1 dimensional case $W = \text{span}\{u\}$. Given v

Find the point $P_u v$ in W closest to v.

$$P_u v = \frac{u \cdot v}{u \cdot u} u = \underbrace{\frac{u u^T}{u^T u}}_{} v$$

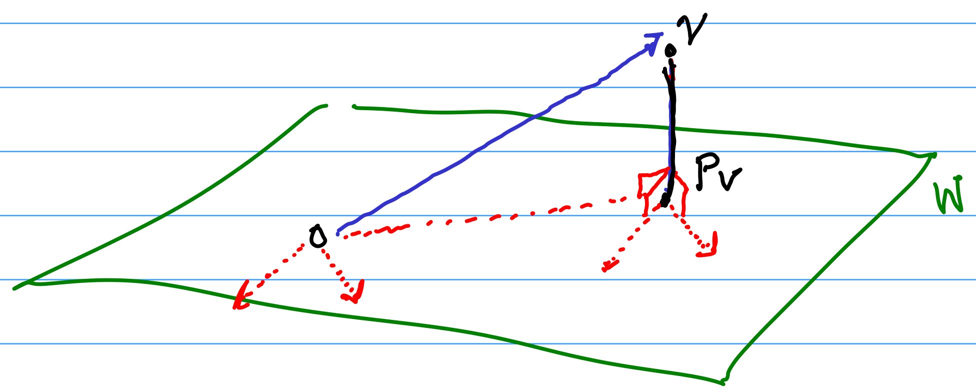
$$P_u = \frac{u u^T}{u^T u}$$

<u>Remark</u> If u is a unit vector then $u \cdot u = 1$ and

$$P_u = u u^T \quad \text{and} \quad P_u v = (u \cdot v) u$$

Move to higher dimensional case.

$$W = \text{span}\{u_1, u_2, \ldots, u_p\} \subseteq \mathbb{R}^n \quad \text{and} \quad v \in \mathbb{R}^n$$



Idea $(v - \underbrace{Pv}_{\text{find this vector.}}) \cdot w = 0$ for every $w \in W$

Note Pv is defined through a dot product.

Think about dot products in general

Suppose $y \cdot u = y \cdot v$ for all $y \in \mathbb{R}^n$

Claim $u = v$. Why?

Let $y = e_k$ where e_k is all zeros except with a 1 in the kth position.

Thus $e_k \cdot u = e_k \cdot v$ for every k

$$e_k \cdot u = e_k^T u = \begin{bmatrix} 0 & 0 & \cdots & \underset{\uparrow \text{ kth element}}{1} & \cdots & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = u_k$$

Therefore

$$u_k = v_k \quad \text{for every } k$$

and so $u = v$.

Idea $\underbrace{P_V}_{\text{find this vector..}} \cdot w = 0$ for every $w \in W = \operatorname{span}\{u_1, u_2, \ldots, u_p\}$

To make things simpler $\boxed{\text{assume}}$ $\{u_1, u_2, \ldots, u_p\}$ is an orthonormal basis.

① If it's not a basis then throw away vectors until you have a independent set that still spans W.

② If they are not orthogonal use the Gram-Schmidt algorithm to make a set of vectors which is orthogonal and still spans the same space.

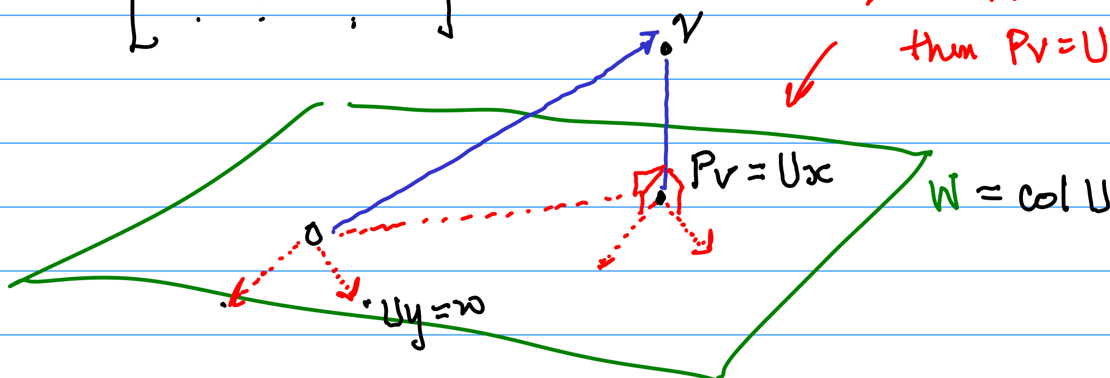
③ If the vectors are not unit vectors then normalized them by dividing them by their lengths.

