Sec. 3.6: Orthogonal bases

① Why are orthogonal bases better than generic bases?

In $\mathbb{R}^2$, vectors are orthogonal (perpendicular) when their dot product is 0:

$$(\underline{v}_1 \perp \underline{v}_2) \Longleftrightarrow (\underline{v}_1^T \underline{v}_2 = 0).$$

The same definition carries over to $\mathbb{R}^n$:

Def: A set $\{\underline{v}_1, \dots \underline{v}_p\}$ of vectors in $\mathbb{R}^n$ is orthogonal if each pair of distinct vectors is orthogonal: $\boxed{\underline{v}_i^T \underline{v}_j = 0 \text{ for } i \neq j.}$

See Ex. 1 in book for numbers.

Note: Our intuitive notion that $(\underline{v}_1 \perp \underline{v}_2) \Longleftrightarrow (\underline{v}_2 \perp \underline{v}_1)$ is supported by the above definition. I.e., we want to show: $(\underline{v}_1^T \cdot \underline{v}_2 = 0) \Rightarrow (\underline{v}_2^T \underline{v}_1 = 0)$. Indeed:

$$(\underline{v}_1^T \underline{v}_2 = 0)^T \Rightarrow (\underline{v}_1^T \underline{v}_2)^T = 0^T \underset{\substack{\uparrow \\ \text{Thm. 10 of} \\ \text{Chap. 1}}}{\Rightarrow} \underline{v}_2^T (\underline{v}_1^T)^T = \underset{\substack{\uparrow \\ \text{scalar}}}{0}$$

$$\Rightarrow \underline{v}_2^T \underline{v}_1 = 0. \quad \checkmark$$

Thm. 13: If $\{\underline{v}_1, \dots \underline{v}_p\}$ = orthogonal set

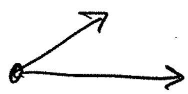
$$\Rightarrow$$

$$\{\underline{v}_1, \dots \underline{v}_p\} = \text{lin. indep. set}$$

Illustration:

← This is clearly lin. indep.

Note that ("lin. indep.") $\not\Rightarrow$ "orthogonal" in general:



← This is lin. indep., but not orthogonal.

Given:

$$\underline{v}_i^T \underline{v}_j = 0 \text{ for } i \neq j$$

Want:

$$c_1 \underline{v}_1 + \dots + c_p \underline{v}_p = \underline{\theta}$$
$$\Rightarrow c_1 = \dots = c_p = 0.$$

Proof:

1) $\underline{v}_1^T (c_1 \underline{v}_1 + c_2 \underline{v}_2 + \dots + c_p \underline{v}_p = \underline{\theta})$

$$c_1 (\cancel{\underline{v}_1^T \underline{v}_1}) + c_2 (\cancel{\underline{v}_1^T \underline{v}_2}) + \dots + c_p (\cancel{\underline{v}_1^T \underline{v}_p}) = (\cancel{\underline{v}_1^T \underline{\theta}}) \rightarrow 0$$

$\|\underline{v}_1\|^2$ ← Sec. 1.6, p. 68 $\qquad$ 0 ← Given → 0

$$c_1 \cdot \underbrace{\|\underline{v}_1\|^2}_{\neq 0} + 0 + \dots + 0 = 0 \quad \Rightarrow \quad c_1 = 0$$

2) If we multiply by $\underline{v}_2^T$, we get $c_2 = 0$.

Similarly, all $c_1 = c_2 = \dots = c_p = 0$.

q.e.d.

Claim: It is much easier to find coordinates of a vector in an orthogonal basis than in a generic basis.

- To find coordinates of $\underline{x}$ in a generic basis $\{\underline{v}_1, \dots \underline{v}_p\}$:

$$\underline{x} = c_1 \underline{v}_1 + \dots + c_p \underline{v}_p$$

$$\underline{x} = [\underline{v}_1, \dots \underline{v}_p] \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}$$

$$\Updownarrow$$

$$V \underline{c} = \underline{x} \quad \Rightarrow \text{solve by REF.}$$

Amount of calculation grows rapidly with p.

