(18.0) Recall: last time we defined orthogonal and orthonormal set of vectors, using dot product on $\mathbb{R}^n$.

- Dot product $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$

$$\rightsquigarrow \quad \vec{u}^T \vec{v} = [u_1 \ u_2 \ \cdots \ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n \in \mathbb{R}.$$

- $\vec{u}$ and $\vec{v}$ are orthogonal to each other (in symbols $\vec{u} \perp \vec{v}$) if $\vec{u}^T \vec{v} = 0$.

- $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subset \mathbb{R}^n$. $\vec{0} \notin S$.

We say S is a set of orthogonal vectors if $\vec{v}_i \perp \vec{v}_j$ for every pair $i \neq j$. S is an orthonormal set of vectors if in addition $\|\vec{v}_i\| = 1$ for every $i = 1, 2, \ldots, p$.

(18.2) Two main points of Orthogonal Set of vectors.

I. If $S \subset \mathbb{R}^n$, $(\vec{0} \notin S)$ is an orthogonal set of vectors, then S is linearly independent.

II. Let $V = \mathrm{Span}(S)$. Given $\vec{w} \in V$, finding $x_1, x_2, \ldots, x_p \in \mathbb{R}$ so that $\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p$ is REALLY easy. Answer: $\boxed{x_j = \dfrac{\vec{v}_j^T \vec{w}}{\vec{v}_j^T \vec{v}_j}}$ $1 \le j \le p$.

(18.3) Gram-Schmidt procedure to construct orthogonal set of vectors.

We saw last time that if $B = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_p\}$ is a linearly independent set, then we can successively modify these vectors to produce another set

$$\tilde{B} = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$$

such that (1) $\mathrm{Span}(B) = \mathrm{Span}(\tilde{B})$ - so they are both bases of $V = \mathrm{Span}(B) = \mathrm{Span}(\tilde{B})$.

(2) $\tilde{B}$ is an orthogonal set of vectors. That is, $\vec{u}_i \perp \vec{u}_j$ for every pair $(i \neq j)$.

The algorithm to build $\tilde{B}$ is as follows:

- $\vec{u}_1 = \vec{w}_1$
- $\vec{u}_2 = \vec{w}_2 - \mathrm{Proj}_{\vec{u}_1}(\vec{w}_2) = \vec{w}_2 - \left(\dfrac{\vec{u}_1^T \vec{w}_2}{\vec{u}_1^T \vec{u}_1}\right)\vec{u}_1$
- $\vec{u}_3 = \vec{w}_3 - \mathrm{Proj}_{\vec{u}_1}(\vec{w}_3) - \mathrm{Proj}_{\vec{u}_2}(\vec{w}_3)$

$$= \vec{w}_3 - \left(\frac{\vec{u}_1^T \vec{w}_3}{\vec{u}_1^T \vec{u}_1}\right)\vec{u}_1 - \left(\frac{\vec{u}_2^T \vec{w}_3}{\vec{u}_2^T \vec{u}_2}\right)\vec{u}_2$$

and continue.

- $$\vec{u}_j = \vec{w}_j - \mathrm{Proj}_{\vec{u}_1}(\vec{w}_j) - \mathrm{Proj}_{\vec{u}_2}(\vec{w}_j) - \dots - \mathrm{Proj}_{\vec{u}_{j-1}}(\vec{w}_j)$$

$$= \vec{w}_j - \left(\frac{\vec{u}_1^T \vec{w}_j}{\vec{u}_1^T \vec{u}_1}\right)\vec{u}_1 - \left(\frac{\vec{u}_2^T \vec{w}_j}{\vec{u}_2^T \vec{u}_2}\right)\vec{u}_2$$

$$- \dots - \left(\frac{\vec{u}_{j-1}^T \vec{w}_j}{\vec{u}_{j-1}^T \vec{u}_{j-1}}\right)\vec{u}_{j-1}.$$

(18.4) Why does this method work?

The proof that this method works is inductive.* Meaning we assume that the method after j steps produced $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_j\}$ so that

- $\mathrm{Span}\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_j\} = \mathrm{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_j\}$
- $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_j\}$ is an orthogonal set.

and we check that $\vec{u}_{j+1}$ also has these properties.

e.g. $$\vec{u}_1^T \vec{u}_{j+1} = \vec{u}_1^T \vec{w}_{j+1} - \left(\frac{\vec{u}_1^T \vec{w}_{j+1}}{\vec{u}_1^T \vec{u}_1}\right)\vec{u}_1^T\vec{u}_1 - \left(\frac{\vec{u}_2^T \vec{w}_{j+1}}{\vec{u}_2^T \vec{u}_2}\right)\left(\cancelto{0}{\vec{u}_1^T \vec{u}_2}\right)$$

$-$ the rest are all 0 since $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_j\}$ is orthogonal set

$$= \vec{u}_1^T \vec{w}_{j+1} - \vec{u}_1^T \vec{w}_{j+1} = 0. \checkmark$$

---

* Optional

(18.5) Example. Let $\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$ $\vec{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 2 \end{bmatrix}$ and $\vec{w}_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$

be three linearly independent vectors in $\mathbb{R}^4$.

