Lecture 14

Projection and least square

Projection

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

which we call it the inner product of $\vec{x}, \vec{y}$

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Which we call it the inner product of $\vec{x}, \vec{y}$

Properties:

- $\|\vec{x}\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x} - y\|^2.$

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Which we call it the inner product of $\vec{x}, \vec{y}$

Properties:

- $\|\vec{x}\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x} - y\|^2$.
- $\cos\theta = \dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}$ if $\vec{x}, \vec{y} \neq 0$.

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

which we call it the inner product of $\vec{x}, \vec{y}$

Properties:

- $\|\vec{x}\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x} - y\|^2.$
- $\cos\theta = \dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|}$ if $\vec{x}, \vec{y} \neq 0.$

Def: $\vec{x} \perp \vec{y}$ if $\langle \vec{x}, \vec{y} \rangle = 0.$

Projection

Recall: $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T y$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

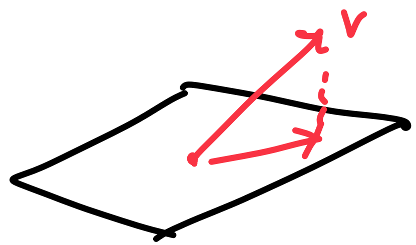
Which we call it the inner product of $\vec{x}, \vec{y}$

Properties:

- $\|\vec{x}\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x} - \vec{y}\|^2.$
- $\cos\theta = \dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}$ if $\vec{x}, \vec{y} \neq 0.$

Def: $\vec{x} \perp \vec{y}$ if $\langle \vec{x}, \vec{y} \rangle = 0.$

Q: $V \leq$ subspace, How we write down the orthogonal projection onto V?



Projection

Projection

Abstractly: $V \subseteq \mathbb{R}^n$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}.$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.

- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{v} = \vec{v}_1 + \vec{v}_2 .$$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.

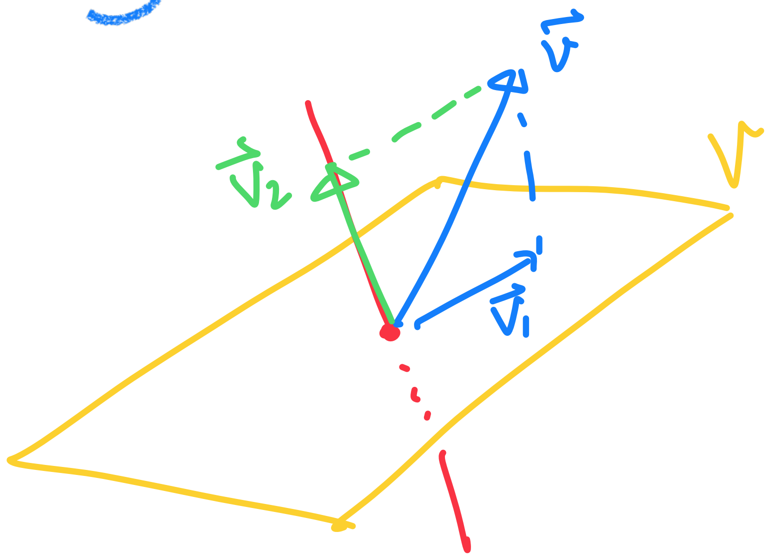
- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{v} = \underset{\substack{\uparrow \\ V}}{\vec{v}_1} + \underset{\substack{\uparrow \\ V^\perp}}{\vec{v}_2}.$$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.



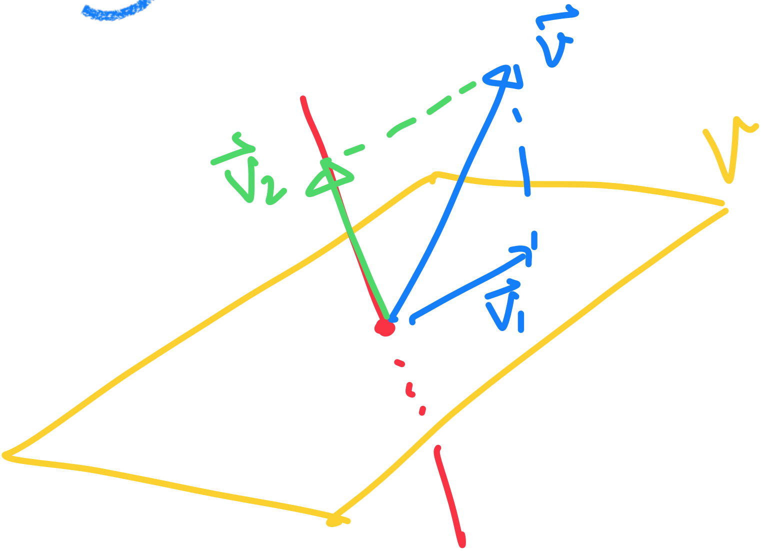
- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\underset{V}{\vec{v} = \vec{v}_1} + \underset{V^\perp}{\vec{v}_2}.$$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.



- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

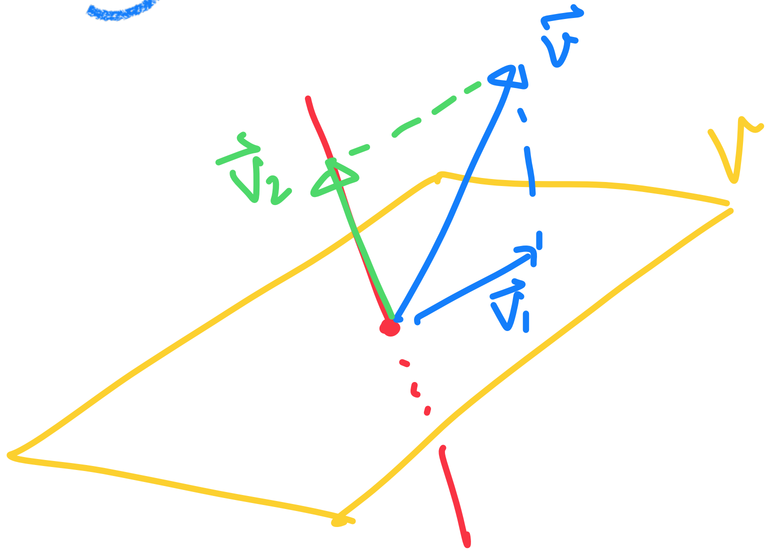
$$\vec{v} = \underset{\in V}{\vec{v}_1} + \underset{\in V^\perp}{\vec{v}_2}.$$

Def: We define the orthogonal projection $P: V \longrightarrow V$. by $P(v) = \vec{v}_1$.

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.



- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{v} = \underset{V}{\vec{v}_1} + \underset{V^\perp}{\vec{v}_2}.$$

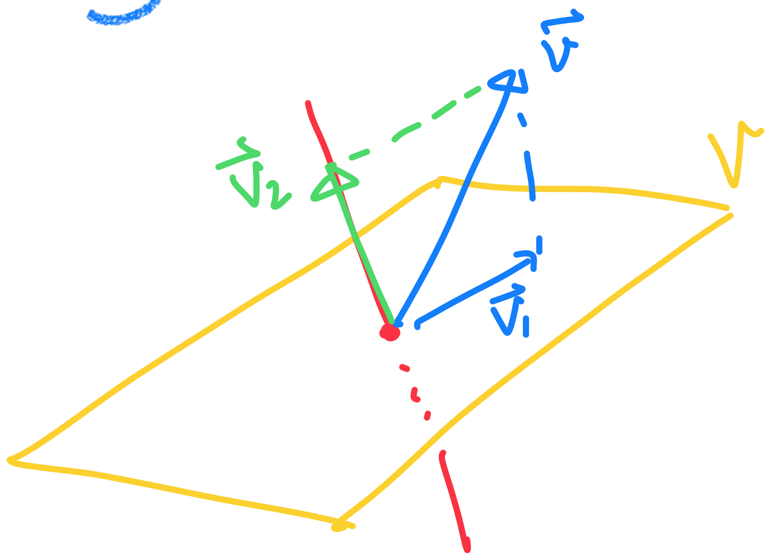
Def: We define the orthogonal projection $P: V \longrightarrow V$. by $P(v) = \vec{v}_1$.

Check: (linearity) Let $v, w \in \mathbb{R}^n$, $v = v_1 + v_2$, $w = w_1 + w_2$.

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^{\perp} = \mathbb{R}^n$

$V \cap V^{\perp} = \{0\}.$



- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\underset{\substack{\uparrow \\ V}}{\vec{v} = \vec{v}_1} + \underset{\substack{\uparrow \\ V^{\perp}.}}{\vec{v}_2}.$$

Def: We define the orthogonal projection $P: V \longrightarrow V.$ by $P(v) = \vec{v}_1.$

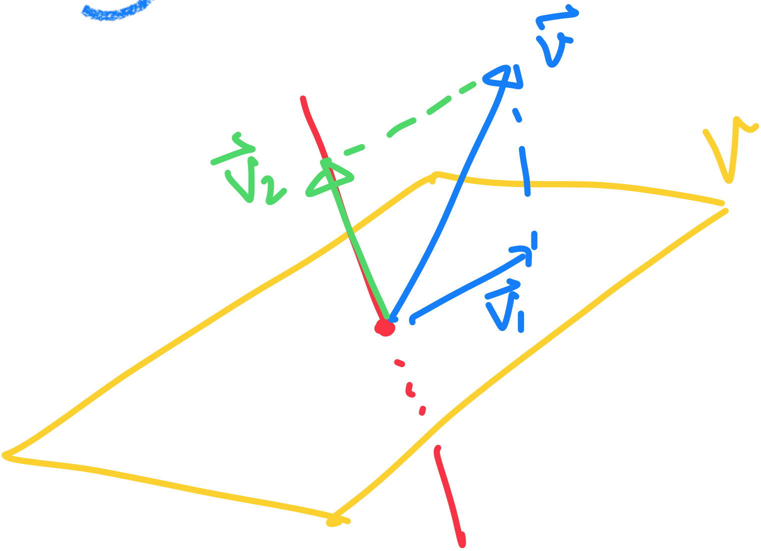
Check: (linearity) Let $v, w \in \mathbb{R}^n$, $v = v_1 + v_2$, $w = w_1 + w_2$.

$cv + dw = (cv_1 + dw_1) + (cv_2 + dw_2) \rightsquigarrow P(cv + dw) = cv_1 + dw_1$

Projection

Abstractly: $V \subseteq \mathbb{R}^n \rightsquigarrow V + V^\perp = \mathbb{R}^n$

$V \cap V^\perp = \{0\}$.



- Any vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{v} = \underset{V}{\vec{v}_1} + \underset{V^\perp}{\vec{v}_2}.$$

Def: We define the orthogonal projection $P : V \longrightarrow V$.

by $P(v) = \vec{v}_1$.

Check: (linearity) Let $v, w \in \mathbb{R}^n$, $v = v_1 + v_2$, $w = w_1 + w_2$.

$$cv + dw = \overbrace{(cv_1 + dw_1)}^{\in V} + \overbrace{(cv_2 + dw_2)}^{\in V^\perp} \rightsquigarrow P(cv+dw) = cv_1 + dw_1$$

Projection

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$.

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V \Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V \Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \ldots, v_k$.

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^{\perp}$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \dots, v_k$.

Step 2: Let $A = \begin{bmatrix} - v_1^T - \\ \vdots \\ - v_k^T - \end{bmatrix}$, $N(A) = V^{\perp}$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \dots, v_k$.

Step 2: Let $A = \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_k^T — \end{bmatrix}$, $N(A) = V^\perp$

find a basis of $V^\perp$; say $v_{k+1}, \dots, v_n$.

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$ $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \dots, v_k$.

Step 2: Let $A = \begin{bmatrix} - v_1^T - \\ \vdots \\ - v_k^T - \end{bmatrix}$, $N(A) = V^\perp$

find a basis of $V^\perp$; say $v_{k+1}, \dots, v_n$.

Step 3: $_{v_i}[P]_{v_i} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V \Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \dots, v_k$.

Step 2: Let $A = \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_k^T — \end{bmatrix}$, $N(A) = V^\perp$

find a basis of $V^\perp$; say $v_{k+1}, \dots, v_n$.

Step 3:

$${}_{v_i}[P]_{v_i} = \begin{bmatrix} \overbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}^{k} & 0 \\ 0 & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}\end{bmatrix}$$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \ldots, v_k$.

Step 2: Let $A = \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_k^T — \end{bmatrix}$, $N(A) = V^\perp$

find a basis of $V^\perp$; say $v_{k+1}, \ldots, v_n$.

Step 3:

$$ {}_{v_i}[P]_{v_i} = \begin{bmatrix} \overbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}^{k} & 0 \\ 0 & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{n-k} \end{bmatrix} $$

Projection

Properties: 1. $P^2 = P$ as linear transformation.

Reason: $P(v) = v_1 \in V$. $\Rightarrow P(P(v)) = P(v_1) = v_1$.

2. $\tilde{P}(v) = v - P(v)$ gives projection onto $V^\perp$

How to write down matrix w.r.t. basis e_i's?

Step 1: Find a basis of V, say $v_1, \dots, v_k$.

Step 2: Let $A = \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_k^T — \end{bmatrix}$, $N(A) = V^\perp$

find a basis of $V^\perp$; say $v_{k+1}, \dots, v_n$.