- To find coordinates of $\underline{x}$ in an orthogonal basis:

derivation

$$\underline{v}_1^T \left(c_1 \underline{v}_1 + c_2 \underline{v}_2 + \dots + c_p \underline{v}_p = \underline{x} \right)$$

$$c_1 \underbrace{(\underline{v}_1^T \underline{v}_1)}_{\|\underline{v}_1\|^2} + c_2 \underbrace{(\underline{v}_1^T \underline{v}_2)}_{0} + \dots + c_p \underbrace{(\underline{v}_1^T \underline{v}_p)}_{0} = \underline{v}_1^T \underline{x}$$

← As in the Proof of Thm. 13

RESULT

$$\Rightarrow \boxed{c_1 = \frac{\underline{v}_1^T \underline{x}}{\|\underline{v}_1\|^2}} \quad \text{Similarly:} \quad \boxed{c_2 = \frac{\underline{v}_2^T \underline{x}}{\|\underline{v}_2\|^2}} \quad \dots \quad \boxed{c_p = \frac{\underline{v}_p^T \underline{x}}{\|\underline{v}_p\|^2}}$$

Coordinates of $\underline{x}$ in an orthogonal basis.

See Ex. 4 in book for numbers.

A simplification occurs if the lengths of all vectors in an orthogonal basis is 1. Hence:

Def: An **orthonormal** basis is:

"⊥" "length"

- an orthogonal basis, where
- lengths of all basis vectors is 1:

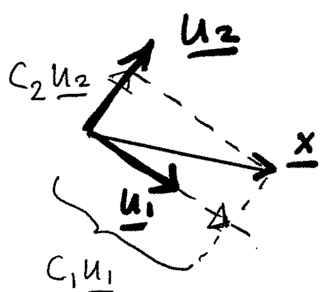
$$\|\underline{u}_i\| = 1 \text{ for } i = 1, \dots p.$$

② Projection of $\underline{x}$ on $\underline{v}$

Def: Let $\{\underline{v}_1, \dots \underline{v}_p\}$ be an **orthogonal** basis; then $\quad \underline{x} = \underbrace{c_1 \underline{v}_1} + \underbrace{c_2 \underline{v}_2} + \dots + \underbrace{c_p \underline{v}_p}$.

Projections of $\underline{x}$ on $\underline{v}_1, \dots \underline{v}_p$ $\rightarrow P_{\underline{v}_1}(\underline{x}) \quad P_{\underline{v}_2}(\underline{x}) \ \cdots \ P_{\underline{v}_p}(\underline{x})$

- Aside note: Geometric meaning of coordinates in an orthonormal basis:



They are the lengths of projections of $\underline{x}$ on the unit basis vectors $\underline{u}_1, \dots \underline{u}_p$.

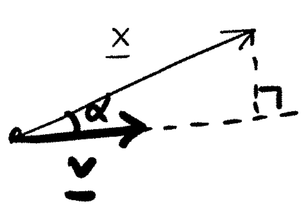
- Above we have derived formulas for coordinates in an orthogonal basis. So we can now write formulas for projections:

$$\boxed{P_{\underline{v}_1}(\underline{x}) = c_1 \underline{v}_1 = \frac{\underline{v}_1^T \underline{x}}{\|\underline{v}_1\|^2} \underline{v}_1}, \text{ etc.}$$

We can write the same formula for **projection of $\underline{x}$ on any one given vector $\underline{v}$:**

$$\boxed{P_{\underline{v}}(\underline{x}) = \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|^2} \cdot \underline{v}}$$

Derivation based on Calculus: $\quad P_{\underline{v}}(\underline{x}) = \underbrace{\|\underline{x}\| \cos\alpha}_{\text{length}} \cdot \underbrace{\frac{\underline{v}}{\|\underline{v}\|}}_{\text{unit vector along } \underline{v}}$



$$= \frac{\|\underline{x}\| \cdot \|\underline{v}\| \cdot \cos\alpha}{\|\underline{v}\|} \cdot \frac{\underline{v}}{\|\underline{v}\|} = \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|} \cdot \frac{\underline{v}}{\|\underline{v}\|}. \ ✓$$

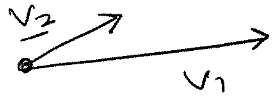
dot product of $\underline{x}$ & $\underline{v}$

③ How to construct an **orthonormal** basis from a generic basis

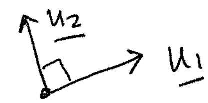
(The Gram-Schmidt orthogonalization)

In $\mathbb{R}^2$:

Given:



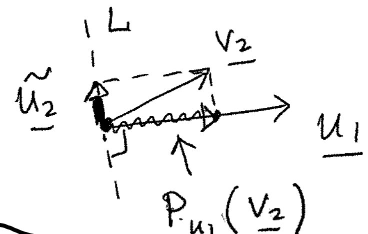
Want:



- $\underline{u}_1 \perp \underline{u}_2$
- $\|\underline{u}_1\| = \|\underline{u}_2\| = 1$
- $\underline{u}_1, \underline{u}_2$ are "made" from $\underline{v}_1, \underline{v}_2$.