Apply Gram-Schmidt orthogonalization procedure to get an ortho~~normal~~gonal set of vectors $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$. Verify that $\begin{bmatrix} 5 \\ 0 \\ -1 \\ 4 \end{bmatrix}$ is in the span of $\vec{w}_1, \vec{w}_2, \vec{w}_3$.

Gram Schmidt: $\vec{u}_1 = \vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$

$$\vec{u}_1^T \vec{w}_2 = 1(2) + 0(1) + 1(0) + 2(2) = 6.$$

$$\vec{u}_1^T \vec{u}_1 = 1^2 + 0^2 + 1^2 + 2^2 = 6.$$

So, $\cancel{\vec{w}_2}$ $\vec{u}_2$

$$\vec{u}_2 = \vec{w}_2 - \left(\frac{\vec{u}_1^T \vec{w}_2}{\vec{u}_1^T \vec{u}_1} \right) \vec{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 2 \end{bmatrix} - \frac{6}{6} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\vec{u}_1^T \vec{w}_3 = 1(1) + 0(-1) + 1(0) + 2(1) = 3$$

$$\vec{u}_2^T \vec{w}_3 = 1(1) + 1(-1) + (-1)(0) + 0(1) = 0$$

$$\vec{u}_2^T \vec{u}_2 = 1^2 + 1^2 + (-1)^2 + 0^2 = 3$$

$$\vec{u}_3 = \vec{w}_3 - \left(\frac{\vec{u}_1^T \vec{w}_3}{\vec{u}_1^T \vec{u}_1} \right) \vec{u}_1 - \left(\frac{\vec{u}_2^T \vec{w}_3}{\vec{u}_2^T \vec{u}_2} \right) \vec{u}_2$$

$$= \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{3}{6}\right) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} - \left(\frac{0}{3}\right) \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - \frac{1}{2} \\ -1 - 0 \\ 0 - \frac{1}{2} \\ 1 - 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ 0 \end{bmatrix}.$$

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} \text{ and } \vec{u}_3 = \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \\ 0 \end{bmatrix}$$

(Check: $\vec{u}_1^T \vec{u}_2 = \vec{u}_1^T \vec{u}_3 = \vec{u}_2^T \vec{u}_3 = 0$.)

> Note: $\vec{w}_1 = \vec{u}_1$; $\vec{w}_2 = \vec{u}_2 + \vec{u}_1$; $\vec{w}_3 = \vec{u}_3 + \frac{1}{2}\vec{u}_1$
>
> $\vec{u}_1 = \vec{w}_1$; $\vec{u}_2 = \vec{w}_2 - \vec{w}_1$; $\vec{u}_3 = \vec{w}_3 - \frac{1}{2}\vec{w}_1$
>
> So $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.

just to make a point, not part of the problem.

Now $\vec{b} = \begin{bmatrix} 5 \\ 0 \\ -1 \\ 4 \end{bmatrix} \in \mathbb{R}^4$. If $\vec{b} \in \text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, then there are real numbers x_1, x_2, x_3 so that $\vec{b} = x_1\vec{u}_1 + x_2\vec{u}_2 + x_3\vec{u}_3$.

$$\vec{u}_1^T \vec{b} = x_1 \vec{u}_1^T \vec{u}_1 + x_2 \underbrace{\cancel{\vec{u}_1^T \vec{u}_2}}_{0} + x_3 \underbrace{\cancel{\vec{u}_1^T \vec{u}_3}}_{0}$$

$$\Rightarrow \quad x_1 = \frac{\vec{u}_1^T \vec{b}}{\vec{u}_1^T \vec{u}_1} = \frac{1(5) + 0(0) + 1(-1) + 2(4)}{6} = \frac{12}{6} = 2.$$

