Linear Algebra I Fall 2023

§12 Orthonormal bases & Gram-Schmidt algorithm

Recall: 1) If $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$, then the dot-product is

$$x \cdot y = x_1 y_1 + \dots + x_n y_n .$$

2) $x, y \in \mathbb{R}^n$ are <u>orthogonal</u> if $x \cdot y = 0$.

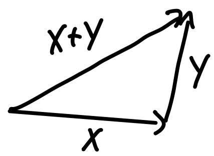
3) The <u>norm</u> of $x \in \mathbb{R}^n$ is $\|x\| = \sqrt{x \cdot x}$.

Properties: 1) $\|\lambda x\| = |\lambda| \|x\|$, $\lambda \in \mathbb{R}$, $x \in \mathbb{R}^n$

2) $|x \cdot y| \leq \|x\| \cdot \|y\|$, $x, y \in \mathbb{R}^n$ (Cauchy-Schwarz inequality)

"$\overset{\uparrow}{=}$" if x, y are lin. dep.

3) $\|x+y\| \leq \|x\| + \|y\|$,



(Triangle inequality)

Definition 12.1 i) A vector $u \in \mathbb{R}^n$ is called a <u>unit vector</u> if $\|u\| = 1$.

ii) Every vector $u \in \mathbb{R}^n$ with $u \neq 0$ can be normalized by

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

$\rightsquigarrow$ $\hat{u}$ is a unit vector since $\|\hat{u}\| = \left\| \frac{1}{\|u\|} u \right\| = \frac{1}{\|u\|} \|u\| = 1$.

iii) Vectors $u_1, \dots, u_\ell \in \mathbb{R}^n$ are called <u>orthonormal</u> (ON) if

$$u_i \cdot u_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \qquad 1 \leq i, j \leq \ell .$$

Definition 12.2 A basis $\mathcal{B} = (b_1, \dots, b_m)$ of a subspace U is called an <u>orthonormal basis (ONB)</u> of U, if $b_1, \dots, b_m$ are orthonormal.

Example 39: The standard basis $(e_1, \dots, e_n)$ of $\mathbb{R}^n$ is an ONB.

Proposition 12.3 i) If $v_1, \dots, v_m \in \mathbb{R}^n$ are orthonormal then they are linearly independent.

ii) Let $\mathcal{B} = (v_1, \dots, v_m)$ be an ONB of $V \subset \mathbb{R}^n$ and $u \in V$. Then

$$[u]_{\mathcal{B}} = \begin{pmatrix} u \cdot v_1 \\ \vdots \\ u \cdot v_m \end{pmatrix} \in \mathbb{R}^m,$$

i.e. $u = \sum_{i=1}^{m} (u \cdot v_i)\, v_i$.

iii) If $\mathcal{B} = (v_1, \dots, v_m)$ is an ONB of $V \subset \mathbb{R}^n$ and $u, v \in V$, then

$$u \cdot v = [u]_{\mathcal{B}} \cdot [w]_{\mathcal{B}}.$$

Proof: i) Assume $v_1, \dots, v_m$ are ON and

$$\lambda_1 v_1 + \dots + \lambda_m v_m = 0.$$

For all $1 \le j \le m$ we get $v_j \bullet (\lambda_1 v_1 + \dots + \lambda_m v_m) = 0$

$$\begin{array}{c} v_j \bullet (\lambda_1 v_1 + \dots + \lambda_m v_m) = 0 \\ \shortparallel \\ \lambda_1 (v_j \bullet v_1) + \dots + \lambda_m (v_j \bullet v_m) = 0 \\ \shortparallel \\ \lambda_j \end{array}$$

$\Rightarrow \quad \lambda_1 = \dots = \lambda_m = 0 \Rightarrow v_1, \dots, v_m$ are lin. indep.

ii) Since $u \in V$ we can write

$$u = \lambda_1 v_1 + \dots + \lambda_m v_m .$$

$$\Rightarrow \quad u \bullet v_j = \lambda_j v_j \bullet v_j = \lambda_j$$

$$\Rightarrow \quad u = (u \bullet v_1) v_1 + \dots + (u \bullet v_m) v_m$$

$$\Rightarrow \quad [u]_B = \begin{pmatrix} u \bullet v_1 \\ \vdots \\ u \bullet v_m \end{pmatrix}.$$

iii) Let $[u]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $[w]_B = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$

$$\begin{aligned} u \bullet w &= (x_1 v_1 + \dots + x_m v_m) \bullet (y_1 v_1 + \dots + y_m v_m) \\ &= x_1 y_1 + \dots + x_m y_m \\ &= [u]_B \bullet [w]_B . \end{aligned}$$

Definition 12.4 For $U \subset \mathbb{R}^n$ subspace, we define the <u>orthogonal complement of U in $\mathbb{R}^n$</u> by

$$U^\perp = \{x \in \mathbb{R}^n \mid x \cdot u = 0 \text{ for all } u \in U\} \subset \mathbb{R}^n.$$