Step 1: Construct $\underline{u}_1$: $\underline{u}_1 = \dfrac{\underline{v}_1}{\|\underline{v}_1\|}$.

Step 2: Construct $\underline{u}_2$:



- Draw line $L \perp \underline{u}_1$
- Write $\underline{v}_2$ as:

$$\underset{\text{given}}{\underline{v}_2} = \underbrace{P_{\underline{u}_1}(\underline{v}_2)}_{\text{find using formula}} + \underset{\text{want}}{\tilde{\underline{u}}_2}$$

ALONG line L

Simplified formula

$\Rightarrow \tilde{\underline{u}}_2 = \underline{v}_2 - \underbrace{P_{\underline{u}_1}(\underline{v}_2)}_{\frac{\underline{u}_1^T \underline{v}_2}{\|\underline{u}_1\|^2} \cdot \underline{u}_1 \,\equiv\, (\underline{u}_1^T \underline{v}_2) \cdot \underline{u}_1}$; $\tilde{\underline{u}}_2 \perp \underline{u}_1$ by design.

(since $\|\underline{u}_1\| = 1$)

$$\boxed{\text{So, when } \|\underline{u}\| = 1, \quad P_{\underline{u}}(\underline{v}) = \underline{u} \cdot (\underline{u}^T \underline{v}).}$$

- Make a unit $\underline{u}_2$: $\underline{u}_2 = \tilde{\underline{u}}_2 / \|\tilde{\underline{u}}_2\|$.

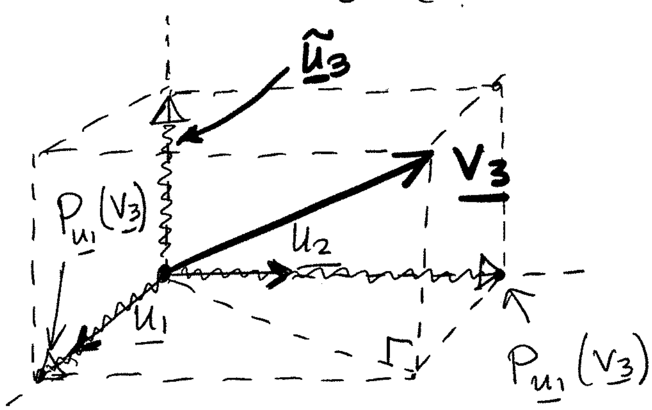
In $\mathbb{R}^3$: Given basis $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$, construct an orthonormal basis $\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$.

Step 1: Same as in $\mathbb{R}^2$: $\underline{u}_1 = \underline{v}_1 / \|\underline{v}_1\|$.

Step 2: $\underline{v_2}$ and $\underline{u_1}$ are in the same plane (because any two vectors are always in the same plane). Therefore, can apply the same process as in $\mathbb{R}^2$ to "make" $\underline{u_2}$:

- $\tilde{\underline{u_2}} = \underline{v_2} - P_{\underline{u_1}}(\underline{v_2})$
- $\underline{u_2} = \tilde{\underline{u_2}} / \|\tilde{\underline{u_2}}\|$.

Step 3 We now have 2 unit orthogonal vectors $\underline{u_1}$ & $\underline{u_2}$, and we can think of them as "our" $\vec{i}$ and $\vec{j}$ (unit coordinate vectors in 3D).



Write

$$\underline{v_3} = P_{\underline{u_1}}(\underline{v_3}) + P_{\underline{u_2}}(\underline{v_3}) + \tilde{\underline{u_3}}$$

Given

formula

$\perp$ to plane made by $\underline{u_1}, \underline{u_2}$

$$\Rightarrow \boxed{\tilde{\underline{u_3}} = \underline{v_3} - P_{\underline{u_1}}(\underline{v_3}) - P_{\underline{u_2}}(\underline{v_3})}$$

- Make a unit $\underline{u_3}$: $\underline{u_3} = \tilde{\underline{u_3}} / \|\tilde{\underline{u_3}}\|$.

Note: Ex. 5 & 6 in book use an equivalent process, but they do not normalize their vectors $\tilde{\underline{u_2}}, \tilde{\underline{u_3}}$ etc. (and do not use the "$\sim$"). You may use either the Notes' or the book's process.