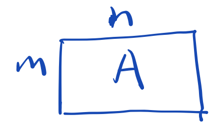
Review of 4 subspaces of $A \in \mathbb{R}^{m \times n}$



① Column Space of A : $Col(A) \subseteq \mathbb{R}^m$

② Row Space of A : $Col(A^T) \subseteq \mathbb{R}^n$

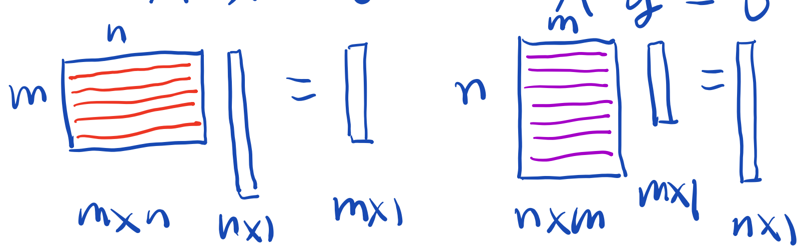
③ Null Space of A : $Null(A) \subseteq \mathbb{R}^n$

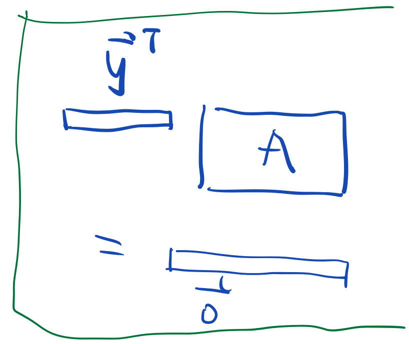
④ Left Null Space of A : $Null(A^T) \subseteq \mathbb{R}^m$

$$Null(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$$

$$Null(A^T) = \{\vec{y} \in \mathbb{R}^m : A^T\vec{y} = \vec{0}\}$$

$A\vec{x} = \vec{0}$ $\qquad$ $A^T\vec{y} = \vec{0}$





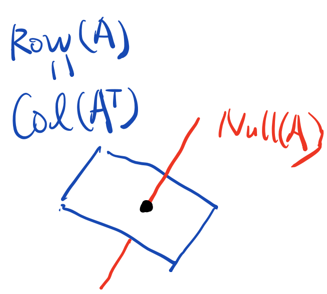
If $A\vec{x} = \vec{0}$, then each row of A is $\perp$ to $\vec{x}$

$\Rightarrow Col(A^T) \perp Null(A)$

If $A^T\vec{y} = \vec{0}$, then each row of A^T is $\perp$ to $\vec{y}$

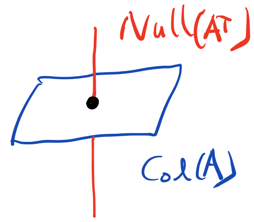
$\Rightarrow$ each col of A is $\perp$ to $\vec{y}$

$\Rightarrow Col(A) \perp Null(A^T)$



Col $(A^T) \perp Null(A)$

Col $(A) \perp Null(A^T)$



① $\dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A)) = \dim(\mathrm{Col}(A^T))$

② $\dim(\mathrm{Null}(A))$ can be different from $\dim(\mathrm{Null}(A^T))$

$$\underset{2\times 3}{\boxed{A^T}}\ \underset{3\times 2}{\boxed{A}} = \underset{2\times 2}{\boxed{A^TA}}$$

If we assume cols of A are independent, then A^TA is invertible

Proof: Want to show $\underline{(A^TA)\vec{x} = \vec{0}}$ has only zero sol

$$A^T\underbrace{A\vec{x}}_{\vec{y}} = \vec{0} \iff \vec{y} = A\vec{x} \text{ is a sol to } A^T\vec{y} = \vec{0}$$

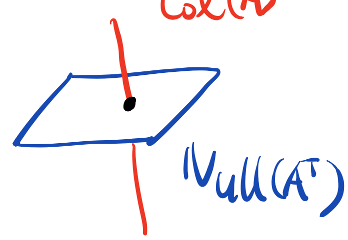
$$\iff A\vec{x} \in \mathrm{Null}(A^T)$$

