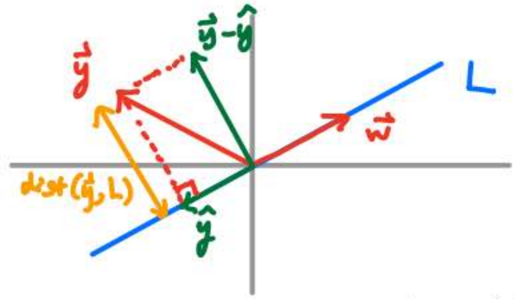
Lecture 31 : Orthogonality

Projections & Reflections



To find the point $\hat{y}$ on $L = \text{span}(\vec{w})$ closest to $\vec{y} \in \mathbb{R}^2$, we can use

Calculus : this is where $\text{dist}(\vec{y}, t\vec{w})$, or equivalently, $\text{dist}(\vec{y}, t\vec{w})^2$ reaches its minimum :

$$0 = \frac{d}{dt}\,\text{dist}(\vec{y}, t\vec{w})^2 = \frac{d}{dt}(\vec{y}-t\vec{w})\cdot(\vec{y}-t\vec{w}) = \frac{d}{dt}\left\{\|\vec{y}\|^2 - 2t\,\vec{y}\cdot\vec{w} + t^2\|\vec{w}\|^2\right\}$$

$$= -2\vec{y}\cdot\vec{w} + 2t\|\vec{w}\|^2$$

$$\Rightarrow \quad t = \frac{\vec{y}\cdot\vec{w}}{\|\vec{w}\|^2} = \frac{\vec{y}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}. \qquad \text{Set} \quad \hat{y} := \frac{\vec{y}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\,\vec{w}.$$

Then we have

$$\vec{w}\cdot(\vec{y}-\hat{y}) = \vec{w}\cdot\vec{y} - \vec{w}\cdot\hat{y} = \vec{w}\cdot\vec{y} - \cancel{\frac{\vec{y}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\,\vec{w}\cdot\vec{w}} = 0,$$

and writing

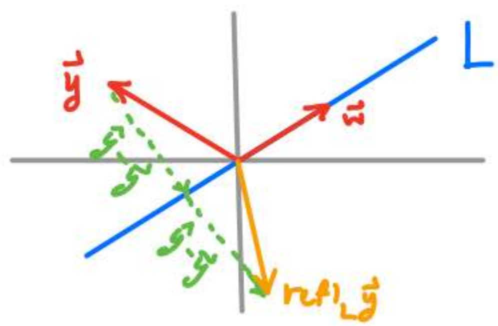
$$\vec{y} = \hat{y} + (\vec{y}-\hat{y}) =: \text{proj}_L(\vec{y}) + (\vec{y} - \text{proj}_L(\vec{y}))$$

decomposes $\vec{y}$ into a vector along L and a vector $\perp$ to L.

This also allows us to find a formula for the reflection of $\vec{y}$ about L :

$$\text{refl}_L\vec{y} := \vec{y} + 2(\hat{y}-\vec{y}) = 2\hat{y} - \vec{y}$$

$$= 2\text{proj}_L(\vec{y}) - \vec{y}.$$



More generally, in $\mathbb{R}^n$, to project $\vec{y}$ to $L = \text{Span}\{\vec{w}\}$, we want $\vec{y} = \hat{y} + \vec{z}$, where $\vec{z} \perp \vec{w}$ and $\hat{y} = \alpha \vec{w}$. Dotting both sides with $\vec{w}$,

which is $\text{proj}_L \vec{y}$

$$\vec{w} \cdot \vec{y} = \vec{w} \cdot \hat{y} + \vec{w} \cdot \vec{z} = \vec{w} \cdot (\alpha \vec{w}) + 0 = \alpha\, \vec{w} \cdot \vec{w}$$

$$\Rightarrow \alpha = \frac{\vec{w} \cdot \vec{y}}{\vec{w} \cdot \vec{w}} \Rightarrow \boxed{\text{proj}_L \vec{y} := \frac{\vec{w} \cdot \vec{y}}{\vec{w} \cdot \vec{w}} \vec{w} \quad \text{and } \vec{z} = \vec{y} - \text{proj}_L \vec{y}.}$$

Moreover, we set

$$\text{dist}(\vec{y}, L) := \|\vec{z}\| = \|\vec{y} - \text{proj}_L(\vec{y})\| = \text{dist}(\vec{y}, \text{proj}_L \vec{y}).$$

Ex 1/ $\vec{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$, $L = \text{Span}\{\vec{w}\}$

$$(\hat{y} =)\, \text{proj}_L \vec{y} = \frac{\begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 6 \end{pmatrix}}{\begin{pmatrix} 8 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 6 \end{pmatrix}} \begin{pmatrix} 8 \\ 6 \end{pmatrix} = \frac{30}{100} \begin{pmatrix} 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 1.8 \end{pmatrix}$$

and

$$\vec{z} = \vec{y} - \hat{y} = \begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix} \Rightarrow \text{dist}(\vec{y}, L) = \left\| \begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix} \right\| = \sqrt{0.6^2 + (-0.8)^2} = 1.$$

We also have the decomposition of $\vec{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 2.4 \\ 1.8 \end{pmatrix}}_{\text{along } L} + \underbrace{\begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix}}_{\perp \text{ to } L}$.

//

Orthogonal bases

Let $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ be nonzero vectors. We call $\{\vec{v}_1, \dots, \vec{v}_k\}$ an orthogonal set if $\vec{v}_i \perp \vec{v}_j$ for all $i \neq j$.

Theorem: Orthogonal sets are linearly independent.