Similarly,
$$x_2 = \frac{\vec{u}_2^T \vec{b}}{\vec{u}_2^T \vec{u}_2} = \frac{1(5) + 1(0) + (-1)(-1) + 0(4)}{3} = \frac{6}{3} = 2.$$

$$x_3 = \frac{\vec{u}_3^T \vec{b}}{\vec{u}_3^T \vec{u}_3} = \frac{\frac{1}{2}(5) + (-1)(0) + (-\frac{1}{2})(-1) + 0(4)}{\frac{1}{4} + 1 + \frac{1}{4} + 0} = \frac{\frac{5}{2} + \frac{1}{2}}{\frac{3}{2}}$$

$$= \frac{6}{2} \cdot \frac{2}{3} = 2.$$

So
$$x_1\vec{u}_1 + x_2\vec{u}_2 + x_3\vec{u}_3 = 2\vec{u}_1 + 2\vec{u}_2 + 2\vec{u}_3 = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ -1 \\ 4 \end{bmatrix}$$

We get $\vec{b} = 2\vec{u}_1 + 2\vec{u}_2 + 2\vec{u}_3$, hence $\vec{b}$ is in the

$\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$.

(18.6) Linear Transformations

Our next topic is linear functions from a vector space to another, called linear transformations. Recall (from Calculus) that a function $\boxed{F : \mathbb{R}^n \to \mathbb{R}^m}$ takes inputs from $\mathbb{R}^n$ and gives outputs in $\mathbb{R}^m$.

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \rightsquigarrow \boxed{F} \rightsquigarrow \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix} \text{ in } \mathbb{R}^m$$

in $\mathbb{R}^n$

For example, $n = m = 1$ (Calculus I) - function of one variable $f(x)$

$x \in \mathbb{R} \rightsquigarrow f(x) \in \mathbb{R}$.

e.g. $\sin(x)$, x^2+4, $\log(x)$, e^x, ... are functions from $\mathbb{R}$ to $\mathbb{R}$. (except for $\log : \mathbb{R}_{>0} \to \mathbb{R}$).

$n = 2$, $m = 1$ (Calculus III) - functions of 2 variables, $f(x,y)$

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \rightsquigarrow f(x,y) \in \mathbb{R}$$

e.g. $f(x,y) = 2x + 4y$, $x^3 - y^2$, $\sin(x+y)$, ... functions from $\mathbb{R}^2$ to $\mathbb{R}$.

- $n=3$, $m=3$ (called vector fields in Calculus III)
  - $\underline{\text{Three}}^{\,m}$ functions $f(x,y,z)$, $g(x,y,z)$, $h(x,y,z)$ each of $\underline{3}_{\,n}$ variables

e.g. $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$ $\leadsto$ $\begin{bmatrix} 2xy+4 \\ \cos(x^2z) \\ 9y^2-4z \end{bmatrix}$ in $\mathbb{R}^3$.

<u>Linear transformations</u> are <u>functions from $\mathbb{R}^n$ to $\mathbb{R}^m$</u> (or from a vector subspace of $\mathbb{R}^n$ to $\mathbb{R}^m$) which <u>respect the vector space structure</u> - i.e., addition and scalar multiplication.

<u>Definition.</u> A <u>linear transformation</u> $\boxed{T : \mathbb{R}^n \to \mathbb{R}^m}$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

- $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for $\vec{u}, \vec{v} \in \mathbb{R}^n$.
- $T(c\vec{u}) = cT(\vec{u})$ for $\vec{u} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

Almost all the functions in the examples above are <u>NOT</u> linear transformations. Only $f(x,y) = 2x+4y : \mathbb{R}^2 \to \mathbb{R}$ is.

$$T : \mathbb{R}^2 \to \mathbb{R} \qquad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = 2x+4y.$$

Let us check that it is a linear transformation.

- $T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right)$

$$= 2(x_1 + x_2) + 4(y_1 + y_2) = (2x_1 + 4y_1) + (2x_2 + 4y_2)$$

$$= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \quad \checkmark$$

- $T\left(c\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) = 2(cx) + 4(cy) = c(2x+4y)$

$$= c \cdot T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right).$$

The point is: in the picture on page 6 $\quad \mathbb{R}^n \rightsquigarrow \boxed{F} \rightsquigarrow \mathbb{R}^m$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \rightsquigarrow \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}$$

Every $f_1, f_2, \dots, f_m$ has to be a linear expression in $x_1, x_2, \dots, x_n$ without a constant term for F to be a linear transformation.

e.g. Let $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be given by

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - y \\ 2y + 1 \end{bmatrix}.$$

This is NOT a linear transformation.

(because of 1 in $2y+1$).

Check: $F\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = F\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 - y_1 - y_2 \\ 2(y_1 + y_2) + 1 \end{bmatrix}$

NOT equal

$$F\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + F\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - y_1 \\ 2y_1 + 1 \end{bmatrix} + \begin{bmatrix} x_2 - y_2 \\ 2y_2 + 1 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - y_1 - y_2 \\ 2(y_1 + y_2) + 2 \end{bmatrix}.$$