Since $A\vec{x} \in \mathrm{Col}(A)$, we know

$$A\vec{x} \in \mathrm{Col}(A) \cap \mathrm{Null}(A^T)$$

$\mathrm{Col}(A) \perp \mathrm{Null}(A^T) \Rightarrow$

$$A\vec{x} \perp A\vec{x}$$

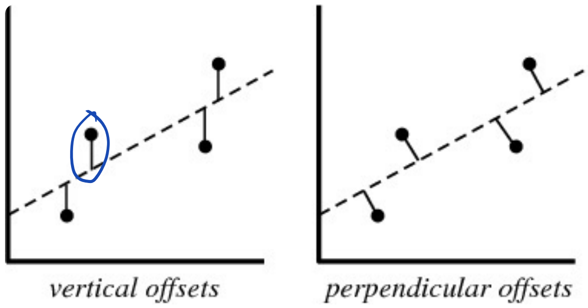


$$\Rightarrow \langle A\vec{x}, A\vec{x}\rangle = 0$$

$$\Rightarrow A\vec{x} = \vec{0}$$

$\left.\begin{array}{l}\text{Independence}\\ \text{of Cols of } A\end{array}\right\} \Rightarrow \vec{x} = \vec{0}$. $\square$

- Line/Curve fitting data points



Given four data points,

(t_1, y_1)
(t_2, y_2)
(t_3, y_3)
(t_4, y_4)

we want a line fitting points.

There are two ways to measure fitting errors

① vertical error: $\sum_{i=1}^{4} |at_i+b-y_i|^2$

This error is defined only if the line is $y=at+b$.

The best line/curve defined by this error is called least square line/curve fitting.

② orthogonal/perpendicular error:

Chapter 7, Principal Component Analysis (PCA).

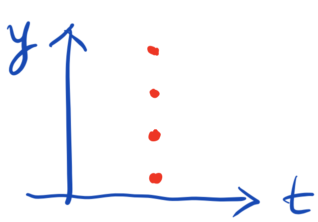
The line by PCA has equation $ay+bt=c$.

③ The best lines by two offsets are different, though very often similar.

1) Least square is used if y depends on t.

2) PCA is used if y and t are not related.

3) Example:



in this case, y cannot be a function of t, PCA should be used.

Chapter 4 is for least squares.

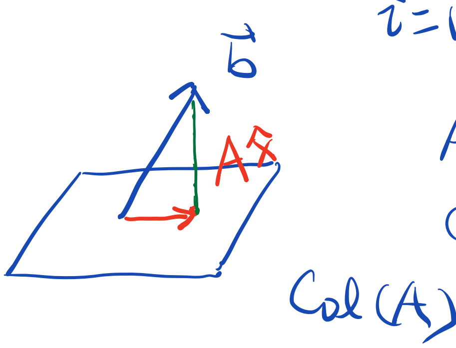
$$\begin{cases} at_1+b = y_1 \\ at_2+b = y_2 \\ at_3+b = y_3 \\ at_4+b = y_4 \end{cases} \Longleftrightarrow \underbrace{\begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ t_3 & 1 \\ t_4 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}}_{\vec{b}}$$

- $A \in \mathbb{R}^{m\times n}$, $m > n \Rightarrow Col(A) \neq \mathbb{R}^m$
- For arbitrary $\vec{b} \in \mathbb{R}^m$, $\vec{b}$ may NOT belong to $Col(A)$, thus no sol to $A\vec{x} = \vec{b}$
- Instead, we seek $\hat{x}$ s.t.

$$\|A\hat{x} - \vec{b}\|^2 = \min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{b}\|^2$$

① $A\vec{x} = \begin{pmatrix} at_1+b \\ at_2+b \\ at_3+b \\ at_4+b \end{pmatrix}$, $\vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$

$$\|A\vec{x} - \vec{b}\|^2 = \sum_{i=1}^{4} |at_i + b - y_i|^2$$

② 

Any vector in $Col(A)$ can be written as $A\vec{x}$ for some $\vec{x} \in \mathbb{R}^n$.