Proof: Suppose $\vec{0} = c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k$. Then for any j,

$$0 = \vec{v}_j \cdot (c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k) = c_1 \cancelto{0}{\vec{v}_j \cdot \vec{v}_1} + \cdots + c_j \vec{v}_j \cdot \vec{v}_j + \cdots + c_k \cancelto{0}{\vec{v}_j \cdot \vec{v}_k}$$
$$= c_j \|\vec{v}_j\|^2$$

$\Longrightarrow c_j = 0.$ Since j was arbitrary, all the c_i's are 0. $\square$

Ex 2/ Are $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ independent?

Check: $\vec{v}_1 \cdot \vec{v}_2 = 0$, $\vec{v}_1 \cdot \vec{v}_3 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$

$\Longrightarrow$ orthogonal $\Rightarrow$ independent! //

Corollary: An orthogonal set in $\mathbb{R}^n$ with n elements is a basis. (We call this an orthogonal basis.)

To find the coordinate vector of $\vec{y}$ with respect to an orthogonal basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ of $\mathbb{R}^n$, write

$$\vec{y} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n \Longrightarrow \vec{v}_j \cdot \vec{y} = c_1 \cancelto{0}{\vec{v}_j \cdot \vec{v}_1} + \cdots + c_j \vec{v}_j \cdot \vec{v}_j + \cdots + c_n \cancelto{0}{\vec{v}_j \cdot \vec{v}_n}$$
$$= c_j \|v_j\|^2$$

$$\Longrightarrow c_j = \frac{\vec{v}_j \cdot \vec{y}}{\|\vec{v}_j\|^2}$$

$$\Longrightarrow [\vec{y}]_{\mathcal{B}} = \begin{pmatrix} \vec{v}_1 \cdot \vec{y} / \|\vec{v}_1\|^2 \\ \vdots \\ \vec{v}_n \cdot \vec{y} / \|\vec{v}_n\|^2 \end{pmatrix},$$

that is,

$$\boxed{\vec{y} = \left(\frac{\vec{v}_1 \cdot \vec{y}}{\vec{v}_1 \cdot \vec{v}_1}\right) \vec{v}_1 + \cdots + \left(\frac{\vec{v}_n \cdot \vec{y}}{\vec{v}_n \cdot \vec{v}_n}\right) \vec{v}_n.}$$

"Fourier decomposition" of $\vec{y}$

(You can use this to recover the formulas for $\hat{y}$ & $\vec{z}$ from above, by taking $\vec{v}_1 = \vec{w}$.)

Ex 3/ $\vec{v}_1, \vec{v}_2, \vec{v}_3$ as in Ex. 2, $y = \begin{pmatrix} 8 \\ -4 \\ -3 \end{pmatrix}$

$$\Rightarrow \vec{y} = \frac{5}{2}\vec{v}_1 + \left(\frac{-3}{2}\right)\vec{v}_2 + 2\vec{v}_3 .$$

//

Orthonormal bases

often abbreviated o.n.

The set $\{\vec{u}_1, \dots, \vec{u}_k\}$ is orthonormal $\overset{\text{def.}}{\iff}$ it is orthogonal and $\vec{u}_j \cdot \vec{u}_j = 1$ $(\forall j)$.

We can write $\vec{u}_i \cdot \vec{u}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. If $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal basis, then $(*)$ simplifies to

$$\vec{y} = (\vec{u}_1 \cdot \vec{y})\,\vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{y})\,\vec{u}_n .$$

An easy example of an orthonormal basis is $\{\vec{e}_1, \dots, \vec{e}_n\}$ in $\mathbb{R}^n$.

How do we get more such bases?

Ex 4/ $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ is orthogonal.

Their lengths are $\|\vec{v}_1\| = \sqrt{1^2+2^2+1^2} = \sqrt{6}$, $\|\vec{v}_2\| = \sqrt{3}$, $\|\vec{v}_3\| = \sqrt{2}$.

So $\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\vec{u}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$, $\vec{u}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ is an o.n. basis. //

Orthogonal matrices

An $n \times n$ matrix U is orthogonal $\overset{\text{def.}}{\iff}$ its columns form an orthonormal basis of $\mathbb{R}^n$. Equivalently,

$$(**) \qquad U^T U = I_n \qquad \text{(why?)}.$$

<u>Properties:</u> Let U be orthogonal. Then

(1) U is invertible (b/c columns are independent)

(2) U preserves the dot product: $U\vec{x} \cdot U\vec{y} = \vec{x} \cdot \vec{y}$.

(note this has $\|U\vec{x}\| = \|\vec{x}\|$ as a special case.)

(3) U has orthonormal rows (b/c $UU^T = I_n$)

(4) $\det(U) = \pm 1$ (take det of both sides of (**))

Ex 5/ Continuing Ex. 4, $U = \begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{3} & 1/\sqrt{2} \\ 2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 1/\sqrt{6} & -1/\sqrt{3} & -1/\sqrt{2} \end{pmatrix}$ is orthogonal. //

So why does (2) hold?

$$(U\vec{x}) \cdot (U\vec{y}) = (U\vec{x})^T (U\vec{y}) = \vec{x}^T U^T U \vec{y} = \vec{x}^T I_n \vec{y}$$
$$= \vec{x}^T \vec{y} = \vec{x} \cdot \vec{y}.$$

In closing, note that we can define orthogonality for transformations more abstractly: $T : \mathbb{R}^n \to \mathbb{R}^n$ is <u>orthogonal</u> if it preserves length, i.e. $\|T(\vec{x})\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$. You could try to show that

$$T \text{ is orthogonal} \iff [T]_{\mathcal{E}} =: U \text{ is orthogonal.}$$

A more difficult fact is that <u>all orthogonal transformations are products of rotations and reflections</u> (this is only easy to see in $\mathbb{R}^2$).