leaves the least possible error. (which is e).



Squared length for any x,

$$\boxed{|Ax-b|^2 = |Ax-p|^2 + |e|^2}$$

* The least squares solution $\hat{x}$ makes $E = |Ax - b|^2$ as small as possible.

By calculus **:** The error function E to be minimized is a sum of squares $e_1^2+e_2^2+e_3^2$ (the square of the error in each equation):

$$E=|Ax-b|^2 = e_1^2+e_2^2+e_3^2$$

Ex:-

$$E=|Ax-b|^2=(m\cdot 0+c-6)^2+(m\cdot 1+c)^2+(m\cdot 2+c)^2$$

The unknowns are m & C. With 2 unknowns there are 2 derivatives - both zero at the * minimum. They are "partial derivatives" because $\frac{\partial E}{\partial C}$ treats m as constant and $\frac{\partial E}{\partial m}$ treats C as constant:

$$\cancel{\frac{\partial E}{\partial c} = 2(\ldots)}$$

$$\frac{\partial E}{\partial c} = 2(m\cdot 0 + c - 6) + 2(m\cdot 1 + c) + 2(m\cdot 2 + c) = 0$$

$$\frac{\partial E}{\partial m} = 2(m\cdot 0 + c - 6)(0) + 2(m\cdot 1 + c)(1) + 2(m\cdot 2 + c)(2) = 0$$

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

$$\frac{\partial E}{\partial c} = 0 \Rightarrow 3m + 3c = 6$$

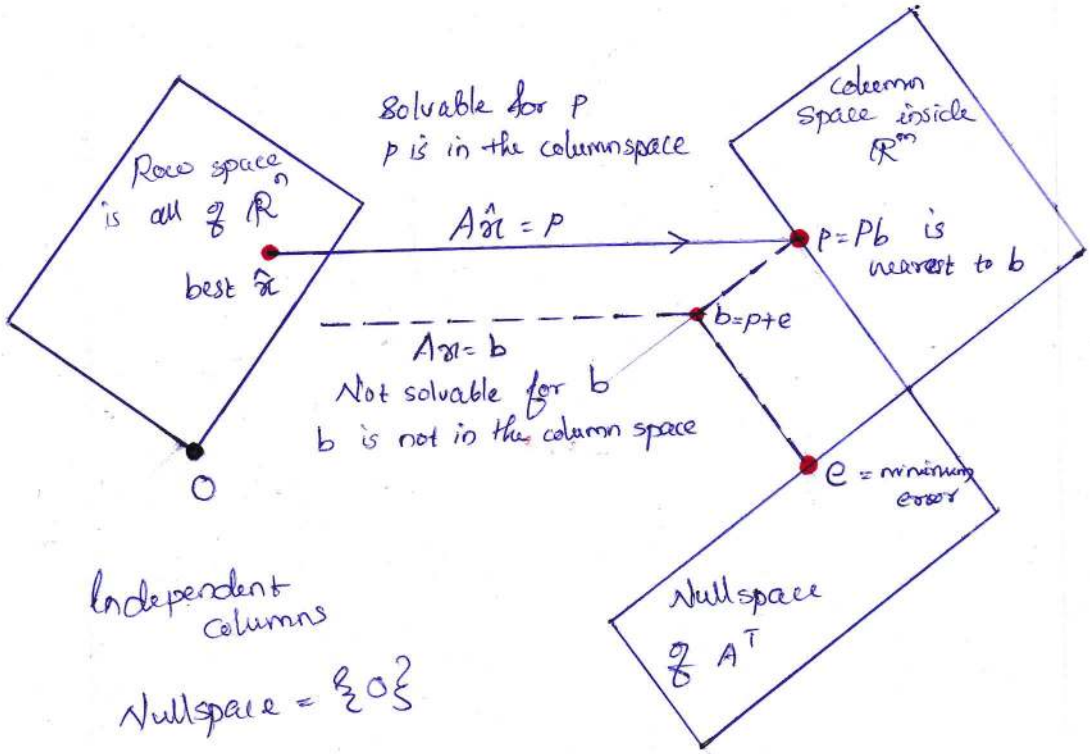
$$\iff \underline{A^T A \hat{x} = A^T b}$$

$$\frac{\partial E}{\partial m} = 0 \Rightarrow 5m + 3c = 0$$

$$A^T A = \begin{bmatrix} 3 & 3 \\ 5 & 3 \end{bmatrix}$$

The best m & c are the components of $\hat{x}$.

* The partial derivatives of $E = |Ax - b|^2$ are zero when $A^T A \hat{x} = A^T b$.



Independent columns

Nullspace = {0}

- The projection $p = A\hat{x}$ is closest to b, so $\hat{x}$ minimizes $E = |b - Ax|^2$

Fitting a Straight Line

- Application of least squares

It starts with $m > 2$ points, hopefully near a straight line.

