- Orthogonal bases

Definition) A basis $\vec{v}_1, \dots, \vec{v}_n$ of a vector space V is an orthogonal basis if the vectors are pairwise orthogonal.

Example) The standard basis of $\mathbb{R}^3$, $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$ is an orthogonal basis.

Example) Suppose $\vec{v}_1, \dots, \vec{v}_n$ is an orthogonal basis of V.
Then, find $c_1, \dots, c_n \in \mathbb{R}$ for $\vec{w} \in V$ such that
$\vec{w} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n$ ——— $(*)$

$\Rightarrow$ Take the dot product of $\vec{v}_i$ on both sides of $(*)$

$$\vec{v}_i \cdot \vec{w} = c_i \|\vec{v}_i\|^2$$
$$\downarrow$$
$$c_i = \frac{\vec{v}_i \cdot \vec{w}}{\vec{v}_i \cdot \vec{v}_i}$$

Definition) A basis $\vec{v}_1, \dots, \vec{v}_n$ of V is an orthonormal basis if the vectors are orthogonal and have length 1.

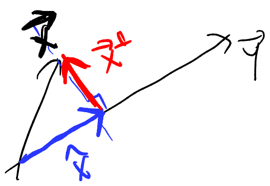
- Orthogonal Projection

Definition) The orthogonal projection of $\vec{x}$ onto $\vec{y}$ is

$$\hat{x} \triangleq \left(\frac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}\right)\vec{y}$$

- Geometric interpretation of the orthogonal projection.

$$\hat{x} = \frac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}\vec{y}$$



i) the vector $\hat{x}$ is the closest vector to $\vec{x}$ which is in span$\{\vec{y}\}$.

ii) The error $\vec{x}^{\perp} = \vec{x} - \hat{x}$ is orthogonal to span$\{\vec{y}\}$. ($\Leftrightarrow$ is in Nul$\{\vec{y}\}$)

iii) $\hat{x}$ can be derived from i) and ii).

1) $\hat{x}$ can be written as $c\cdot\vec{y}$.

2) Then the error $(\vec{x} - \hat{x}) \in$ Nul$\{\vec{y}\}$.

$\rightarrow$ $(\vec{x} - \hat{x})\cdot\vec{y} = 0$

$(\vec{x} - c\vec{y})\cdot\vec{y} = 0$

$\vec{x}\cdot\vec{y} - c\cdot\vec{y}\cdot\vec{y} = 0$ $\quad\therefore\ c = \dfrac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}$

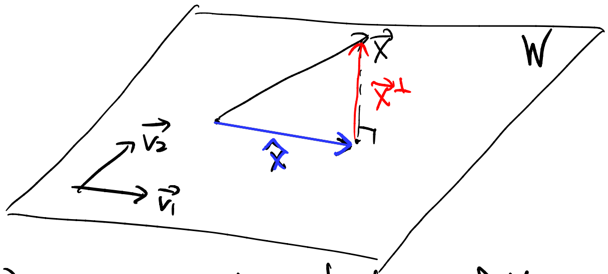
- Orthogonal projection on subspaces.

Theorem) Let W be a subspace of $\mathbb{R}^n$. Then each $\vec{x} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{x} = \underbrace{\hat{x}}_{\in W} + \underbrace{\vec{x}^{\perp}}_{\in W^{\perp}}$$

$\hat{x}$ is the orthogonal projection of $\vec{x}$ onto W.

↳ $\hat{x}$ is the point in W closest to $\vec{x}$. For any other $\vec{y} \in W$, $\mathrm{dist}(\vec{x}, \hat{x}) < \mathrm{dist}(\vec{x}, \vec{y})$



If $\vec{v}_1, \dots, \vec{v}_m$ is an orthogonal basis of W. Then,

$$\hat{x} = \left(\frac{\vec{x} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 + \cdots + \left(\frac{\vec{x} \cdot \vec{v}_m}{\vec{v}_m \cdot \vec{v}_m}\right)\vec{v}_m$$

Once, $\hat{x}$ is determined, $\vec{x}^{\perp} = \vec{x} - \hat{x}$.