③ Intuition: Shortest distance $\Leftrightarrow$ Orthogonality

④ Orthogonality $\Rightarrow$ $\underbrace{A^TA}\hat{x} = A^T\vec{b}$

$$\Rightarrow \hat{x} = \underbrace{(A^TA)^{-1}A^T\vec{b}}$$

1) least square sol to $A\vec{x} = \vec{b}$ is

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

2) Projection of $\vec{b}$ onto Col (A) is

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

Example: Find the least square line

$y = at+b$ for four data points (y_i, t_i)

$$\begin{cases} at_1+b = y_1 \\ at_2+b = y_2 \\ at_3+b = y_3 \\ at_4+b = y_4 \end{cases} \Leftrightarrow \underbrace{\begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ t_3 & 1 \\ t_4 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}}_{\vec{b}}$$

$$A^TA\vec{x} = A^T\vec{b}$$

$$\begin{pmatrix} t_1 & t_2 & t_3 & t_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ t_3 & 1 \\ t_4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

$$\begin{pmatrix} t_1^2+t_2^2+t_3^2+t_4^2 & t_1+t_2+t_3+t_4 \\ t_1+t_2+t_3+t_4 & 4 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} t_1y_1+t_2y_2+t_3y_3+t_4y_4 \\ y_1+y_2+y_3+y_4 \end{pmatrix}$$

Solve the linear system.

Example: Find the best quadratic fit to four data points (y_i, t_i)

$y = at^2+bt+c$

$A \quad \vec{x} = \vec{b}$

$$\begin{aligned} at_1^2 + bt_1 + c &= y_1 \\ at_2^2 + bt_2 + c &= y_2 \\ at_3^2 + bt_3 + c &= y_3 \\ at_4^2 + bt_4 + c &= y_4 \end{aligned} \iff \begin{pmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ t_3^2 & t_3 & 1 \\ t_4^2 & t_4 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

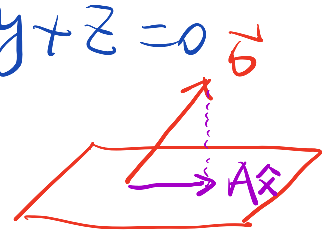
$$A^TA\vec{x} = A^T\vec{b}$$

Solve the linear system.

Example: Find the projection of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ onto the plane $x+y+z=0$.

Sol: 1) Find a basis for $x+y+z=0$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \end{array}\right]$$



$$\left.\begin{matrix} y = s \\ z = t \end{matrix}\right\} \Rightarrow x = -s - t$$

$$\Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ t \end{pmatrix} = \begin{pmatrix} -s \\ s \\ 0 \end{pmatrix} + \begin{pmatrix} -t \\ 0 \\ t \end{pmatrix}$$

$$= s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$\Rightarrow$ The plane is Span$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$

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

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

$$A\vec{x} = \vec{b}$$

$$A^T A \hat{x} = A^T \vec{b}$$

$$\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 2 & 1 & 1 \\ 1 & 2 & 2 \end{array}\right]$$

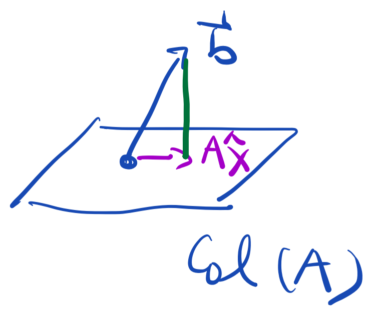
$$\Rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 1/2 \\ 1 & 2 & 2 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 1/2 \\ 0 & 3/2 & 3/2 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 1/2 \\ 0 & 1 & 1 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right]$$

$$\Rightarrow \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$\Rightarrow$ The projection is $A\hat{x}$

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

Verify: $\underbrace{\begin{bmatrix}1\\2\\3\end{bmatrix}}_{\vec{b}} - \underbrace{\begin{bmatrix}-1\\0\\1\end{bmatrix}}_{A\hat{x}} = \begin{bmatrix}2\\2\\2\end{bmatrix}$ $x+y+z=0$

Example: Find projection of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$ onto the hyperplane $\begin{cases} x+y+t=0 \\ x+y+z=0 \end{cases}$

Sol: ① Find a basis for the hyperplane

Solve $\begin{cases} x+y+t+0\cdot z=0 \\ x+y+0\cdot t+z=0 \end{cases}$

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right]$$

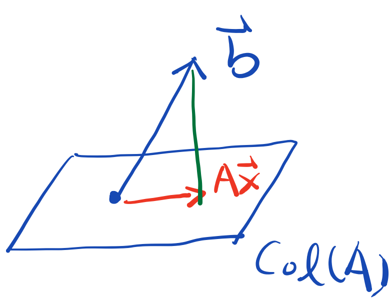
② Form A by putting basis vectors together. A is 4×2

$$\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solve $A^T A \hat{x} = A^T \vec{b}$

The projection is $A\hat{x} = A(A^TA)^{-1}A^T\vec{b}$.