Lemma 12.5 Let $U \subset \mathbb{R}^n$ be a subspace

i) $U^\perp \subset \mathbb{R}^n$ is a subspace.

ii) We have $U \cap U^\perp = \{0\}$.

iii) If $(u_1, \ldots, u_r)$ is a basis of U, $x \in \mathbb{R}^n$, then

$$x \in U^\perp \iff x \cdot u_1 = \ldots = x \cdot u_r = 0.$$

iv) Let $(f_1, \ldots, f_r)$ be an ONB of U and $x \in \mathbb{R}^n$. Then

$$x = x_\parallel + x_\perp,$$

where

$$x_\parallel = \sum_{i=1}^{r} (x \cdot f_i) f_i \in U, \quad x_\perp = x - x_\parallel \in U^\perp.$$

Proof: i) Clearly $0 \in U^\perp$ since $0 \cdot v = 0$ for any $u \in \mathbb{R}^n$.
If $x, y \in U$, i.e. $x \cdot u = y \cdot u$ for all $u \in U$, then

$$(x+y) \cdot u = x \cdot u + y \cdot u = 0 + 0 = 0 \Rightarrow x+y \in U.$$

$(\lambda \in \mathbb{R})$ $$(\lambda x) \cdot u = \lambda (x \cdot u) = \lambda \cdot 0 = 0 \Rightarrow \lambda x \in U.$$

ii) If $x \in U \cap U^\perp$ then $x \cdot x = 0$

$$\underset{\displaystyle x_1^2 + \dots + x_n^2}{\overset{\parallel}{x \cdot x}} = 0 \quad \Rightarrow x_1 = \dots = x_n = 0.$$

iii) "$\Rightarrow$" is clear.

"$\Leftarrow$": For all $w \in U$: $w = \lambda_1 u_1 + \dots + \lambda_m u_m$

$$x \cdot w = x \cdot (\lambda_1 u_1 + \dots + \lambda_m u_m) = \lambda_1 \underbrace{(x \cdot u_1)}_{=0} + \dots + \lambda_r \underbrace{(x \cdot u_r)}_{=0}$$

$$= 0$$

$$\Rightarrow x \in U^\perp.$$

vi) Clearly $x_\parallel = \sum_{i=1}^{r} (x \cdot f_i) f_i \in U$. Want to show $x_\perp \in U^\perp$.

For all $1 \le j \le r$:

$$f_j \cdot x_\perp = f_j \cdot (x - x_\parallel) = f_j \cdot x - f_j \cdot \sum_{i=1}^{r} (x \cdot f_i) f_i$$

$$= f_j \cdot x - x \cdot f_j = 0.$$

$$\overset{\text{iii)}}{\Longrightarrow} \quad x_\perp \in U^\perp.$$

$\square$

The Gram-Schmidt algorithm / process

For a subspace V with basis $\mathcal{B}$ we want to create an ONB $\mathcal{F}$ of V.

Example 41 Consider $b_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $b_2 = \begin{pmatrix}1\\0\\2\end{pmatrix}$, $b_3 = \begin{pmatrix}-1\\2\\1\end{pmatrix}$.

$\mathcal{B} = (b_1, b_2, b_3)$ is a basis of $\mathbb{R}^3$.

Want to construct ONB $\mathcal{F} = (f_1, f_2, f_3)$. $\left(f_i \cdot f_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \right)$

Step 1: $f_1 = \hat{b}_1 = \frac{1}{\|b_1\|} b_1 = \frac{1}{\sqrt{3}} \begin{pmatrix}1\\1\\1\end{pmatrix}$.

Step 2.1: Make b_2 orthogonal to f_1

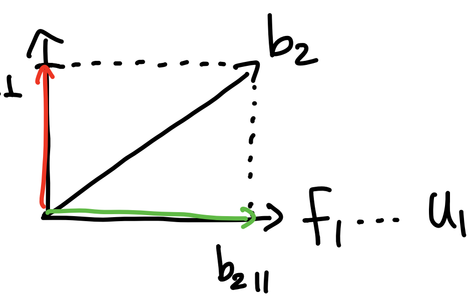
Set $U_1 = \operatorname{span}\{f_1\} = \operatorname{span}\{b_1\}$

Lemma 12.5 iv) $\Rightarrow$ $b_2 = \underbrace{b_{2\parallel}}_{\in\, U_1} + \underbrace{b_{2\perp}}_{\in\, U_1^{\perp}} = (b_2 \cdot f_1) f_1 + b_{2\perp}$

$\begin{pmatrix}1\\0\\2\end{pmatrix} \cdot \frac{1}{\sqrt{3}} \begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{\sqrt{3}} \cdot 3$

$w_2 = b_{2\perp} = b_2 - (b_2 \cdot f_1) f_1$ $\quad w_2 = b_{2\perp}$

$= \begin{pmatrix}1\\0\\2\end{pmatrix} - \frac{3}{\sqrt{3}} f_1 = \begin{pmatrix}1\\0\\2\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\-1\\1\end{pmatrix}$.