First, we will show that,

Let W be a subspace of a vector space V and $\vec{v} \in V$.

Then, $\vec{w} \in W$ is closest to $\vec{v}$ iff $(\vec{v} - \vec{w}) \in W^{\perp}$.

a) If $(\vec{v} - \vec{w}) \in W^{\perp}$, then $\|\vec{v} - \vec{w}\| \leq \|\vec{v} - \vec{u}\|$ for all $\vec{u} \in W$ and $\|\vec{v} - \vec{w}\| = \|\vec{v} - \vec{u}\|$ iff $\vec{w} = \vec{u}$.

→ First, remark that

$$\|\vec{v} - \vec{w}\| < \|\vec{v} - \vec{u}\| \iff \|\vec{v} - \vec{w}\|^2 < \|\vec{v} - \vec{u}\|^2$$

Hence,

$$\begin{aligned} \|\vec{v} - \vec{u}\|^2 &= \|(\vec{v} - \vec{w}) + (\vec{w} - \vec{u})\|^2 \qquad (\vec{v} \in V,\ \vec{w} \in W,\ \vec{u} \in W,\ \vec{v} - \vec{w} \in W^{\perp}) \\ &= \|\vec{v} - \vec{w}\|^2 + \|\vec{w} - \vec{u}\|^2 + 2 \cdot \underbrace{\langle \vec{v} - \vec{w}, \vec{w} - \vec{u} \rangle}_{=0} \\ &\geq \|\vec{v} - \vec{w}\|^2 \end{aligned}$$

where the equality holds when $\|\vec{w} - \vec{u}\| = 0 \iff \vec{w} = \vec{u}$

(b) If $\vec{w} \in W$ is the closest to $\vec{v} \in V$, then $\vec{v} - \vec{w} \in W^{\perp}$

→ We know that $\|\vec{v} - \vec{w}\|^2 \leq \|\vec{v} - \vec{u}\|^2$ for all $\vec{u} \in W$.

Therefore the function $f: \mathbb{R} \to \mathbb{R}$,

$$F(t) \triangleq \|\vec{v} - \vec{w} + t\vec{x}\|^2 \qquad (\vec{x} \in W)$$

has a minimum at $t=0$ for all $\vec{x} \in W$.

⇓

$$\begin{aligned} F(t) &= \langle \vec{v} - \vec{w} + t\vec{x}, \sim \rangle \\ &= \|\vec{v} - \vec{w}\|^2 + 2t\langle \vec{v} - \vec{w}, \vec{x} \rangle + t^2\|\vec{x}\|^2 \end{aligned}$$

and

$$F'(t) = 2\langle \vec{v} - \vec{w}, \vec{x} \rangle + 2t\|\vec{x}\|^2 = 0 \Big|_{t=0}$$

→ $\underline{\langle \vec{v} - \vec{w}, \vec{x} \rangle = 0 \text{ for all } \vec{x} \in W}$

→ $\vec{v} - \vec{w} \in W^{\perp}$

Example) $W = \text{span}\left\{ \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, $\vec{x} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$,

Find the orthogonal projection of $\vec{x}$ onto W.

$$\rightarrow \quad \hat{x} = \frac{\vec{x}\cdot\vec{w}_1}{\|\vec{w}_1\|^2}\vec{w}_1 + \frac{\vec{x}\cdot\vec{w}_2}{\|\vec{w}_2\|^2}\vec{w}_2 = \frac{10}{10}\cdot\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \frac{3}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

$$\vec{x}^{\perp} = \vec{x} - \hat{x} = \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix} \qquad \because\ \hat{x}\cdot\vec{x}^{\perp} = 0$$

$$\hat{x} + \vec{x}^{\perp} = \vec{x}$$

Definition) Let $\vec{v}_1, \dots, \vec{v}_m$ be an orthogonal basis of W, a subspace of $\mathbb{R}^n$.
Then, the projection map $\pi_W : \mathbb{R}^n \rightarrow W$ is given by

$$\pi_W : \underset{\in \mathbb{R}^n}{\vec{x}} \mapsto \underset{\in W}{\hat{x}} = \left(\frac{\vec{x}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \cdots + \frac{\vec{x}\cdot\vec{v}_m}{\vec{v}_m\cdot\vec{v}_m}\vec{v}_m \right).$$