At times $t_1, \dots, t_m$ those m points are at heights $b_1, \dots, b_m$. The best line $C + Dt$ misses the points by vertical distances $e_1, \dots, e_m$. No line is perfect, and the least square line minimizes, $E = e_1^2 + \dots + e_m^2$.

To fit the m points, we are trying to solve 'm' equations (and we only have 2 unknowns).

$Ax = b$ is

$$\left.\begin{aligned} C + Dt_1 &= b_1 \\ C + Dt_2 &= b_2 \\ &\vdots \\ C + Dt_m &= b_m \end{aligned}\right\} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}$$

$C(A)$ is so thin that almost certainly b is outside of it.

When b happens to lie in the $C(A)$, the points happen to lie on a line. In that case, $b = p$. Then $Ax = b$ is solvable and the errors are $e = (0, \ldots, 0)$.

* The closest line $C+Dt$ has heights $p_1, \ldots, p_m$ with errors $e_1, \ldots, e_m$.

* Solve $A^TA\hat{x} = A^Tb$ for $\hat{x} = (C, D)$. The errors are $e_i = b_i - C - Dt_i$

The 2 columns of A are independent (unless all times t_i are the same). So we turn to least squares and solve $A^TA\hat{x} = A^Tb$.

$$A^TA = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

In a specific problem these numbers are given. The best $\hat{x} = (C, D)$ is $(A^TA)^{-1}A^Tb$.

The line $C+Dt$ minimizes $e_1^2+\cdots+e_m^2 = |Ax-b|^2$ when $A^T A\hat{x} = A^T b$:

$$A^T A\hat{x} = A^T b \iff \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

The vertical errors at the 'm' points on the line are the components of $e = b-p$. This error vector $b - A\hat{x}$ is $\perp$ to the columns of A (geometry). The error is in the nullspace of A^T (linear algebra). The best $\hat{x} = (C, D)$ minimizes the total error E, the sum of squares (calulus):

$$E(x) = |Ax-b|^2 = (C+Dt_1-b_1)^2 + \cdots + (C+Dt_m-b_m)^2$$

Calulus sets the derivatives $\frac{\partial E}{\partial C}$ and $\frac{\partial E}{\partial D}$ to zero, and produces $A^T A\hat{x} = A^T b$.

In general,

- we are fitting m data points by n parameters $x_1, \ldots, x_n$. The matrix A has n columns and $n<m$. The derivatives of $|Ax-b|^2$ give the n equations $A^TA\hat{x} = A^Tb$

- The derivative of a square is linear. This is why the method of least squares is so popular.

- $mC + D\sum t_i = \sum b_i \longrightarrow C + D\hat{t} = \hat{b}$

check 6-3(11)

$\Rightarrow$ The best line goes thro' the center point $(\hat{t}, \hat{b})$

Ex: 2. A has orthogonal columns when the measurement times t_i add to zero.

Gram-Schmidt

Suppose,

$b = (1, 2, 4)$ at times $t = -2, 0, 2$. Those times add to zero.

The columns of A has zero dot product:

Ans

$$C + D(-2) = 1$$
$$C + D(0) = 2 \quad \text{(or)} \quad Ax = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$
$$C + D(2) = 4$$

When the Columns of A are orthogonal, $A^T A$ will be a diagonal matrix

$$A^T A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} |a_1|^2 & 0 & \cdots & 0 \\ 0 & |a_2|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |a_n|^2 \end{bmatrix}$$

$$A^T A \hat{x} = A^T b \longrightarrow \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \quad (4)$$

$$\Longrightarrow C = \frac{7}{3}, \quad D = \frac{3}{4}$$

* Orthogonal columns are so helpful that it is worth shifting the times by subtracting the average time, $\hat{t} = \frac{t_1 + t_2 + \cdots + t_m}{m}$.

If the original times were 1, 3, 5, then $\hat{t} = 3$. The shifted times $T = t - \hat{t} = t - 3$ adds up to zero.

- The sum of the deviations taken from the mean is zero.

$$T_1 = 1-3 = -2, \quad T_2 = 3-3 = 0, \quad T_3 = 5-3 = 2$$

$$A_{new} = \begin{bmatrix} 1 & T_1 \\ 1 & T_2 \\ 1 & T_3 \end{bmatrix}, \quad A_{new}^T A_{new} = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}$$

Now, C & D comes from the easy equation (9)
The best straight line uses $C+DT$ which is
$C + D(t-\hat{t}) = C + D(t-3)$.

→ Nothing wrong with shifting the data points left or right.

3(30)

$$T = t-\hat{t} : \quad C = \frac{b_1 + \cdots + b_m}{m} \quad \text{and} \quad D = \frac{b_1T_1 + \cdots + b_mT_m}{T_1^2 + \cdots + T_m^2}$$

⇒ Make the columns orthogonal in advance.
Then $A_{new}^T A_{new}$ is diagonal and $\hat{x}_{new}$ is easy.