Step 3: $${}_{v_i}[P]_{v_i} = \begin{bmatrix} \overbrace{1 \; \ddots \; 1}^{k} & 0 \\ 0 & \underbrace{0 \; \ddots \; 0}_{n-k} \end{bmatrix}$$

Step 4: Transition matrix ${}_{e_i}[id]_{v_i}$

Projection

Projection

We want more geometric way of writing it :

Projection

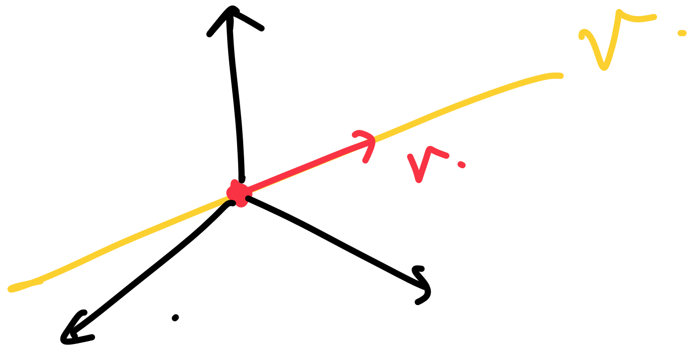
We want more geometric way of writing it:

<u>Easy case:</u> $\dim V = 1$, <u>i.e.</u> $V = \mathrm{span}\{v\}$.

Projection

We want more geometric way of writing it:

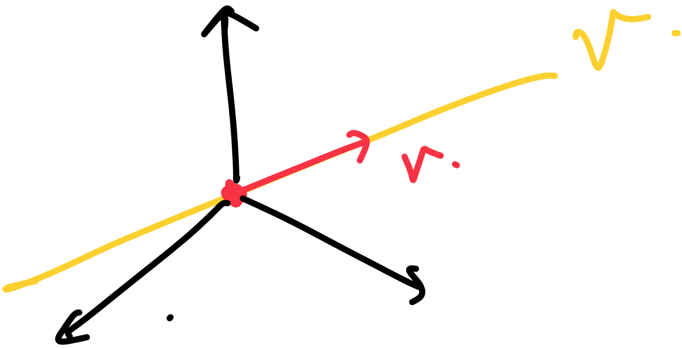
<u>Easy case:</u> $\dim V = 1$, i.e. $V = \mathrm{span}\{v\}$.



Projection

We want more geometric way of writing it:

Easy case: $\dim V = 1$, i.e. $V = \operatorname{span}\{v\}$.



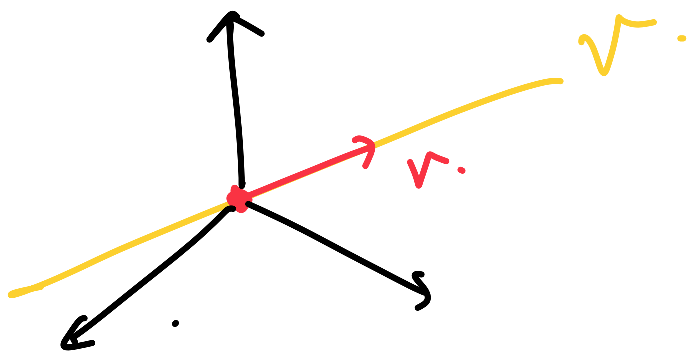
Step 1: normalize v, write

$$\hat{v} = \frac{1}{\|v\|} v$$

Projection

We want more geometric way of writing it:

Easy case: $\dim V = 1$, i.e. $V = \text{span}\{v\}$.



Step 1: normalize v, write

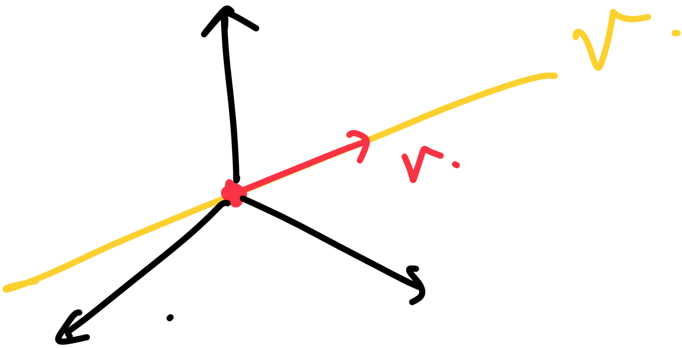
$$\hat{v} = \frac{1}{\|v\|} v$$

$\leftarrow$ non-zero.

Projection

We want more geometric way of writing it :

Easy case : $\dim V = 1$, i.e. $V = \operatorname{span}\{v\}$.



Step 1: normalize v, write

$$\hat{v} = \frac{1}{\|v\|} v$$

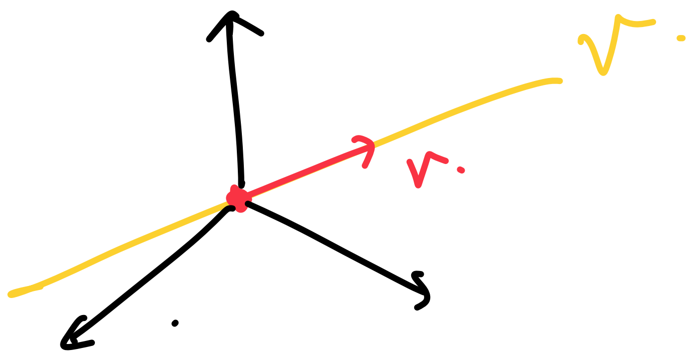
← non-zero.

$$\|\hat{v}\| = \frac{1}{\|v\|}\|v\| = 1.$$

Projection

We want more geometric way of writing it:

Easy case: $\dim V = 1$, i.e. $V = \mathrm{span}\{v\}$.



Step 1: normalize v, write

$$\hat{v} = \frac{1}{\|v\|} v$$

non-zero.

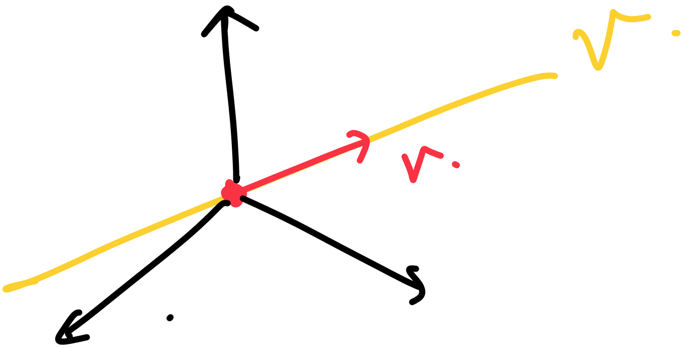
$$\|\hat{v}\| = \frac{1}{\|v\|}\|v\| = 1.$$

- For any other vector x, $P(x) = c \cdot \hat{v}$.

Projection

We want more geometric way of writing it:

Easy case: $\dim V = 1$, i.e. $V = \mathrm{span}\{v\}$.



Step 1: normalize v, write

$$\hat{v} = \frac{1}{\|v\|} v$$ ← non-zero.

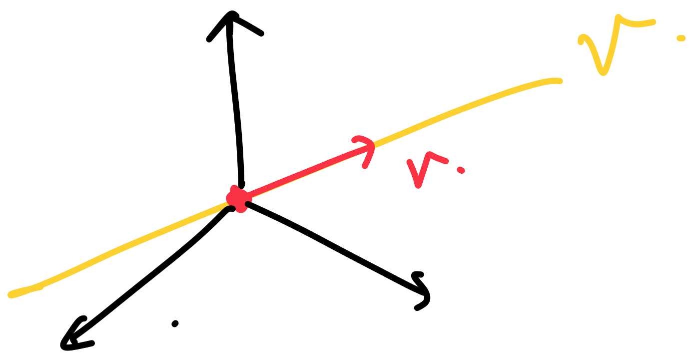
$$\|\hat{v}\| = \frac{1}{\|v\|} \|v\| = 1.$$

- For any other vector x, $P(x) = c \cdot \hat{v}$. ← need to find this.

Projection

We want more geometric way of writing it:

Easy case: $\dim V = 1$, i.e. $V = \text{span}\{v\}$.



Step 1: normalize v, write

$$\hat{v} = \frac{1}{\|v\|} v$$

non-zero.

$$\|\hat{v}\| = \frac{1}{\|v\|} \|v\| = 1.$$

- For any other vector x, $P(x) = c \cdot \hat{v}$.

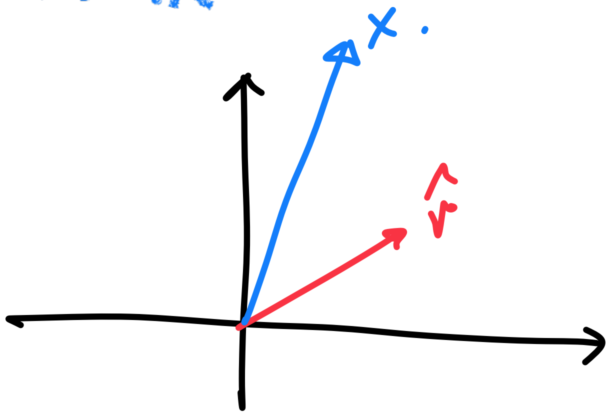
need to find this.

- Guess a formula for c

Projection

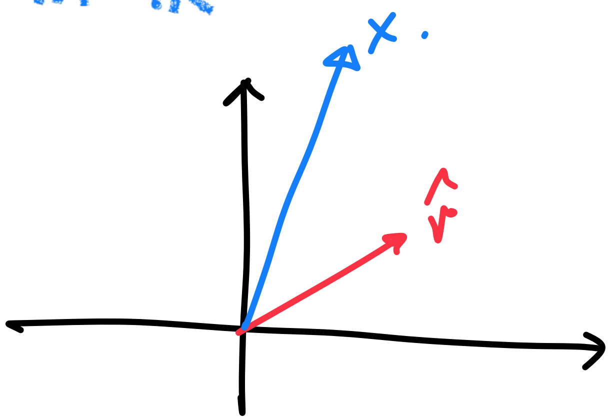
Projection

Say in $\mathbb{R}^2$:



Projection

Say in $\mathbb{R}^2$:

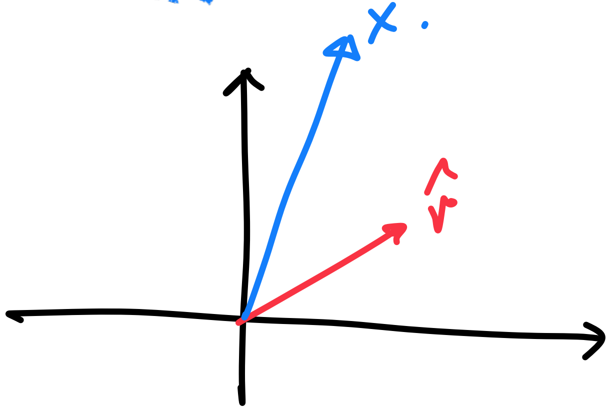


rotation

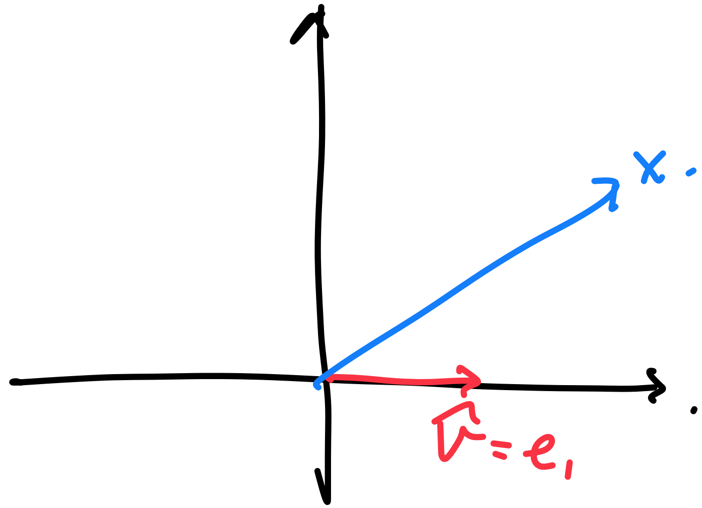
+ reflection.