Def: $A(A^TA)^{-1}A^T$ is the projection matrix onto $Col(A)$.



- $\min\limits_{\vec{x}\in\mathbb{R}^n} \|A\vec{x}-\vec{b}\|^2$

$\vec{x}, \vec{y} \in \mathbb{R}^n$ $\quad \vec{x}\cdot\vec{y} = \langle \vec{x}, \vec{y}\rangle = \vec{y}^T\vec{x}$

$$= \min_{\vec{x}\in\mathbb{R}^n} \langle A\vec{x}-\vec{b},\ A\vec{x}-\vec{b}\rangle$$

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

$$(ABC)^T = C^T B^T A^T$$

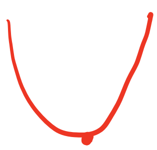
$$= \min_{\vec{x}\in\mathbb{R}^n} (A\vec{x}-\vec{b})^T (A\vec{x}-\vec{b})$$

$$= \min_{\vec{x}\in\mathbb{R}^n} (\vec{x}^T A^T - \vec{b}^T)(A\vec{x}-\vec{b})$$

$$f(\vec{x}) = (\vec{x}^T A^T - \vec{b}^T)(A\vec{x}-\vec{b})$$

$$= \vec{x}^T A^T A\vec{x} - \vec{b}^T A\vec{x} - \vec{x}^T A^T \vec{b} - \vec{b}^T\vec{b}$$

$$\vec{x}^T A^T \vec{b} = (\vec{x}^T A^T \vec{b})^T = \vec{b}^T A\vec{x}$$



$$= \vec{x}^T A^T A\vec{x} - 2\vec{x}^T A^T \vec{b} - \vec{b}^T\vec{b}$$

$$= \vec{x}^T A^T A\vec{x} - 2\langle A^T\vec{b}, \vec{x}\rangle - \|\vec{b}\|^2$$

$$\nabla f(\vec{x}) = 2A^T A\vec{x} - 2A^T\vec{b}$$

So minizing $f(\vec{x}) \iff \nabla f(\vec{x}) = 0$

$\iff A^T A\vec{x} = A^T\vec{b}$

$\iff$ Orthogonality.

- A set of column vectors $\{\vec{a}_1, \dots, \vec{a}_n\}$ is called orthogonal if $\vec{a}_i \perp \vec{a}_j$ for any i, j.

$$\langle \vec{a}_i, \vec{a}_j \rangle = 0$$

$$\vec{a}_j^T \vec{a}_i = 0$$

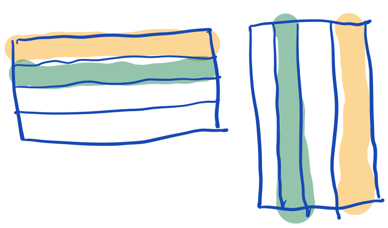
Ex: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$

- A vector $\vec{u}$ is unit if $\|\vec{u}\| = 1$.

- A set of column vectors $\{\vec{a}_1, \dots, \vec{a}_n\}$ is called orthonormal if $\begin{cases} \vec{a}_i \perp \vec{a}_j \text{ for any } i, j. \\ \|\vec{a}_i\| = 1, \ \forall i \end{cases}$

$$\vec{a}_j^T \vec{a}_i = \begin{cases} 0 & , i \neq j \\ 1 & , i = j \end{cases}.$$

- Let $\vec{a}_1, \dots, \vec{a}_n$ be cols of $A \in \mathbb{R}^{m \times n}$, assume they are orthonormal, then

$A^T A =$  $= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

In this case, the projection matrix is

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

Ex: $A = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ $\qquad A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$