* The best line is $C+DT$ or $C+D(t-\hat{t})$

The time shift that makes A^TA diagonal is an example of the Gram-Schmidt process: orthogonalize the columns of A in advance.

Dependent columns in A : What is $\hat{x}$?

variables: t, b

$C + Dt = b$

Which $\hat{x}$ is best if 'A' has dependent columns?

L: $x_1 + x_2 t = b$

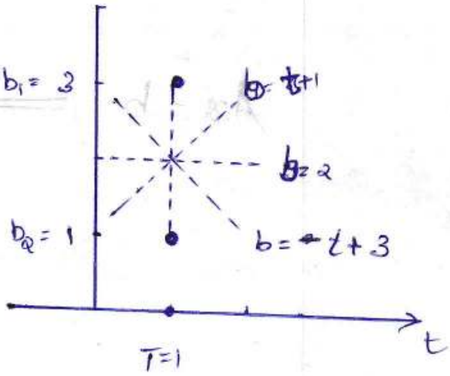
Ex :- points : $(1,3)$, $(1,1)$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = b$$

$$Ax = b$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = P$$

$$A\hat{x} = P$$



The measurements $b_1 = 3$ and $b_2 = 1$ are at the same time $T = 1$.

A straight line $C + Dt$ can not go through both points

Projecting $b = (3,1)$ onto the column space of A

$C(A) = C(1,1)$, ~~Va/Ve~~ $a = (1,1)$

$p = \cancel{Pb = A\hat{x}\, b = \quad 2}$

$$p = Pb = \frac{aa^T}{a^T a} b = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$\longrightarrow$ $\underline{\underline{p = (2,2)}}$

$$Ax \stackrel{?}{=} b \Longrightarrow A\hat{x} = p$$

An equation with no solution has become an equation with infinitely many solutions.

The problem is that A has dependent columns and $(1,-1)$ is in its ~~col~~ null space

$$e = b - p = (1,-1) = (3,1) - (2,2)$$

What solution $\hat{x}$ should we choose ?

$$\hat{x}_1 + \hat{x}_2 = 2 \implies \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \hat{x}_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$(\hat{x}_1, \hat{x}_2) = (2,0)$, $(1,1)$, $(3,-1)$, - - -

$b = 2$ $\qquad$ $b = t+1$ $\qquad$ $b = -t+3$

All these lines (drawn as dashed lines in fig) have the same 2 errors 1 and -1 at time T.
These errors $e = b - p = (1,-1)$ are as small as possible.
But, this doesn't tell us which dash line is best.

In section 7.4,

the "pseudoinverse" of A will choose the shortest solution to $A\hat{x} = p$. Here, the shortest solution will be $x^+ = (1,1)$.

Revisit 7.4

Fitting by a Parabola

If we throw a ball, it would be crazy to fit the path by a straight line.

A parabola $b = C + Dt + Et^2$ allow the ball to go up and come down again (b is the height at time t).

The actual path is not a perfect parabola, but the whole theory of projectiles starts with that approximation.

Problem : Fit heights $b_1, \dots, b_m$ at times $t_1, \dots, t_m$ by a parabola $C + Dt + Et^2$

solution : With $m > 3$, the m equations for an exact fit are generally unsolvable :

$$\left.\begin{array}{c} C + Dt_1 + Et_1^2 = b_1 \\ \vdots \\ C + Dt_m + Et_m^2 = b_m \end{array}\right\} \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & t_m^2 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

where,

$Ax = b$ with $A_{m\times 3} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & t_m^2 \end{bmatrix}$

Least Squares : The closest parabola $C+Dt+Et^2$ chooses $\hat{x} = (C, D, E)$ to satisfy the 3 ? normal equations $A^TA\hat{x} = A^Tb$.

$C(A)$ has dimension 3 in $\mathbb{R}^m$.

The projection of b is $p = A\hat{x}$, which combines the 3 columns using the coefficients C, D, E.

The error at the 1st data point is

$e_1 = b_1 - C - Dt_1 - Et_1^2$.

The total squared error is, $E = e_1^2 + e_2^2 + \cdots e_m^2$.

If you prefer to minimize by calculus, take the partial derivatives of E w.r.t C, D, E. These 3 derivatives will be zero when $\hat{x} = (C, D, E)$ solves the 3×3 system of equations $A^T A \hat{x} = A^T b$.

* The closest parabola $C + Dt + Et^2$ has heights $p_1, p_2, \dots, p_m$ with errors $e_1, \dots, e_m$.

* Solve $A^T A \hat{x} = A^T b$ for $\hat{x} = (C, D, E)$. The errors are $e_i = b_i - C - Dt_i - Et_i^2$.