Projection

Say in $\mathbb{R}^2$:

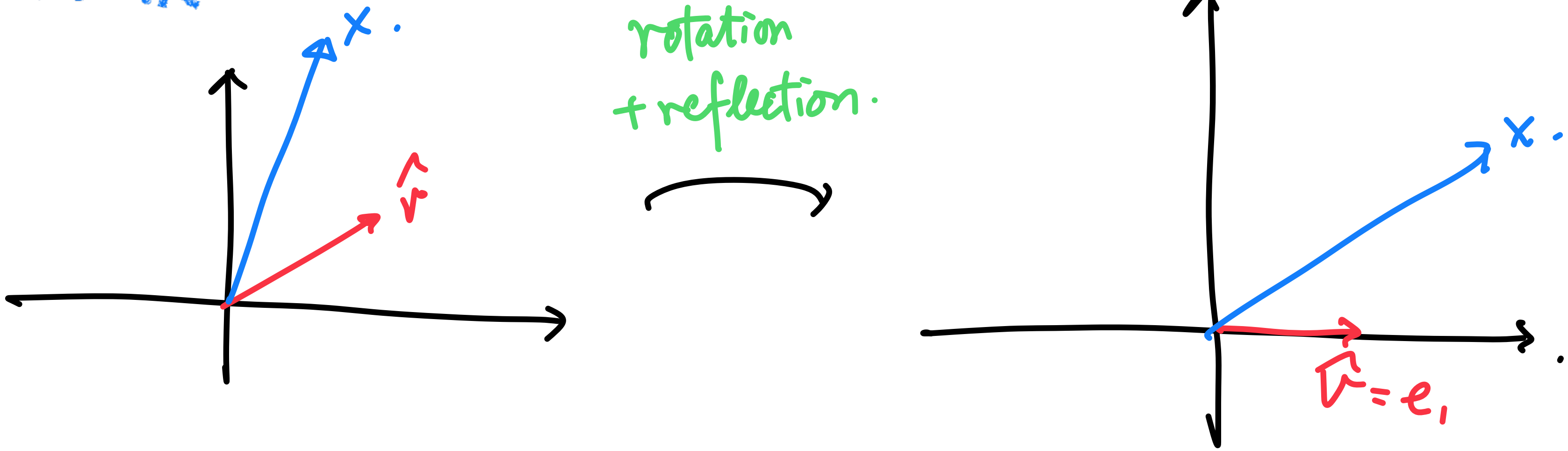


rotation + reflection.



Projection

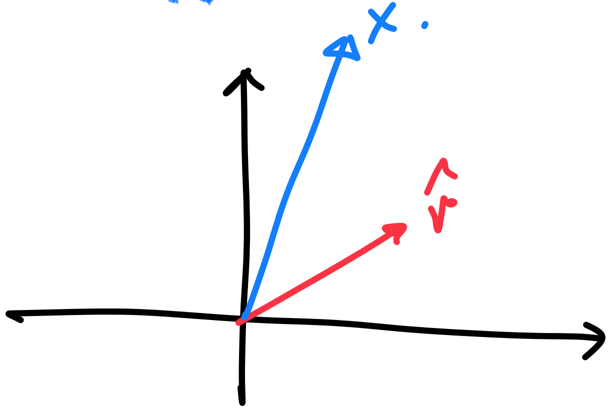
Say in $\mathbb{R}^2$:



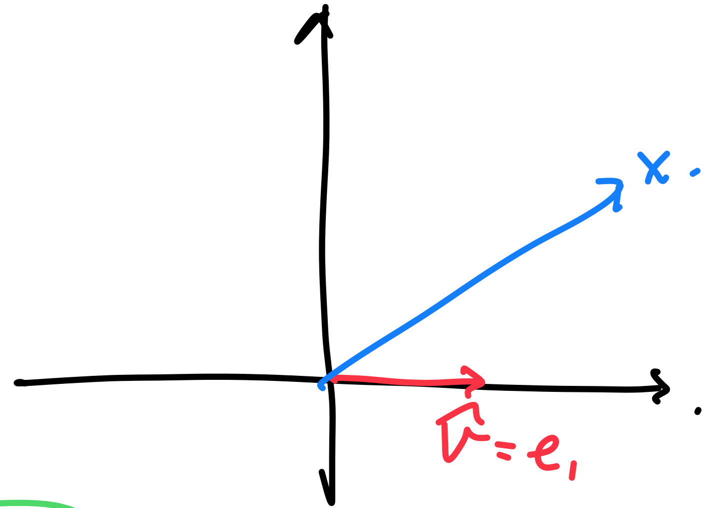
In the later picture: $P(x) = \langle x, \hat{v} \rangle \hat{v}$

Projection

Say in $\mathbb{R}^2$:



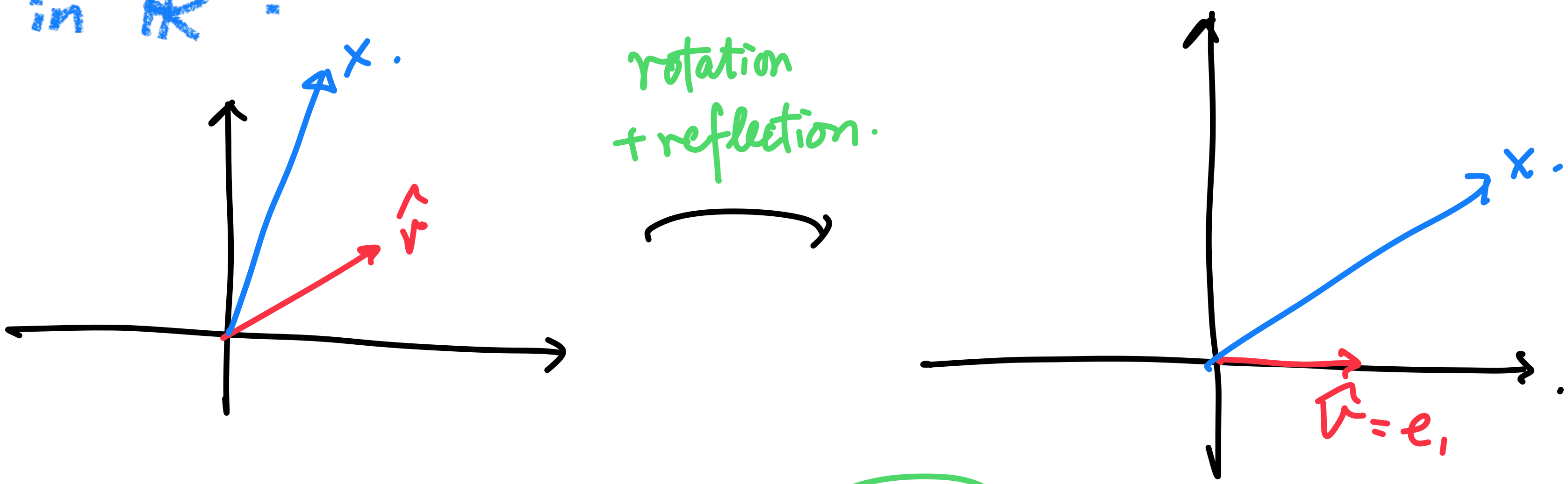
rotation + reflection.



In the later picture: $P(x) = \underbrace{\langle x, \hat{v} \rangle}_{= c.} \hat{v}$

Projection

Say in $\mathbb{R}^2$:

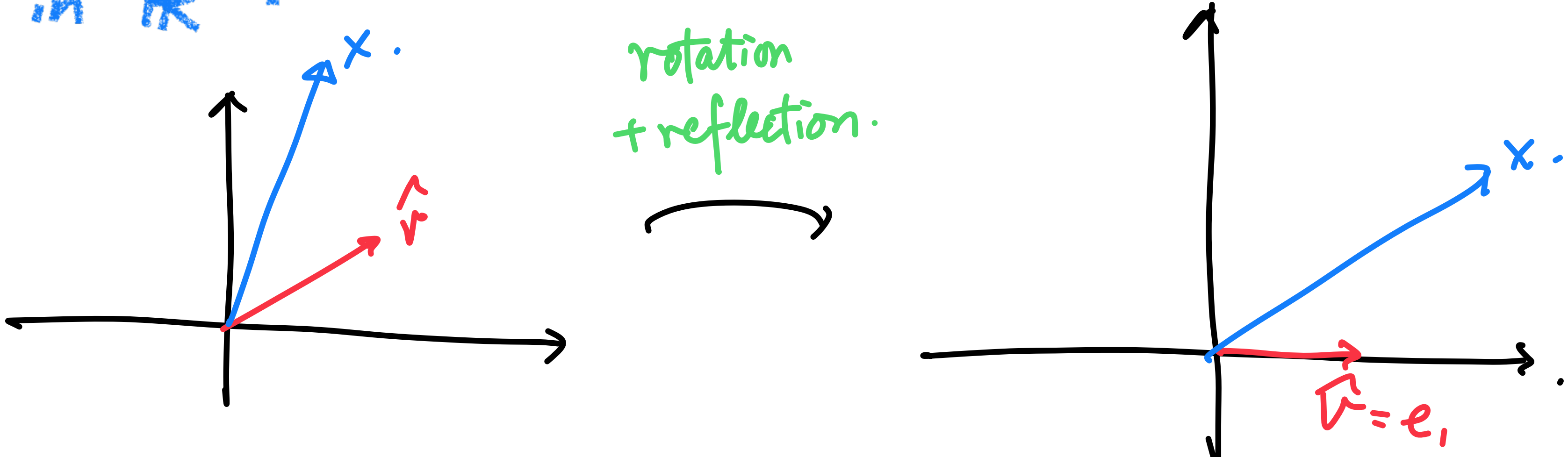


In the later picture: $P(x) = \underbrace{\langle x, \hat{v} \rangle}_{= c} \hat{v}$

Observation: This formula holds for any dimension.

Projection

Say in $\mathbb{R}^2$:



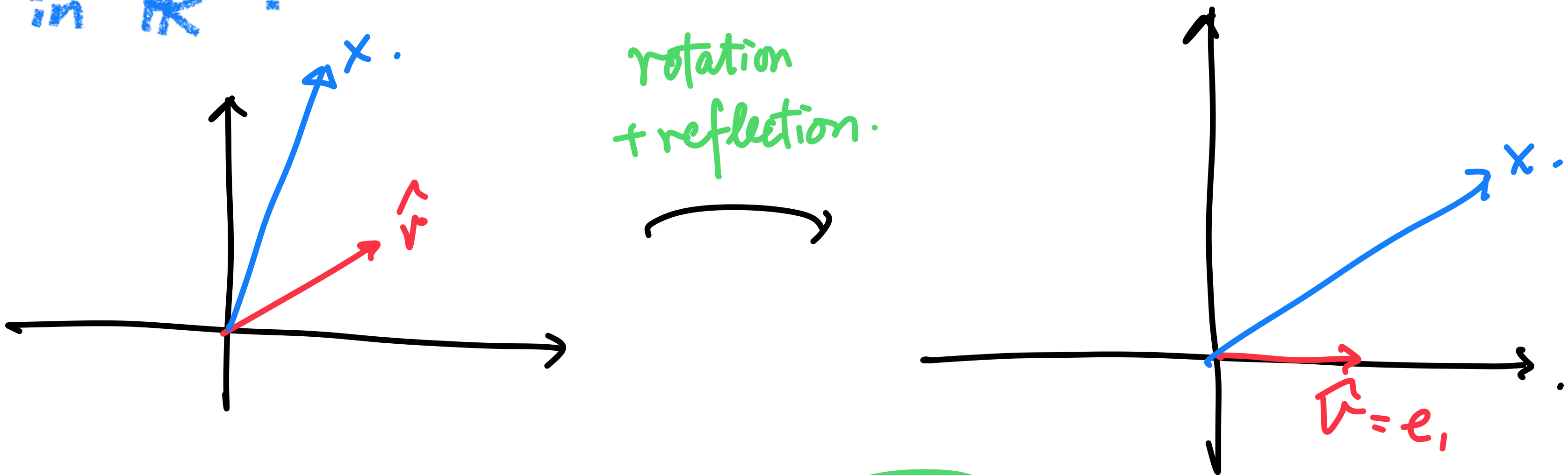
In the later picture: $P(x) = \underbrace{\langle x, \hat{v} \rangle}_{= c.} \hat{v}$

Observation: This formula holds for any dimension.

Claim: $P(x) = \langle x, \hat{v} \rangle \hat{v} = \dfrac{\langle x, v \rangle}{\|v\|^2} v$ is the required projection!

Projection

Say in $\mathbb{R}^2$:



In the later picture: $P(x) = \underbrace{\langle x, \hat{v} \rangle}_{= c.} \hat{v}$

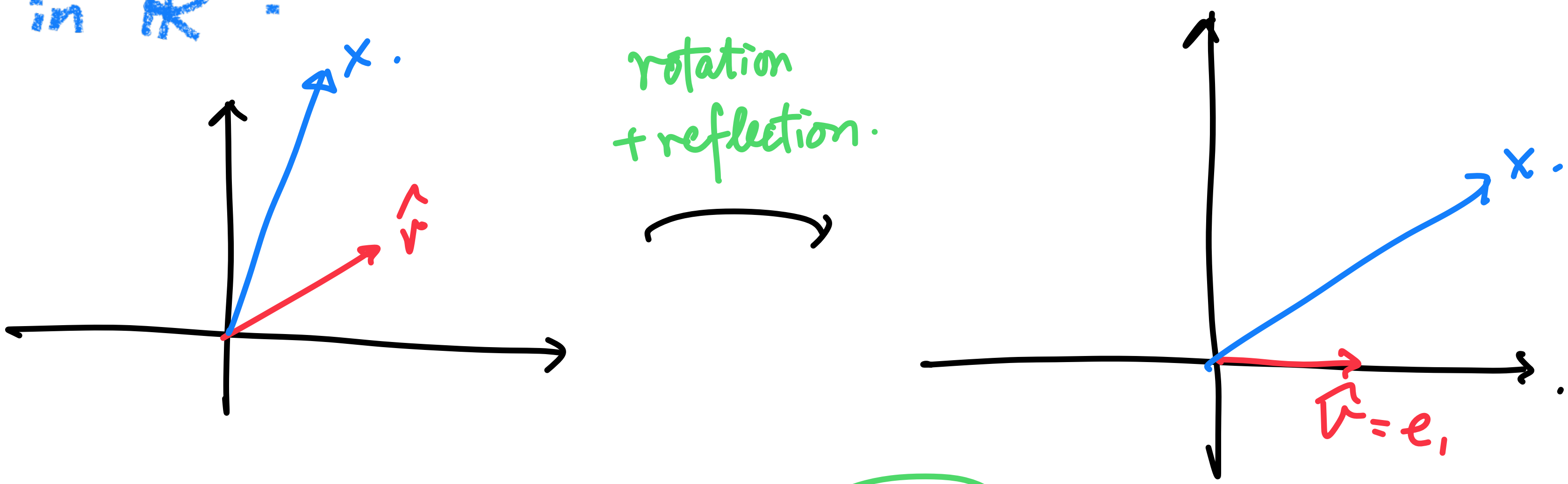
Observation: This formula holds for any dimension.