Thus $\{u_1, u_2, \dots, u_p\}$ is an orthonormal basis

$$u_i \cdot u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$U = \left[u_1 \,\vdots\, u_2 \,\vdots\, \cdots \,\vdots\, u_p \right] \in \mathbb{R}^{n \times p}$$

Since $Pv \in W$ then $Pv = Ux$ for some x



$$w \cdot (v - Pv) = 0$$

$$Uy \cdot (v - Ux) = 0 \quad \text{for all } y \in \mathbb{R}^p$$

$$(Uy)^T (v - Ux) = 0$$

$$y^T U^T (v - Ux) = 0$$

$$y \cdot (U^T v - U^T U x) = 0$$

$$y \cdot U^T v = y \cdot x \quad \text{for all } y \in \mathbb{R}^p$$

Thus $U^T v = x$ and so $Pv = Ux = UU^T v$

so $P = UU^T$

$$P = \underset{n \times \cancel{p}}{U} \; \underset{\cancel{p} \times n}{U^T} \in \mathbb{R}^{n \times n}$$

one dimensional case

$$P_u = uu^T$$

$$P_v = UU^T v = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_p^T \end{bmatrix} v$$

row representation of matrix vector mult

$$= \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix} \begin{bmatrix} u_1 \cdot v \\ u_2 \cdot v \\ \vdots \\ u_p \cdot v \end{bmatrix}$$

column representation of matrix vector mult

$$= (u_1 \cdot v) u_1 + (u_2 \cdot v) u_2 + \cdots + (u_p \cdot v) u_p$$

$$= u_1 (u_1 \cdot v) + u_2 (u_2 \cdot v) + \cdots + u_p (u_p \cdot v)$$

$$= u_1 u_1^T v + u_2 u_2^T v + \cdots + u_p u_p^T v$$

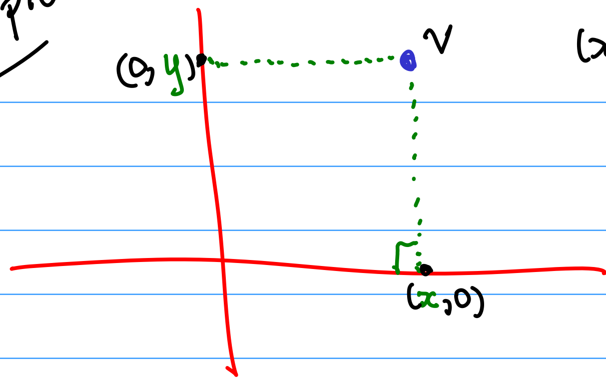
$$= P_{u_1} v + P_{u_2} v + \cdots P_{u_p} v$$

$$= (P_{u_1} + P_{u_1} + \cdots + P_{u_p}) v$$

Thus $P = UU^T = P_{u_1} + P_{u_1} + \cdots + P_{u_p}$

the projection onto W is the sum of 1 dimensional projections onto the different basis vectors.

2D picture



$$(x,y) = P_{e_1}v + P_{e_2}v$$
$$=$$

③ If the vectors are not unit vectors then normalized them by dividing them by their lengths.

Now $\left\{ \frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, \dots, \frac{u_p}{\|u_p\|} \right\}$ is an ortho-normal basis.

$$\|u_1\|^2 = u_1 \cdot u_1$$

$$P_v = \left(\frac{u_1}{\|u_1\|}\cdot v\right)\frac{u_1}{\|u_1\|} + \left(\frac{u_2}{\|u_2\|}\cdot v\right)\frac{u_2}{\|u_2\|} + \dots + \left(\frac{u_p}{\|u_p\|}\cdot v\right)\frac{u_p}{\|u_p\|}$$

$$= \frac{u_1\cdot v}{u_1\cdot u_1}u_1 + \frac{u_2\cdot v}{u_2\cdot u_2}u_2 + \dots + \frac{u_p\cdot v}{u_p\cdot u_p}u_p$$

$$= \frac{u_1u_1^T}{\|u_1\|^2}v + \frac{u_2u_2^T}{\|u_2\|^2}v + \dots + \frac{u_pu_p^T}{\|u_p\|^2}v$$

Let $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix}$ columns are orthogonal but not unit vectors

$D = \begin{bmatrix} 1/\|u_1\| & & & 0 \\ & 1/\|u_2\| & & \\ & & \ddots & \\ 0 & & & 1/\|u_p\| \end{bmatrix}$

$UD = \begin{bmatrix} \frac{u_1}{\|u_1\|} & \frac{u_2}{\|u_2\|} & \cdots & \frac{u_p}{\|u_p\|} \end{bmatrix}$ ← orthogonal matrix columns are orthonormal.

$$P = UD(UD)^T = UDD^TU^T = UD^2U^T$$