Section 10.5,

Fourier Series – approximating functions instead of vectors.

The function to be minimized changes from a sum of squared errors $e_1^2+\cdots+e_m^2$ to an integral of the squared error.

Ex:3. For a parabola $b = C + Dt + Et^2$ to go through the 3 heights $b = 6, 0, 0$ with $t = 0, 1, 2$, the

Ans: eq's for $C, D\ E$ are:

$$\left.\begin{aligned} C + D\cdot 0 + E\cdot 0^2 &= 6 \\ C + D\cdot 1 + E\cdot 1^2 &= 0 \\ C + D\cdot 2 + E\cdot 2^2 &= 0 \end{aligned}\right\} Ax = b$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$|A| = 4 - 2 = 2 \implies \begin{bmatrix} C \\ D \\ E \end{bmatrix} = A^{-1} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$C = 6$, $\quad \left.\begin{aligned} 6 + D + E &= 0 \\ 6 + 2D + 4E &= 0 \end{aligned}\right\}$ $\quad \begin{aligned} 2D + 4E &= -6 \\ 2E + 2E &= -12 \\ \hline 2E &= 6 \Rightarrow E = 3 \\ D &= -9. \end{aligned}$

$(C, D, E) = (6, -9, 3)$

The parabola thro' the 3 points is $b = 6 - 9t + 3t^2$

The $C(A) = \mathbb{R}^3$.

Projection matrix, $P = I$, The projection of b is b.

The error is zero.

We didn't need $A^T A \hat{x} = A^T b$ because we solved $Ax = b$

4.3(A) Start with 9 measurements b_1 to b_9, all zero, at times $t = 1, \cdots, 9$. The 10^{th} measurement $b_{10} = 40$ is an outlier. Find the best horizontal line $y = C$ to fit the 10 points $(1,0), (2,0) \cdots (9,0), (10,40)$ using 3 options for the error E:

(a) Least squares, $E_2 = e_1^2 + \cdots + e_{10}^2$

Ans:

$C + Dt = b$

$$\begin{aligned} C + D &= 0 \\ C + 2D &= 0 \\ &\vdots \\ C + 9D &= 0 \\ C + 10D &= 40 \end{aligned} \Longrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 9 \\ 1 & 10 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 40 \end{bmatrix}$$

$A x = b.$

$$A^T A = \underset{2 \times 10}{\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & & 9 & 10 \end{bmatrix}} \underset{10 \times 2}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 9 \\ 1 & 10 \end{bmatrix}} = \begin{bmatrix} 10 & 55 \\ 55 & 385 \end{bmatrix}$$

$$\frac{n(n+1)(2n+1)}{6} \qquad \frac{\overset{5}{\cancel{10}} \times 11 \times \overset{7}{\cancel{21}}}{\cancel{6}} = \frac{770}{2}$$

$$A^T b = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix} = \begin{bmatrix} 40 \\ 400 \end{bmatrix}$$

$$A^T A \hat{x} = A^T b \Longrightarrow C = -8, \; D = \frac{24}{11}$$

$\underline{y = C}$

$$\left.\begin{matrix} C = 0 \\ \vdots \\ C = 40 \end{matrix}\right\} \quad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} C = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 40 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 10$$

$$\hat{x} = (A^T A)^{-1} A^T b$$

$$C = \frac{1}{10} * \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 40 \end{bmatrix} = \frac{1}{10} * 40 = 4$$

- The least square fit to $0, 0, \ldots, 40$ by a horizontal line is $C = 4$

Best horizontal Line: $y = 4$

(b) Least max. error, $E_\infty = |e_{max}|$

Qu:

?

(c) Least sum of errors, $E_1 = |e_1| + \cdots + |e_{10}|$

Ans: $|e_1| + \cdots |e_{10}| = 9|c| + |c-40|$

- Sum of deviation is minimized when taken from the median

$\therefore$ $c = 0$

$\boxed{4.3(B)}$ Find the parabola $C+Dt+Et^2$ that comes closest (least squares error) to the values $b=(0,0,1,0,0)$ at the times $t=-2,-1,0,1,2$.

Ans:

$$\left.\begin{aligned} C+D(-2)+E(-2)^2 &= 0 \\ C+D(-1)+E(-1)^2 &= 0 \\ C+D(0)+E(0)^2 &= 1 \\ C+D(1)+E(1)^2 &= 0 \\ C+D(2)+E(2)^2 &= 0 \end{aligned}\right\} \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$Ax = b$.

$$A^T A = \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix} \Rightarrow$$ col (2) of A is orthogonal to col(1) & col(3).

A: rectangular Vandermonde matrix.

$$A^T A\hat{x} = A^T b \implies \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$C = \frac{34}{70}, \quad D = 0, \quad E = \frac{-10}{70}.$$