Claim: $P(x) = \langle x, \hat{v} \rangle \hat{v} = \dfrac{\langle x, v \rangle}{\|v\|^2} v$ is the required projection!

- Let any x, write $x = \left(x - \langle x, \hat{v} \rangle \hat{v}\right) + \langle x, \hat{v} \rangle \hat{v}$

Projection

Say in $\mathbb{R}^2$:



In the later picture: $P(x) = \underbrace{\langle x, \hat{v} \rangle}_{= c} \hat{v}$

Observation: This formula holds for any dimension.

Claim: $P(x) = \langle x, \hat{v} \rangle \hat{v} = \dfrac{\langle x, v \rangle}{\|v\|^2} v$ is the required projection!

- Let any x, write $x = \big(\underbrace{x - \langle x, \hat{v} \rangle \hat{v}}_{\text{in } V^{\perp}\,?}\big) + \underbrace{\langle x, \hat{v} \rangle \hat{v}}_{V}$

Projection

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^\perp$?

we have $\left\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \right\rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \langle \hat{v}, \hat{v} \rangle$

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?

we have $\left\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \right\rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{=1.}$

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^\perp$?

we have $\left\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \right\rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{=1.}$

$= \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle = 0 .$

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?

we have $\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{=1.}$

$= \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle = 0.$

e.g.: Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we compute the projection of x onto $\mathrm{Span}\{v\}$.

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?

we have $$\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{= 1.}$$
$$= \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle = 0 .$$

e.g.: Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we compute the projection of x onto $\mathrm{Span}\{v\}$.

- $\|v\|^2 = 1^2 + 1^2 + 1^2 = 3.$

Projection

Check: $x - \langle x, \hat{v}\rangle \hat{v}$ is in $V^{\perp}$?

we have $\left\langle x - \langle x,\hat{v}\rangle \hat{v}, \hat{v} \right\rangle = \langle x,\hat{v}\rangle - \langle x,\hat{v}\rangle \underbrace{\langle \hat{v},\hat{v}\rangle}_{=1.}$

$$= \langle x,\hat{v}\rangle - \langle x,\hat{v}\rangle = 0.$$

e.g.: Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we compute the projection of x onto $\mathrm{Span}\{v\}$.

- $\|v\|^2 = 1^2 + 1^2 + 1^2 = 3.$

- $\langle x, v\rangle = 1\cdot 1 + 2\cdot 1 + 3\cdot 1 = 6.$

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^\perp$?

we have $\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{= 1.}$

$= \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle = 0.$

e.g.: Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we compute the projection of x onto $\mathrm{Span}\{v\}$.

- $\|v\|^2 = 1^2 + 1^2 + 1^2 = 3.$

- $\langle x, v \rangle = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 = 6.$

$$P(x) = \frac{\langle x, v \rangle}{\|v\|^2} v = 2 \cdot v = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

Projection

Check: $x - \langle x, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?

we have $\left\langle x - \langle x, \hat{v} \rangle \hat{v}, \hat{v} \right\rangle = \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle \underbrace{\langle \hat{v}, \hat{v} \rangle}_{= 1.}$

$= \langle x, \hat{v} \rangle - \langle x, \hat{v} \rangle = 0$.

e.g.: Let $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we compute the projection of x onto $\mathrm{Span}\{v\}$.

- $\|v\|^2 = 1^2 + 1^2 + 1^2 = 3$.
- $\langle x, v \rangle = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 = 6$.

$$P(x) = \frac{\langle x, v \rangle}{\|v\|^2} v = 2 \cdot v = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

independent on the choice of v.

Projection

Projection

Q: How above the matrix correspond to projection onto Span{v}?

Projection

Q: How does the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v \rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

Projection

Q: How does the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v \rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

Projection

Q: How above the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v\rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

$$= \frac{v \cdot v^T \cdot x}{\|v\|^2} = \left(\frac{1}{\|v\|^2} v \cdot v^T\right) \cdot x.$$

Projection

Q: How above the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v\rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

$$= \frac{v \cdot v^T \cdot x}{\|v\|^2} = \underbrace{\left(\frac{1}{\|v\|^2} v \cdot v^T\right)}_{} \cdot x .$$

$n \times n$ matrix of $rk = 1$.

Projection

Q: How above the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v\rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

$$= \frac{v \cdot v^T \cdot x}{\|v\|^2} = \underbrace{\left(\frac{1}{\|v\|^2} v \cdot v^T \right)}_{n \times n \text{ matrix of } rk = 1.} \cdot x.$$

- The matrix is given by $\frac{1}{\|v\|^2} v \cdot v^T$

Projection

Q: How above the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v\rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

$$= \frac{v \cdot v^T \cdot x}{\|v\|^2} = \underbrace{\left(\frac{1}{\|v\|^2} v \cdot v^T\right)}_{} \cdot x.$$

$n \times n$ matrix of $rk = 1$.

- The matrix is given by $\frac{1}{\|v\|^2} v \cdot v^T$

e.g.: $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\frac{1}{\|v\|^2} = \frac{1}{3}$, $v \cdot v^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Projection

Q: How above the matrix correspond to projection onto Span{v}?

$$P(x) = \frac{\langle x, v\rangle}{\|v\|^2} v = \frac{(v^T \cdot x)}{\|v\|^2} \cdot v$$

a number.

$$= \frac{v \cdot v^T \cdot x}{\|v\|^2} = \underbrace{\left(\frac{1}{\|v\|^2} v \cdot v^T\right)} \cdot x.$$

$n \times n$ matrix of $rk = 1$.

- The matrix is given by $\frac{1}{\|v\|^2} v \cdot v^T$

e.g.: $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\frac{1}{\|v\|^2} = \frac{1}{3}$, $v \cdot v^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

then $[P] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Projection

Observation: For the projection P onto a line:

$[P]$ is a symmetric matrix:

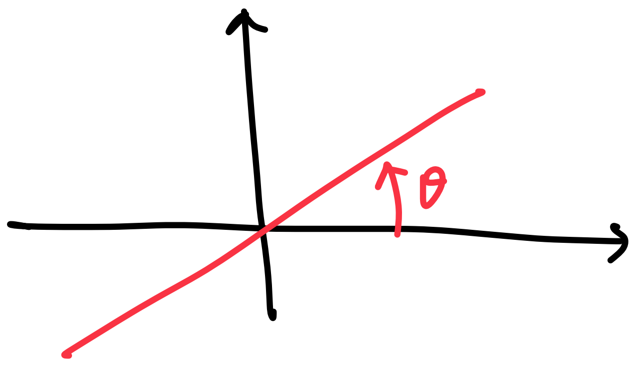
$$[P]^T = \left(\frac{1}{\|v\|^2} v \cdot v^T\right)^T = \frac{1}{\|v\|^2} v \cdot v^T = [P].$$

- Indeed: Any projection P onto V has its matrix representation under e_i's being symmetric (prove later).

e.g. We may check the formula for projection onto the θ-line in $\mathbb{R}^2$.

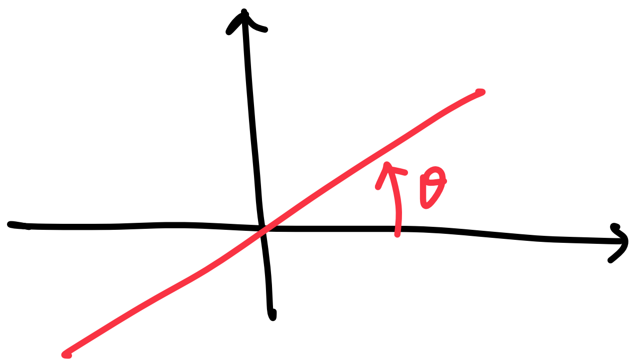
Projection

Projection



Find projection P onto the θ-line?

Projection

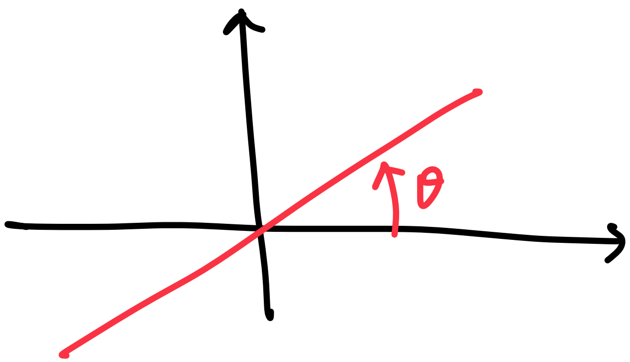


Find projection P onto the θ-line?

- take $v \neq 0$ parallel to the line

$$v = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

Projection



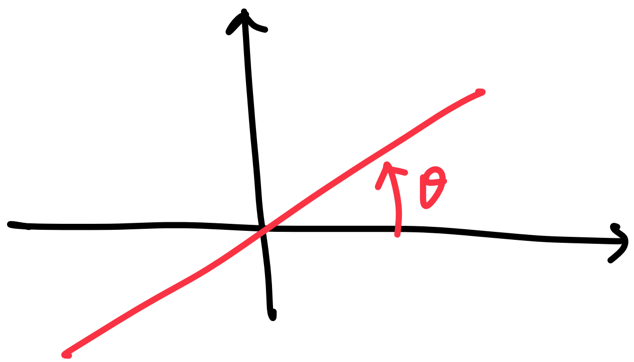
Find projection P onto the θ-line?

- take $v \neq 0$ parallel to the line

  $$v = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

- $$\|v\|^2 = \cos^2\theta + \sin^2\theta = 1, \quad v \cdot v^T = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$

Projection



Find projection P onto the θ-line?

- take $v \neq 0$ parallel to the line

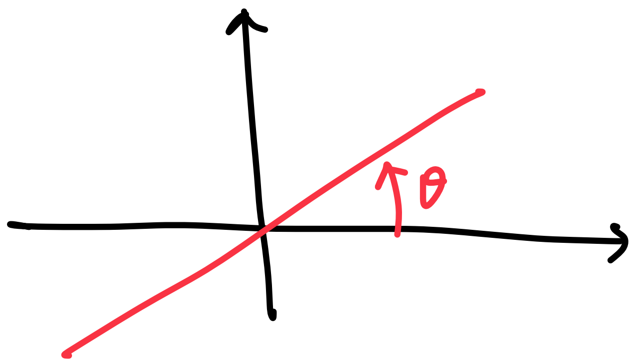
$$v = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

- $\|v\|^2 = \cos^2\theta + \sin^2\theta = 1, \quad v \cdot v^T = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$

- As a result

$$[P] = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

Projection



Find projection P onto the θ-line?

- take $v \neq 0$ parallel to the line

$$v = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

- $\|v\|^2 = \cos^2\theta + \sin^2\theta = 1$, $v \cdot v^T = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$

- As a result

$$[P] = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

which agrees with the one we get using transition matrix.

Projection

Projection

How to write down a general projection onto V ?

Projection

How to write down a general projection onto V ?

Def: • $\vec{q}_1, \ldots, \vec{q}_k$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal (標準正交) if

Projection

How to write down a general projection onto V?

Def: • $\vec{q}_1, \dots, \vec{q}_k$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal (標準正交) if

$$\langle \vec{q}_i, \vec{q}_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Projection

How to write down a general projection onto V?

Def: • $\vec{q}_1, \ldots, \vec{q}_k$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal (標準正交) if

$$\langle \vec{q}_i, \vec{q}_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

• If furthermore it is a basis for $V \leq \mathbb{R}^n$, we call it a orthonormal basis for V

Projection

How to write down a general projection onto V?

Def: • $\vec{q}_1, \dots, \vec{q}_k$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal (標準正交) if

$$\langle \vec{q}_i, \vec{q}_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

• If furthermore it is a basis for $V \leq \mathbb{R}^n$, we call it a orthonormal basis for V

e.g.: 1. $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ o.n. basis for $\mathbb{R}^3$.

Projection

How to write down a general projection onto V?

Def: • $\vec{q}_1, \cdots, \vec{q}_k$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal (標準正交) if

$$\langle \vec{q}_i, \vec{q}_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

• If furthermore it is a basis for $V \leq \mathbb{R}^n$, we call it a orthonormal basis for V

e.g.: 1. $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ o.n. basis for $\mathbb{R}^3$.

2. 

$v_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$, $v_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ o.n. basis for $\mathbb{R}^2$.

Projection

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).
- We can write the orthogonal projection onto V:

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_k$ (we should learn how to find it later).

- We can write the orthogonal projection onto V:

$$P(v) := \langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \cdots + \langle v, \vec{q}_k \rangle \vec{q}_k .$$

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).