Step 2.2: Set $f_2 = \hat{w}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix}0\\-1\\1\end{pmatrix}$.

Set $U_2 = \text{span}\{f_1, f_2\} = \text{span}\{b_1, f_2\}$
$= \text{span}\{b_1, b_2\}$.

(f_1, f_2) are ONB of U_2.

<u>Step 3.1</u> $\quad b_3 = \underset{\substack{\cap \\ U_2}}{b_{3\parallel}} + \underset{\substack{\cap \\ U_2^\perp}}{b_{3\perp}}$

$\begin{pmatrix}-1\\2\\1\end{pmatrix} \cdot \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{\sqrt{3}} \cdot 2 \qquad \begin{pmatrix}-1\\2\\1\end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\end{pmatrix} = \frac{1}{\sqrt{2}} \cdot (-1)$

Set $w_3 = b_{3\perp} = b_3 - \underline{(b_3 \cdot f_1) f_1 - (b_3 \cdot f_2) f_2}$

$$= \begin{pmatrix}-1\\2\\1\end{pmatrix} - \frac{2}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0\\-1\\1\end{pmatrix} = \frac{5}{6}\begin{pmatrix}-2\\1\\1\end{pmatrix}$$

<u>Step 3.2</u> $\quad$ Set $f_3 = \hat{w}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}-2\\1\\1\end{pmatrix}$.

$U_3 = \text{span}\{f_1, f_2, f_3\} = \text{span}\{b_1, b_2, f_3\} = \text{span}\{b_1, b_2, b_3\}$.

$F = (f_1, f_2, f_3)$ are ONB of $U_3 = \mathbb{R}^3$.

General description of the Gram-Schmidt alg.

Given: Basis $\mathcal{B} = (b_1, \dots, b_m)$ of U.

Want construct ONB $\mathcal{F} = (f_1, \dots, f_m)$.

Step 1: Set $f_1 = \hat{b}_1 = \frac{1}{\|b_1\|} b_1$.

$U_1 = \operatorname{span}\{f_1\} = \operatorname{span}\{b_1\}$.

Step ℓ: Have ON $(f_1, \dots, f_{\ell-1})$, $U_{\ell-1} = \operatorname{span}\{f_1, \dots, f_{\ell-1}\}$
$2 \le \ell \le m$

$= \operatorname{span}\{b_1, \dots, b_{\ell-1}\}$.

(i.e. $b_\ell \notin U_{\ell-1}$)

We can write $b_\ell = \underset{\in U_{\ell-1}}{(b_\ell)_\parallel} + \underset{\in U_{\ell-1}^\perp}{(b_\ell)_\perp}$

Set $w_\ell = (b_\ell)_\perp = b_\ell - (b_\ell \cdot f_1) f_1 - \dots - (b_\ell \cdot f_{\ell-1}) f_{\ell-1}$

$f_\ell = \hat{w}_\ell = \frac{1}{\|w_\ell\|} w_\ell$.

$\Rightarrow$ $(f_1, \dots, f_\ell)$ ON, $U_\ell = \operatorname{span}\{f_1, \dots, f_\ell\}$

$= \operatorname{span}\{b_1, \dots, b_{\ell-1}, f_\ell\}$

$= \operatorname{span}\{b_1, \dots, b_\ell\}$.

After m steps $U_m = \text{span}\{f_1, \dots, f_m\}$
$= \text{span}\{b_1, \dots, b_m\} = U.$

$\Rightarrow \quad F = (f_1, \dots, f_m)$ ONB of U.

Theorem 12.6 Every subspace has an ONB.

Proof: Every subspace has a basis (Thm. 10.4).
Using GSA we get an ONB. $\square$

Corollary 12.7 Let $U \subset \mathbb{R}^n$ be a subspace.
For all $x \in \mathbb{R}^n$ there exist unique $x_\parallel \in U$ and $x_\perp \in U^\perp$ with
$$x = x_\parallel + x_\perp.$$

Proof: Existence: By Thm. 12.6 there exists an ONB $(f_1, \dots, f_m)$ of U. And by Lemma 12.5 we get $x_\parallel$ and $x_\perp$.

Uniqueness: Let $x = x_\parallel + x_\perp = y_\parallel + y_\perp$ for $x_\parallel, y_\parallel \in U$, $x_\perp, y_\perp \in U^\perp$.

$\Rightarrow U \ni x_\parallel - y_\parallel = x_\perp - y_\perp \in U^\perp$

$\Rightarrow \quad x_\parallel - y_\parallel = x_\perp - y_\perp \underset{\text{Lemma 12.5}}{=} 0 \Rightarrow x_\parallel = y_\parallel$ and $x_\perp = y_\perp$ $\square$