* π_W is linear!

$$\rightarrow \pi_W(\alpha\vec{x} + \beta\vec{y}) = \left(\frac{(\alpha\vec{x}+\beta\vec{y})\cdot\vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots \right) = \alpha\left(\frac{\vec{x}\cdot\vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots \right) + \beta\left(\frac{\vec{y}\cdot\vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots \right)$$

$$= \alpha\pi_W(\vec{x}) + \beta\pi_W(\vec{y})$$

* The matrix P representing π_W w.r.t. the standard basis is called the projection matrix.

Example) Find the projection matrix P of $W = \text{span}\left\{ \underset{\vec{w}_1}{\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}}, \underset{\vec{w}_2}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \right\} \subseteq \mathbb{R}^3$

$$\pi_W : V \rightarrow W$$

$\rightarrow$ Standard basis of $\mathbb{R}^3$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

i) The first column of P : projection of $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\frac{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\vec{w}_1 + \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\vec{w}_2 = \begin{bmatrix} 9/10 \\ 0 \\ 3/10 \end{bmatrix} \qquad \rightarrow P = \begin{bmatrix} \frac{9}{10} & \square & \square \\ 0 & \square & \square \\ 3/10 & \square & \square \end{bmatrix}$$

ii)

$$\frac{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\cdot\vec{w}_1}{\cancel{\vec{w}_1\cdot\vec{w}_1}}\vec{w}_1 + \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\cdot\vec{w}_2}{\cancel{\vec{w}_2\cdot\vec{w}_2}}\vec{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \rightarrow P = \begin{bmatrix} 0.9 & 0 & \cdot \\ 0 & 1 & \cdot \\ 0.3 & 0 & \cdot \end{bmatrix}$$

iii)

$$\frac{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\cdot\vec{w}_1}{\cancel{\vec{w}_1\cdot\vec{w}_1}}\vec{w}_1 + \frac{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\cdot\vec{w}_2}{\cancel{\vec{w}_2\cdot\vec{w}_2}}\vec{w}_2 = \begin{bmatrix} 0.3 \\ 0 \\ 0.1 \end{bmatrix}$$

$$\therefore P = \begin{bmatrix} 0.9 & 0 & 0.3 \\ 0 & 1 & 0 \\ 0.3 & 0 & 0.1 \end{bmatrix}$$

Example) $P^2 = P$?

$\rightarrow$ Yes $\because \forall x \in \mathbb{R}^n$, $P^2\vec{x} = P\vec{x}$.

• Least squares.

Definition) $\hat{x}$ is a least squares solution of the system $A\vec{x} = \vec{b}$ if $(A\vec{x} - \vec{b})$ is as small as possible

Idea, $A\vec{x} = \vec{b}$ is consistent, $\Leftrightarrow$ $\vec{b} \in \mathrm{Col}(A)$

If $A\vec{x} = \vec{b}$ is inconsistent,

1) We replace $\vec{b}$ with its projection $\hat{b}$ onto $\mathrm{Col}(A)$

2) and solve $A\vec{x} = \hat{b}$ $\leftarrow$ always consistent.

• The normal equations

Theorem) $\hat{x}$ is a least squares solution of $A\vec{x} = \vec{b}$,

$\Leftrightarrow$ $A^TA\hat{x} = A^T\vec{b}$ — (*)

(the normal equation)

$\Rightarrow$ $\hat{x}$ is a LS solution of $A\vec{x} = \vec{b}$

$\Leftrightarrow$ $(A\hat{x} - \vec{b})$ is as small as possible,

$\Leftrightarrow$ $(A\hat{x} - \vec{b}) \perp \mathrm{Col}(A)$

$\Leftrightarrow$ $(A\hat{x} - \vec{b}) \in \mathrm{Nul}(A^T)$

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

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