- We can write the orthogonal projection onto V:

$$P(v) := \langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle v, \vec{q}_k \rangle \vec{q}_k .$$

$\in V$

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).

- We can write the orthogonal projection onto V:

$$P(v) := \underbrace{\langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle v, \vec{q}_k \rangle \vec{q}_k}_{\in V}.$$

- <u>Check:</u> $v_2 = v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k \in V^{\perp}.$

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_k$ (we should learn how to find it later).
- We can write the orthogonal projection onto V:

$$P(v) := \langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \cdots + \langle v, \vec{q}_k \rangle \vec{q}_k \;\in V$$

- Check: $v_2 = v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \cdots - \langle v, \vec{q}_k \rangle \vec{q}_k \in V^{\perp}$.
- we only need to check $\langle v_2, \vec{q}_i \rangle = 0$ for all $i = 1, \ldots, k$.

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).

- We can write the orthogonal projection onto V:

$$P(v) := \underbrace{\langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle v, \vec{q}_k \rangle \vec{q}_k}_{\in V}.$$

- <u>Check:</u> $v_2 = v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k \in V^{\perp}$.

- We only need to check $\langle v_2, \vec{q}_i \rangle = 0$ for all $i = 1, \dots, k$.

$$\langle v_2, \vec{q}_i \rangle = \langle v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k, \vec{q}_i \rangle$$

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).
- We can write the orthogonal projection onto V:

$$P(v) := \langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle v, \vec{q}_k \rangle \vec{q}_k . \quad \in V$$

- <u>Check:</u> $v_2 = v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k \in V^{\perp}$.
- we only need to check $\langle v_2, \vec{q}_i \rangle = 0$ for all $i = 1, \dots, k$.

$$\langle v_2, \vec{q}_i \rangle = \langle v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k , \vec{q}_i \rangle$$
$$= \langle v, \vec{q}_i \rangle - \langle v, \vec{q}_i \rangle \langle \vec{q}_i, \vec{q}_i \rangle = 0 .$$

Projection

- Let say give V, and given o.n. basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_k$ (we should learn how to find it later).

- We can write the orthogonal projection onto V:

$$P(v) := \langle v, \vec{q}_1 \rangle \vec{q}_1 + \langle v, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle v, \vec{q}_k \rangle \vec{q}_k \in V$$

- <u>Check:</u> $v_2 = v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k \in V^{\perp}$.

- we only need to check $\langle v_2, \vec{q}_i \rangle = 0$ for all $i = 1, \dots, k$.

$$\langle v_2, \vec{q}_i \rangle = \langle v - \langle v, \vec{q}_1 \rangle \vec{q}_1 - \dots - \langle v, \vec{q}_k \rangle \vec{q}_k, \vec{q}_i \rangle$$

$$= \langle v, \vec{q}_i \rangle - \langle v, \vec{q}_i \rangle \underbrace{\langle \vec{q}_i, \vec{q}_i \rangle}_{=1} = 0.$$

Projection

Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

- $[P] = q_1 q_1^T + q_2 q_2^T + \dots + q_k q_k^T$.

Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).
- $$[P] = q_1 q_1^T + q_2 q_2^T + \cdots + q_k q_k^T.$$

Observation: $[P]^T = [P]$, i.e. it is a symmetric matrix.

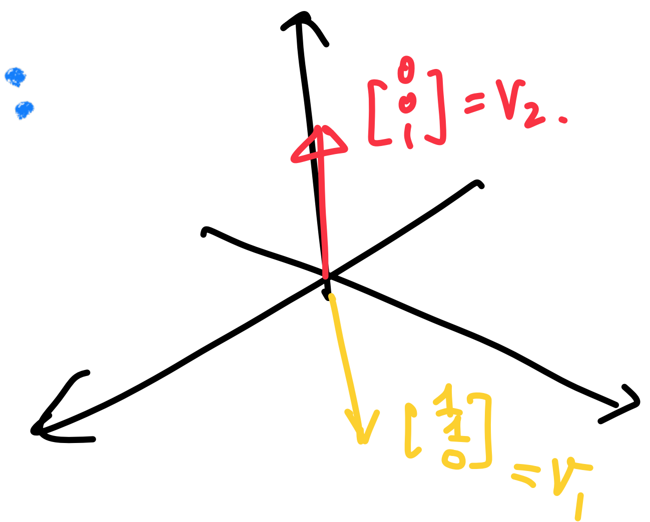
Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

- $$[P] = q_1 q_1^T + q_2 q_2^T + \cdots + q_k q_k^T.$$

Observation: $[P]^T = [P]$, i.e. it is a symmetric matrix.

e.g.:



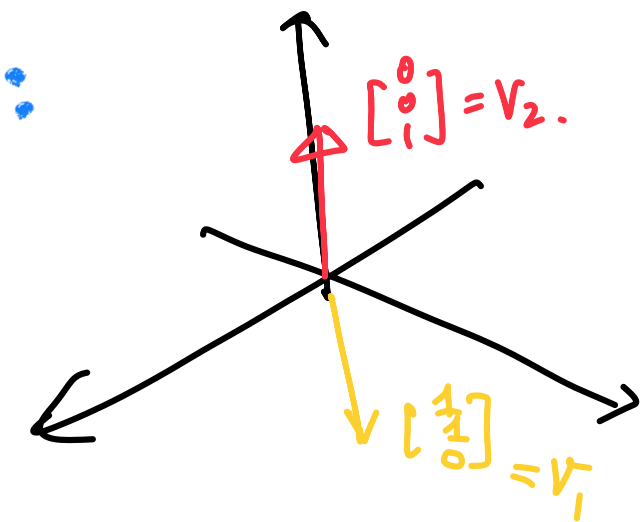
Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

- $$[P] = q_1 q_1^T + q_2 q_2^T + \dots + q_k q_k^T .$$

Observation: $[P]^T = [P]$, i.e. it is a symmetric matrix.

e.g.:



- $\mathrm{Span}\{v_1, v_2\} = V$.

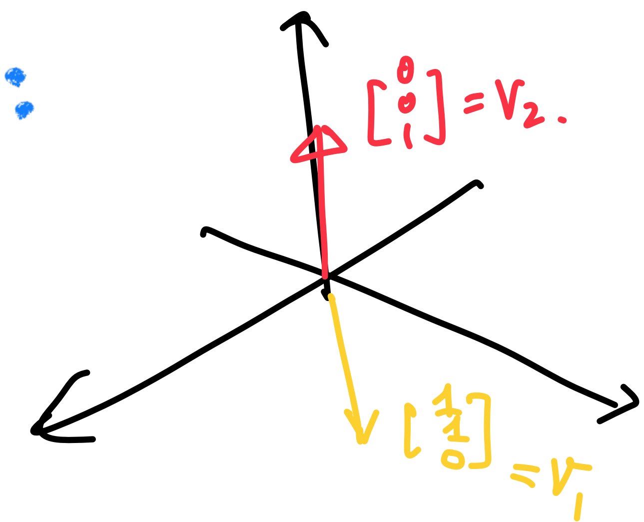
Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

- $[P] = q_1 q_1^T + q_2 q_2^T + \dots + q_k q_k^T$.

Observation: $[P]^T = [P]$, i.e. it is a symmetric matrix.

e.g.:



- $\mathrm{Span}\{v_1, v_2\} = V.$

- $\langle v_1, v_2 \rangle = 0$, $q_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}$

  $q_2 = \frac{v_2}{\|v_2\|} = \begin{bmatrix}0\\0\\1\end{bmatrix}$

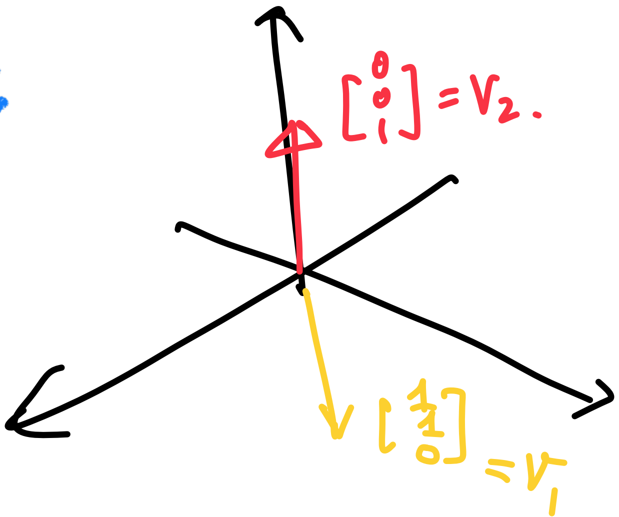
Projection

- Suppose we want to write down matrix representation of P: (w.r.t. e_i's).

- $[P] = q_1 q_1^T + q_2 q_2^T + \dots + q_k q_k^T$.

Observation: $[P]^T = [P]$, i.e. it is a symmetric matric.

e.g.:



- $\mathrm{Span}\{v_1, v_2\} = V$.

- $\langle v_1, v_2 \rangle = 0$, $q_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

  $q_2 = \frac{v_2}{\|v_2\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

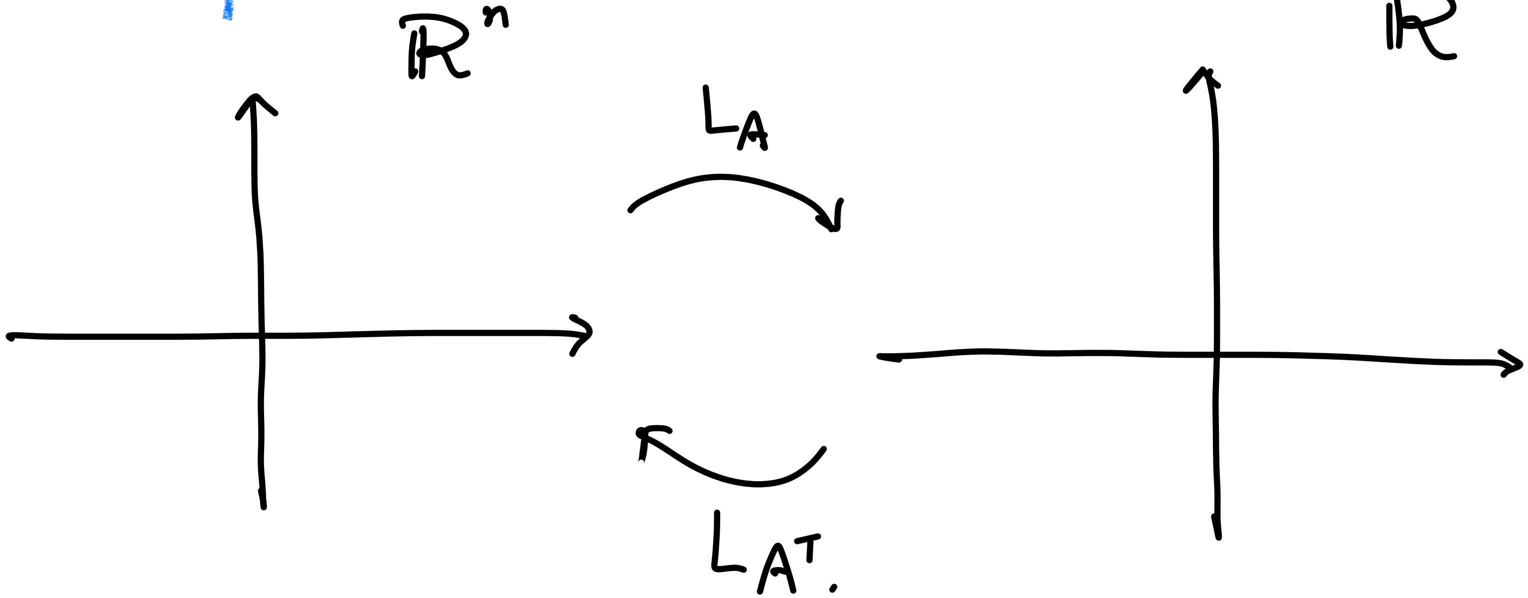
- $[P] = q_1 q_1^T + q_2 q_2^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Transpose

Transpose

What is Transpose?

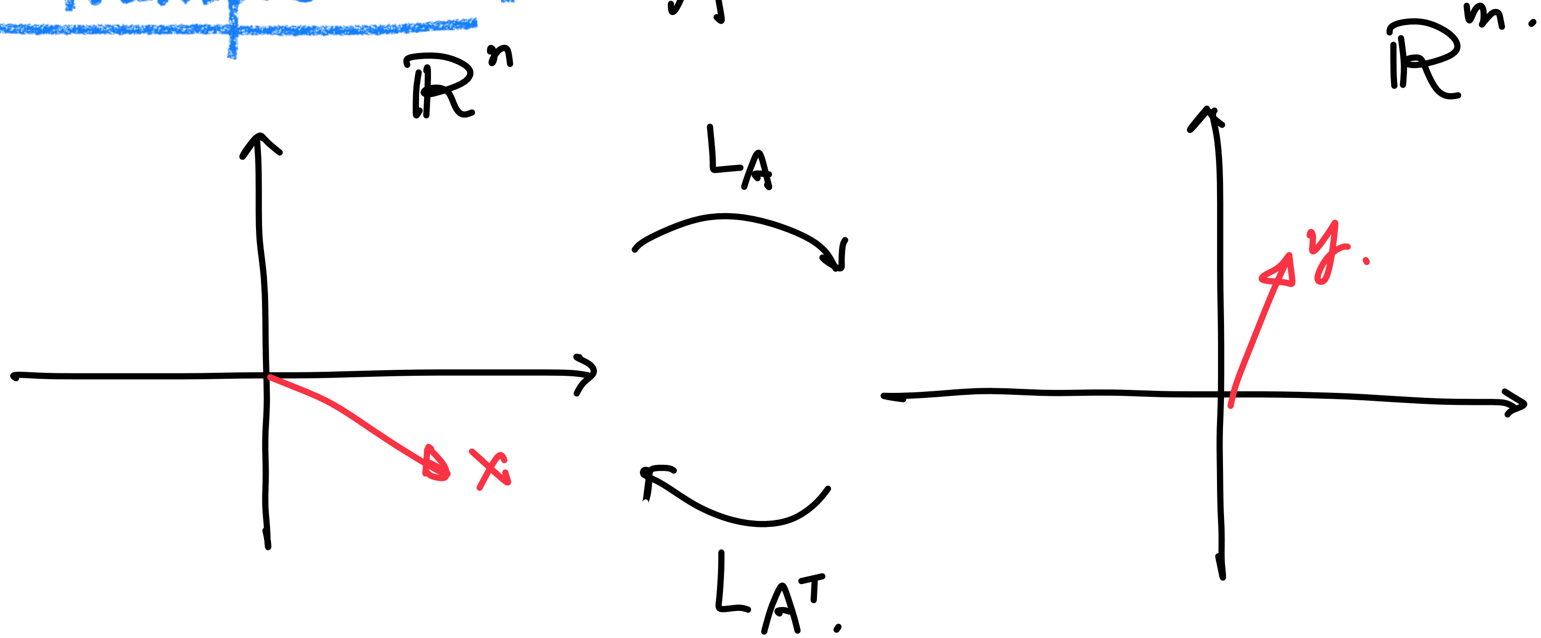
A: $m \times n$ matrix.



Transpose

<u>What is Transpose?</u>

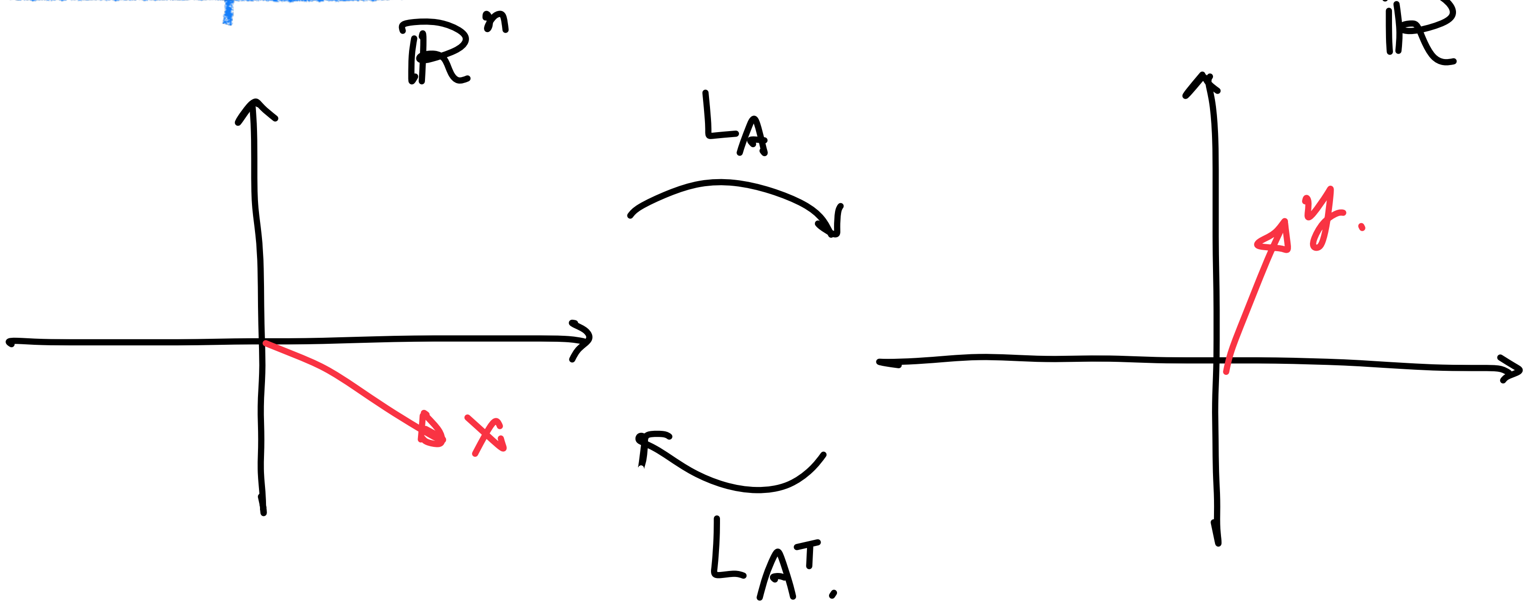
A : $m \times n$ matrix.



Transpose

What is Transpose?

A : $m \times n$ matrix.



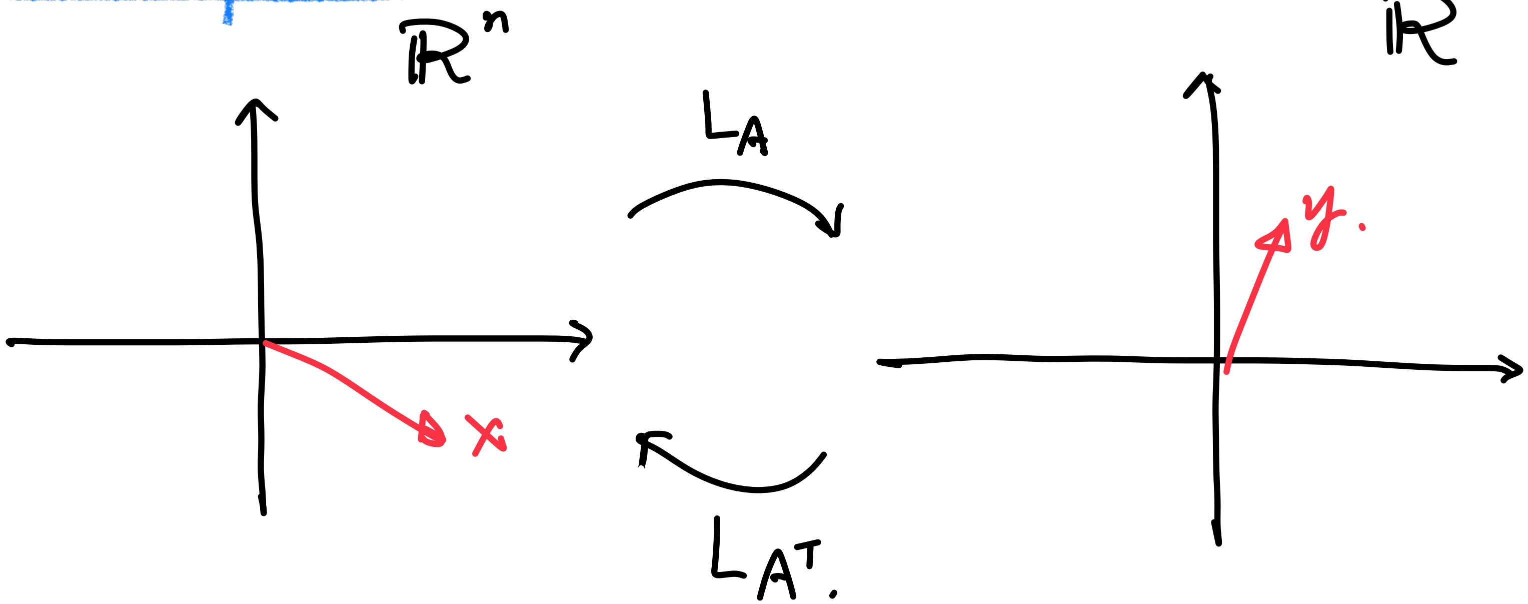
- A^T is characterized by the property

$$\langle Ax, y \rangle_{\mathbb{R}^m} = \langle x, A^T y \rangle_{\mathbb{R}^n}.$$

Transpose

What is Transpose?

A : $m \times n$ matrix.



- A^T is characterized by the property

$$\langle Ax, y \rangle_{\mathbb{R}^m} = \langle x, A^T y \rangle_{\mathbb{R}^n}.$$

Why: L.H.S. $= (Ax)^T y = x^T (A^T y) = \langle x, A^T y \rangle$

Transpose

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation:

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation: if $\langle Ax, y \rangle = \langle x, By \rangle$ for some B.

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation: if $\langle Ax, y \rangle = \langle x, By \rangle$ for some B.

$\Rightarrow \quad x^T A^T y = x^T B y$ for all $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation: if $\langle Ax, y\rangle = \langle x, By\rangle$ for some B.

$$\Rightarrow \quad x^T A^T y = x^T B y \quad \text{for all } x \in \mathbb{R}^n,\ y \in \mathbb{R}^m.$$

$$\Rightarrow \quad x^T C y = 0 \quad \text{for } C = (A^T - B)$$

Transpose

- A^T is the unique $n\times m$ matrix satisfying the above equation: if $\langle Ax, y\rangle = \langle x, By\rangle$ for some B.

$\Rightarrow$ $x^TA^Ty = x^TBy$ for all $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$.

$\Rightarrow$ $x^TCy = 0$ for $C = (A^T - B)$

$\because$ $\langle x, Cy\rangle = 0$ for all $x\in\mathbb{R}^n$

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation:

$$\text{if } \langle Ax, y\rangle = \langle x, By\rangle \text{ for some } B.$$

$$\Rightarrow \quad x^T A^T y = x^T B y \quad \text{for all } x \in \mathbb{R}^n,\ y \in \mathbb{R}^m.$$

$$\Rightarrow \quad x^T C y = 0 \quad \text{for } C = (A^T - B)$$

$$\because \ \langle x, Cy\rangle = 0 \text{ for all } x \in \mathbb{R}^n \Rightarrow Cy = 0. \text{ for all } y$$

Transpose

- A^T is the unique $n \times m$ matrix satisfying the above equation:

$$\text{if } \langle Ax, y \rangle = \langle x, By \rangle \text{ for some } B.$$

$$\Rightarrow \quad x^T A^T y = x^T B y \quad \text{for all } x \in \mathbb{R}^n,\ y \in \mathbb{R}^m.$$

$$\Rightarrow \quad x^T C y = 0 \quad \text{for } C = (A^T - B)$$

$$\because \ \langle x, Cy \rangle = 0 \text{ for all } x \in \mathbb{R}^n \Rightarrow Cy = 0. \text{ for all } y \Rightarrow C = 0.$$

Transpose

- A^T is the unique $n\times m$ matrix satisfying the above equation:

  if $\langle Ax, y\rangle = \langle x, By\rangle$ for some B.

  $\Rightarrow$ $x^TA^Ty = x^TBy$ for all $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$.

  $\Rightarrow$ $x^TCy = 0$ for $C = (A^T - B)$

  $\because$ $\langle x, Cy\rangle = 0$ for all $x\in\mathbb{R}^n \Rightarrow Cy = 0$. for all $y \Rightarrow C = 0$.

- That is the reason we have $(AB)^T = B^TA^T$.

Transpose

- A^T is the unique $n\times m$ matrix satisfying the above equation:

$$\text{if } \langle Ax, y\rangle = \langle x, By\rangle \text{ for some } B.$$

$$\Rightarrow \quad x^T A^T y = x^T B y \quad \text{for all } x\in\mathbb{R}^n,\ y\in\mathbb{R}^m.$$

$$\Rightarrow \quad x^T C y = 0 \quad \text{for } C = (A^T - B)$$

$$\because \ \langle x, Cy\rangle = 0 \text{ for all } x\in\mathbb{R}^n \Rightarrow Cy = 0. \text{ for all } y \Rightarrow C = 0.$$

- That is the reason we have $(AB)^T = B^T A^T$.

$$\langle ABx, y\rangle = \langle Bx, A^T y\rangle = \langle x, B^T A^T y\rangle$$

Transpose

- A^T is the unique $n\times m$ matrix satisfying the above equation: if $\langle Ax, y\rangle = \langle x, By\rangle$ for some B.

$$\Rightarrow \quad x^T A^T y = x^T B y \quad \text{for all } x\in\mathbb{R}^n,\ y\in\mathbb{R}^m.$$

$$\Rightarrow \quad x^T C y = 0 \quad \text{for } C = (A^T - B)$$

$\because$ $\langle x, Cy\rangle = 0$ for all $x\in\mathbb{R}^n \Rightarrow Cy = 0$. for all $y \Rightarrow C = 0$.

- That is the reason we have $(AB)^T = B^T A^T$.

$$\langle ABx, y\rangle = \langle Bx, A^T y\rangle = \langle x, B^T A^T y\rangle$$

by this property.

- Therefor we must have $(AB)^T = B^T A^T$

Least square

Least square

- Once we can measure length, we can talk about approximation

Least square

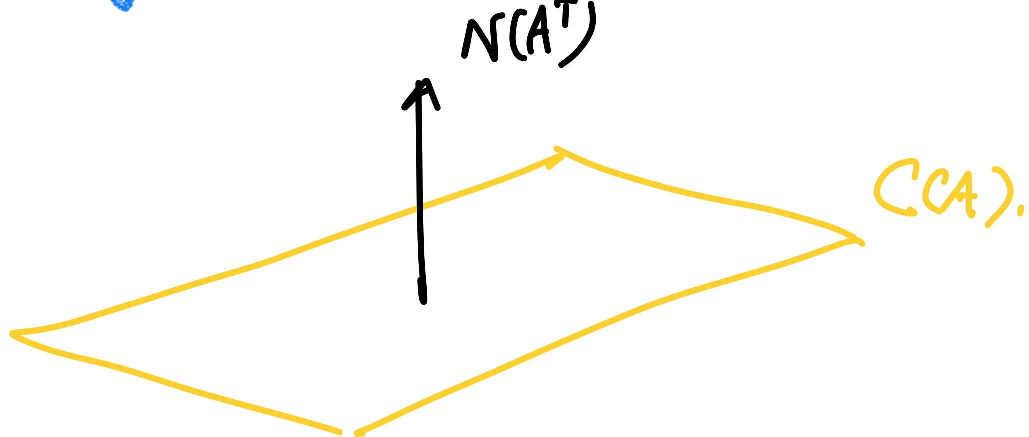
- Once we can measure length, we can talk about approximation

Inconsistent equation: $Ax=b$ has no solution.

Least square

- Once we can measure length, we can talk about approximation

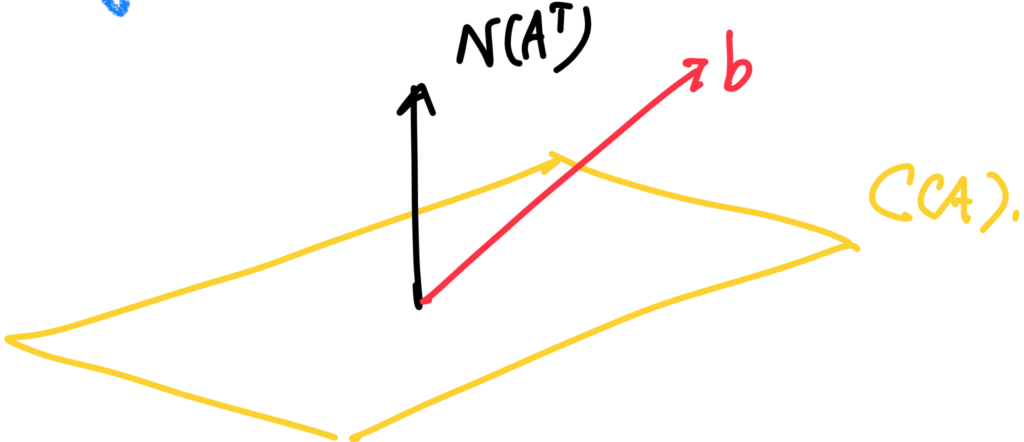
Inconsistent equation : $Ax = b$ has no solution.



Least square

- Once we can measure length, we can talk about approximation

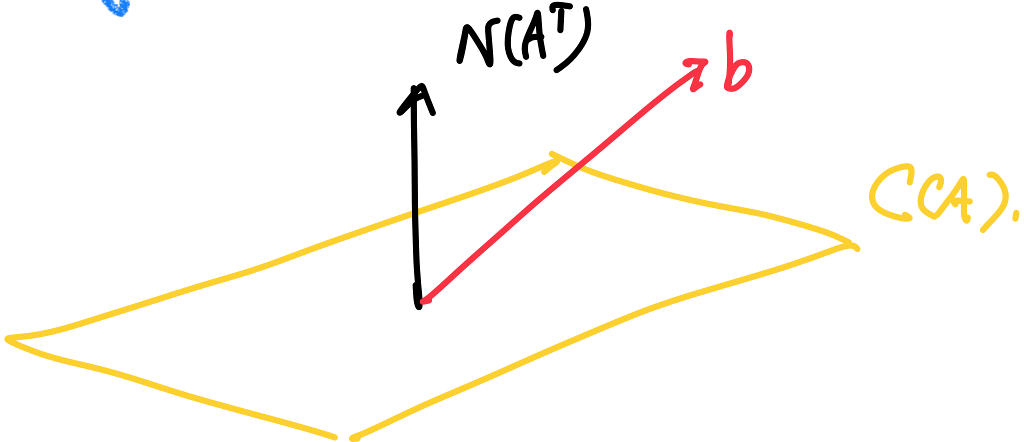
<u>Inconsistent equation :</u> $Ax=b$ has no solution.



Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation: $Ax=b$ has no solution.

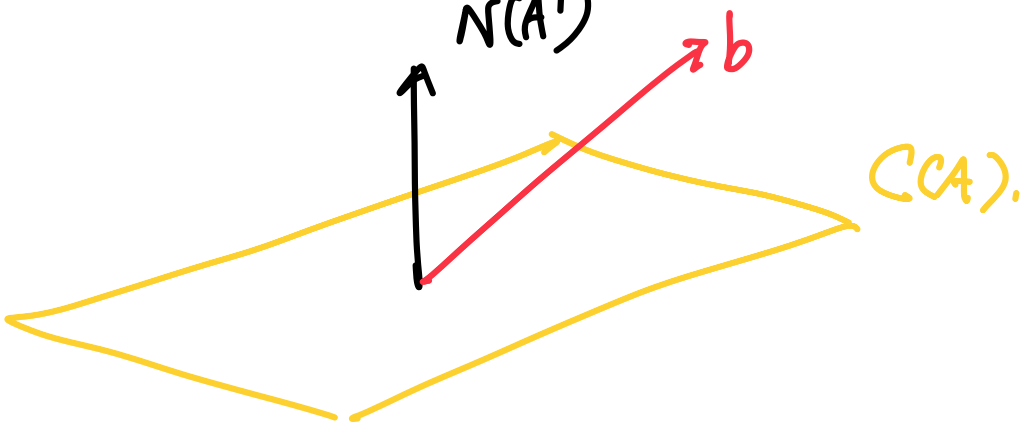


i.e. b is outside $C(A)$.

Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation : $Ax=b$ has no solution.



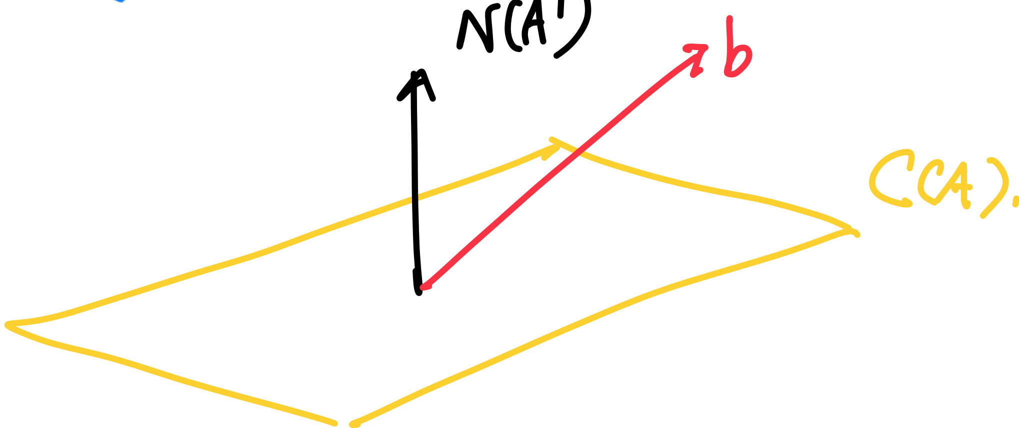
i.e. b is outside $C(A)$.

- We want best approximation: $\|Ax-b\|^2$ attain minimum

Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation: $Ax = b$ has no solution.



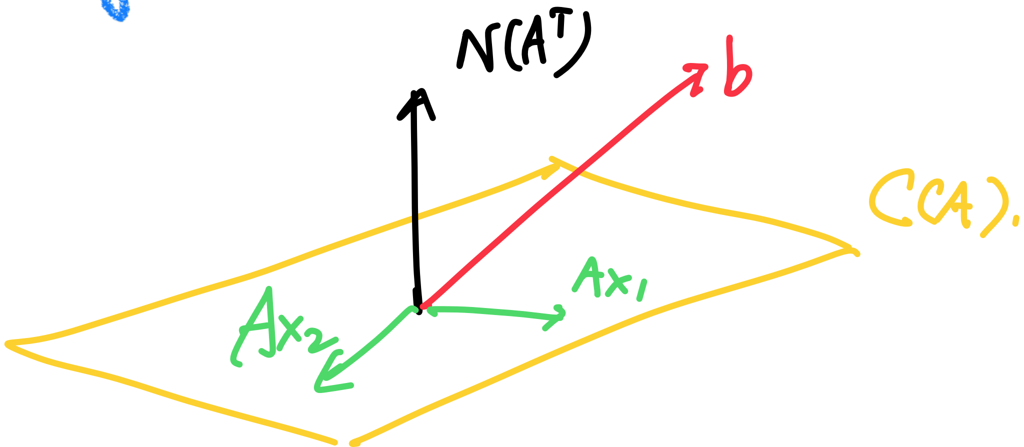
i.e. b is outside $C(A)$.

- We want best approximation: $\|Ax - b\|^2$ attain minimum
- We want to find a point $A\hat{x}$ s.t. it is closest to b.

Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation : $Ax=b$ has no solution.



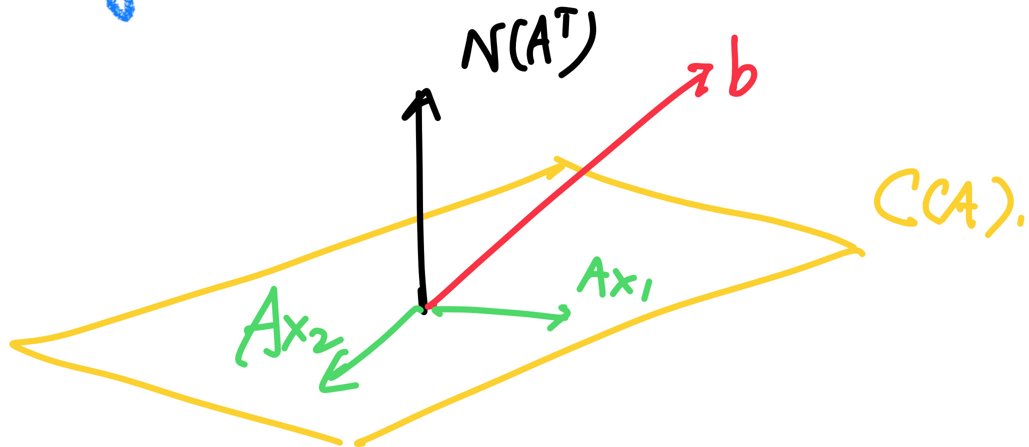
i.e. b is outside $C(A)$.

- We want best approximation: $\|Ax-b\|^2$ attain minimum
- We want to find a point $A\hat{x}$ s.t. it is closest to b.

Least square

- Once we can measure length, we can talk about approximation

<u>Inconsistent equation</u> : $Ax = b$ has no solution.



<u>i.e.</u> b is outside $C(A)$.

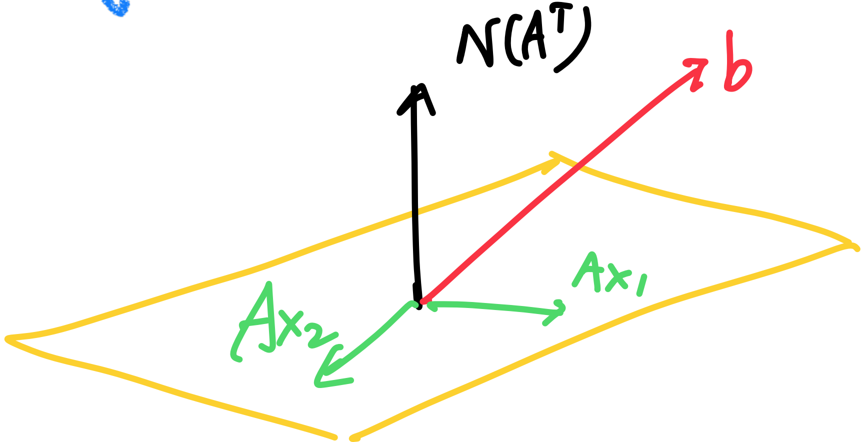
- We want best approximation: $\|Ax - b\|^2$ attain minimum
- We want to find a point $A\hat{x}$ s.t. it is closest to b.

$$\Longrightarrow \quad A\hat{x} = b_1$$

Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation : $Ax=b$ has no solution.



i.e. b is outside $C(A)$.

- We want best approximation: $\|Ax-b\|^2$ attain minimum
- We want to find a point $A\hat{x}$ s.t. it is closest to b.

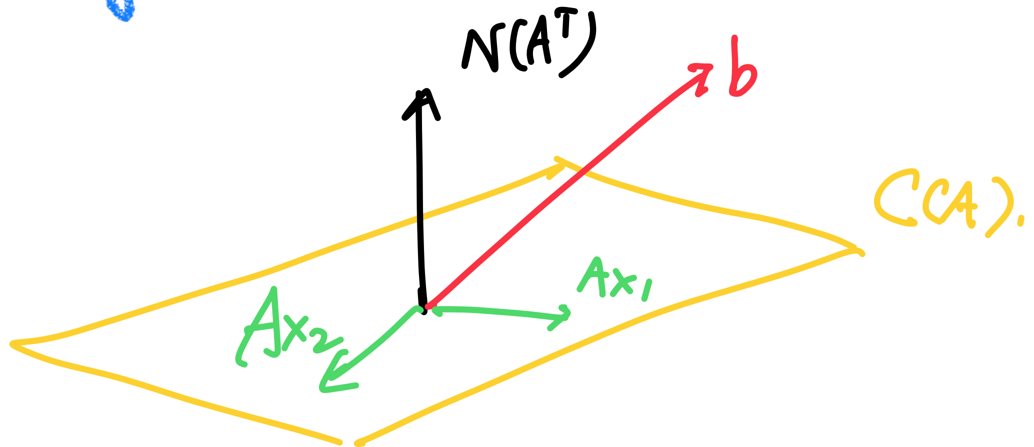
$$\Longrightarrow \quad A\hat{x} = b_1$$

orthogonal projection onto $C(A)$.

Least square

- Once we can measure length, we can talk about approximation

Inconsistent equation : $Ax = b$ has no solution.



i.e. b is outside $C(A)$.

- We want best approximation: $\|Ax - b\|^2$ attain minimum
- We want to find a point $A\hat{x}$ s.t. it is closest to b.

$$\Longrightarrow \quad A\hat{x} = b_1$$ ← orthogonal projection onto $C(A)$.

- Then $b - A\hat{x}$ is in $C(A)^\perp = N(A^T)$.

Least square

Least square

- That means we have:

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

Least square

- That means we have:

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

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\ :\ \text{error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

Least square

- That means we have:

$$\underbrace{A^T(b - A\hat{x})}_{} = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

e : error vector.

- Normal equation : $(A^TA)x = A^Tb.$

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e})=0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

e : error vector.

- Normal equation : $(A^TA)x = A^Tb.$ always solvable!

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\ :\ \text{error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

- Normal equation : $(A^TA)x = A^Tb.$ always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

e : error vector.

- Normal equation : $(A^TA)x = A^Tb$. always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Cross-product matrix

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

e : error vector.

- Normal equation : $(A^TA)x = A^Tb$. always solvable!

A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Cross-product matrix $n \times n$ matrix

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\text{ : error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

- Normal equation : $(A^TA)x = A^Tb.$ always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Cross-product matrix nxn matrix

~~Property~~ :

- A^TA symmetric

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\ :\ \text{error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

- Normal equation: $(A^TA)x = A^Tb.$ always solvable!

A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Cross-product matrix nxn matrix

Property: • A^TA symmetric (warning: $(A^TA)^T \neq AA^T$).

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

e : error vector.

- Normal equation : $(A^TA)x = A^Tb$. always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

Cross-product matrix $n \times n$ matrix

Property:

- A^TA symmetric (warning: $(A^TA)^T \neq AA^T$).
- $N(A^TA) = N(A)$

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\ :\ \text{error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

- Normal equation : $(A^TA)x = A^Tb.$ always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

  Cross-product matrix nxn matrix

Property:

- A^TA symmetric (warning: $(A^TA)^T \neq AA^T$).
- $N(A^TA) = N(A)$
- If $N(A) = \{0\} \Rightarrow A^TA$ invertible.

Least square

- That means we have:

$$A^T(\underbrace{b - A\hat{x}}_{e\ :\ \text{error vector.}}) = 0 \quad \text{or} \quad A^TA\hat{x} = A^Tb.$$

- Normal equation: $(A^TA)x = A^Tb.$ always solvable!

  A solution $\hat{x}$ would be a minimizer of $\|Ax-b\|^2$.

  Cross-product matrix $n \times n$ matrix

Property:

- A^TA symmetric (warning: $(A^TA)^T \neq AA^T$).
- $N(A^TA) = N(A)$
- If $N(A) = \{0\} \Rightarrow A^TA$ invertible.

  $\hat{x}$ has unique solution $\hat{x} = (A^TA)^{-1}A^Tb.$

Least square

Least square

Proof of $N(A^TA) = N(A)$:

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

.

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

.

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

- $\Rightarrow x^T \cdot A^T A x = 0$.

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

- $\Rightarrow x^T \cdot A^T A x = 0 \Longrightarrow (Ax)^T \cdot Ax = 0$.

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

- $\Rightarrow x^T \cdot A^TAx = 0 \Longrightarrow (Ax)^T \cdot Ax = 0.$

$$\Longrightarrow \|Ax\|^2 = 0.$$

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

- $\Rightarrow x^T \cdot A^TAx = 0 \Longrightarrow (Ax)^T \cdot Ax = 0.$

$$\Longrightarrow \|Ax\|^2 = 0.$$

$$\Longrightarrow Ax = 0.$$

Least square

Proof of $N(A^TA) = N(A)$:

- For product of matrices, we generally have

$$N(A) \subseteq N(B \cdot A).$$

- For $B = A^T$, we show $N(A^TA) \subseteq N(A)$.

  take $x \in N(A^TA)$, by definition $A^TAx = 0$.

- $\Rightarrow x^T \cdot A^TAx = 0 \Longrightarrow (Ax)^T \cdot Ax = 0.$

$$\Longrightarrow \|Ax\|^2 = 0.$$

$$\Longrightarrow Ax = 0.$$

$$\Longrightarrow x \in N(A).$$

Least square

Least square

e.g.: $\cdot$ $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.

Least square

e.g.:

- $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.
- $C(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Least square

e.g.:

- $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.
- $C(A) = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation:

- Solving $A^T A x = A^T b$

Least square

e.g.:

- $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.
- $C(A) = \mathrm{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation:

- Solving $A^T A x = A^T b = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$

Least square

e.g.:

- $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.
- $C(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation:

- Solving $A^TAx = A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$
- $N(A) = \{0\}$,

Least square

e.g.: • $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.

• $C(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation: • Solving $A^TAx = A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$

• $N(A) = \{0\}$, $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$

Least square

e.g.: • $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax = b$ has no solution.

• $C(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation: • Solving $A^TAx = A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$

• $N(A) = \{0\}$, $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$

• $(A^TA)^{-1} = \frac{1}{\det(A^TA)} \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$

Least square

e.g.: • $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $Ax=b$ has no solution.

• $C(A) = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$ not containing b.

Approximation: • Solving $A^TAx = A^Tb = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$

• $N(A) = \{0\}$, $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$

• $$(A^TA)^{-1} = \frac{1}{\det(A^TA)} \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$$

$$\Longrightarrow \quad \hat{x} = (A^TA)^{-1}A^Tb = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 23 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Least square

Least square

Projection: • Look at $A^TA\hat{x} = A^Tb$.

Recall that $A\hat{x}$ is the projection of b onto $C(A)$,

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x}$$

Least square

<u>Projection</u>:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$,

- $P(b) = A\hat{x} \underset{\text{If } N(A)=\{0\}}{=\!=}$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$,

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} A(A^TA)^{-1}A^Tb.$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$,

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}}b.$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$,

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}}b.$$

Property: • $[P]^2 = A(A^TA)^{-1}A^T A(A^TA)^{-1}A^T$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}}b.$$

Property:

- $$[P]^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T$$
  $$= A(A^TA)^{-1}A^T = [P].$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}}b.$$

Property:

- $$[P]^2 = A(A^TA)^{-1}A^T A(A^TA)^{-1}A^T$$
  $$= A(A^TA)^{-1}A^T = [P].$$

- $$[P]^T = \left(A(A^TA)^{-1}A^T\right)^T$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}} b$.

Property:

- $[P]^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T$

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

- $[P]^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left((A^TA)^{-1}\right)^TA^T$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}} b.$$

Property:

- $$[P]^2 = A(A^TA)^{-1}A^T A(A^TA)^{-1}A^T$$
  $$= A(A^TA)^{-1}A^T = [P].$$

- $$[P]^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left((A^TA)^{-1}\right)^T A^T$$

- $$(Q^{-1})^T = (Q^T)^{-1}$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}}b.$$

Property:

- $$[P]^2 = A(A^TA)^{-1}A^T A(A^TA)^{-1}A^T$$
  $$= A(A^TA)^{-1}A^T = [P].$$

- $$[P]^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left((A^TA)^{-1}\right)^T A^T$$

- $$(Q^{-1})^T = (Q^T)^{-1} \Rightarrow \left((A^TA)^{-1}\right)^T = (A^TA)^{-1}.$$

Least square

Projection:

- Look at $A^TA\hat{x} = A^Tb$.

  Recall that $A\hat{x}$ is the projection of b onto $C(A)$.

- $$P(b) = A\hat{x} \overset{\text{If } N(A)=\{0\}}{=\!=} \underbrace{A(A^TA)^{-1}A^T}_{\text{is the matrix for projection}} b.$$

Property:

- $$[P]^2 = A(A^TA)^{-1}A^T A(A^TA)^{-1}A^T$$
  $$= A(A^TA)^{-1}A^T = [P].$$

- $$[P]^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left((A^TA)^{-1}\right)^T A^T$$

- $$(Q^{-1})^T = (Q^T)^{-1} \Rightarrow \left((A^TA)^{-1}\right)^T = (A^TA)^{-1}.$$

- $$[P]^T = [P].$$

Projection

Projection

Prop: P is a projection iff 1. $P^2 = P$.
2. $P^T = P$.

Projection

Prop: P is a projection iff 1. $P^2 = P$.
2. $P^T = P$.

- Take $V \leq \mathbb{R}^n$, there are two ways to obtain projection matrix.

Projection

Prop: P is a projection iff
1. $P^2 = P$.
2. $P^T = P$.

- Take $V \le \mathbb{R}^n$, there are two ways to obtain projection matrix.

1. Find an orthonormal basis $q_1, \dots, q_k$ of V.

$$P = q_1 q_1^T + \dots + q_k q_k^T$$

Projection

Prop: P is a projection iff
1. $P^2 = P$.
2. $P^T = P$.

- Take $V \leq \mathbb{R}^n$, there are two ways to obtain projection matrix.

1. Find an orthonormal basis $q_1, \ldots, q_k$ of V.

$$P = q_1 q_1^T + \cdots + q_k q_k^T$$

2. Find a basis $v_1, \ldots, v_k$ of V: and let

Projection

Prop: P is a projection iff 1. $P^2 = P$.
2. $P^T = P$.

- Take $V \le \mathbb{R}^n$, there are two ways to obtain projection matrix.

1. Find an orthonormal basis $q_1, \dots, q_k$ of V.

$$P = q_1 q_1^T + \dots + q_k q_k^T$$

2. Find a basis $v_1, \dots, v_k$ of V : and let

- $A = [v_1, \dots, v_k] : \mathbb{R}^k \longrightarrow \mathbb{R}^n$ with

Projection

Prop: P is a projection iff 1. $P^2 = P$.
2. $P^T = P$.

- Take $V \leq \mathbb{R}^n$, there are two ways to obtain projection matrix.

1. Find an orthonormal basis $q_1, \dots, q_k$ of V.

$$P = q_1 q_1^T + \dots + q_k q_k^T$$

2. Find a basis $v_1, \dots, v_k$ of V : and let

- $A = [v_1, \dots, v_k] : \mathbb{R}^k \longrightarrow \mathbb{R}^n$ with
- $C(A) = V$, then $P = A(A^T A)^{-1} A^T$

Projection

Prop: P is a projection iff
1. $P^2 = P$.
2. $P^T = P$.

- Take $V \leq \mathbb{R}^n$, there are two ways to obtain projection matrix.

  1. Find an orthonormal basis $q_1, \ldots, q_k$ of V.

  $$P = q_1 q_1^T + \cdots + q_k q_k^T$$

  2. Find a basis $v_1, \ldots, v_k$ of V: and let
     - $A = [v_1, \ldots, v_k] : \mathbb{R}^k \longrightarrow \mathbb{R}^n$ with
     - $C(A) = V$, then $P = A(A^TA)^{-1}A^T$

- Both satisfy the two properties.

Projection

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

consider $v = Pv + (v - Pv)$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

consider $v = Pv + (v - Pv)$.

want it to be in $N(P^T)$ $\overset{\shortparallel}{N(P)}$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

consider $v = Pv + (v - Pv)$.

want it to be in $N(P^T)$ $\overset{\shortparallel}{N(P)}$.

• Take $P(v - Pv) = Pv - P^2 v = 0$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

consider $v = Pv + (v - Pv)$.

want it to be in $N(P^T)$ $=$ $N(P)$.

• Take $P(v - Pv) = Pv - P^2 v = 0$.

• Hence $Pv \in C(P)$, $v - Pv \in C(P)^\perp$.

Projection

Pf: (for the converse) Assuming $P^2 = P$, $P^T = P$.

Claim: • P is the projection onto $C(P)$.

consider $v = Pv + (v - Pv)$.

want it to be in $N(P^T)$ $\overset{\shortparallel}{N(P)}$.

• Take $P(v - Pv) = Pv - P^2 v = 0$.

• Hence $Pv \in C(P)$, $v - Pv \in C(P)^{\perp}$.

Therefore P is a projection.