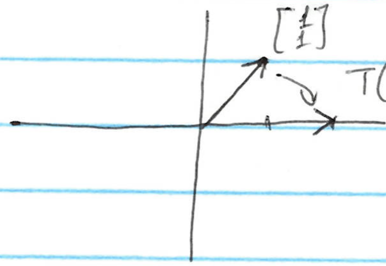
5.1) ~~A~~ Eigenvalues.

A linear transformation $T: V \to V$ tends to move vectors around.

E.g. $T: \mathbb{R}^2 \to \mathbb{R}^2$

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



Def: Let $T: V \to V$ be a linear transformation. If there is a non-zero $\vec{v} \in V$ such that

$$T(\vec{v}) = \lambda \vec{v}.$$

for some scalar $\lambda \in \mathbb{R}$, we call $\vec{v}$ an <u>eigenvector</u> of T. The scalar λ is the corresponding <u>eigenvalue</u>.

(We similarly can talk about eigenvectors / eigenvalues of square matrices.)

E.g. $A = \begin{bmatrix} 1 & 2 \\ -3 & 6 \end{bmatrix}$

$\lambda = 3$ is an eigenvalue of A and its corresponding eigenvector is $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. since

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Next lecture, we will learn how to find eigenvalues and eigenvectors of a given matrix.

Thm: For any $\lambda \in \mathbb{R}$, the set of vectors $\vec{v}$ for which

$$T(\vec{v}) = \lambda \vec{v}$$

is a subspace of V, called the <u>eigenspace</u> of λ.

(We only call λ an eigenvalue if its eigenspace $\neq \{\vec{0}\}$)

<u>Pf.</u> $T(\vec{v}) = \lambda \vec{v}$

$\Leftrightarrow T(\vec{v}) - \lambda \vec{v} = \vec{0}.$

$\Leftrightarrow (T - \lambda I)\vec{v} = \vec{0}.$

$\Rightarrow \vec{v} \in Nul(T - \lambda I).$

E.g. The rotation matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has no eigenvalues.

Recall λ is an eigenvalue if $\exists \vec{v} \neq \vec{0}$ such that

$$A\vec{v} = \lambda \vec{v}$$

$$(A - \lambda I)\vec{v} = \vec{0}.$$

$$\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \vec{0}.$$

$\Rightarrow -\lambda v_1 - v_2 = 0$

$v_1 - \lambda v_2 = 0 \Rightarrow v_1 = \lambda v_2$

$$-\lambda^2 v_2 - v_2 = 0.$$

$$(\lambda^2 + 1) v_2 = 0.$$

$$v_2 = 0$$

$$\Rightarrow v_1 = 0.$$

$\therefore \vec{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$

Given an $n \times n$ matrix A, an eigenvector is a nonzero vector $\vec{v} \in \mathbb{R}^n$ with the property

$$A\vec{v} = \lambda \vec{v}$$

for some scalar λ, called the eigenvalue.

For a given eigenvalue λ, the set of eigenvectors corresponding to λ is $Nul(A - \lambda I)$. So it is a subspace of $\mathbb{R}^n$, called the eigenspace for λ.

E.g. Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$ and $\lambda = 2$ is its eigenvalue. Find a basis for the eigenspace for λ.

That is, the basis of $Nul(A - 2I)$.

$$A - 2I = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \longrightarrow \xrightarrow{RREF} \begin{bmatrix} 1 & -1/2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{ccc} & \uparrow & \uparrow \\ & \text{free} & \end{array}$$

$$Null(A - 2I) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} \right\} = span\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

$$\begin{array}{cc} \uparrow & \uparrow \end{array}$$

basis

E.g. Find eigenvalue and eigenvector of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

i.e. Find λ such that $Nul(A - \lambda I) \neq \{\vec{0}\}$.

i.e. $A - \lambda I$ not invertible

i.e. $\det(A - \lambda I) = 0$.

$$\Rightarrow \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 = 0$$

$$\lambda = 1.$$

For $\lambda = 1$,

$$A - 1 \cdot I = \begin{array}{cc} \text{free} & \\ \downarrow & \\ \end{array} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow Nul(A - 1I) = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \right\} = span\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}.$$

~~Thm: Eigenvectors corresponding to distinct eigenvalues are linearly independent.~~

Def: The scalar equation $\det(A - \lambda I) = 0$ is called the <u>characteristic equation</u> of A.

Fact: λ is a eigenvalue of A
$\Leftrightarrow$ λ is a solution of $\det(A - \lambda I) = 0$.

E.g. Find the characteristic equation of

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix}$$

$$= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

Thm: Eigenvectors with distinct eigenvalues are linearly independent.

Corollary: If A is an $n \times n$ matrix with n distinct eigenvalues, then there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

$$\beta = \{ \overset{\lambda_1}{\vec{v}_1}, \overset{\lambda_2}{\vec{v}_2}, \ldots, \overset{\lambda_n}{\vec{v}_n} \} \quad \text{eigenvectors of } A.$$

Any vector $\vec{w} \in \mathbb{R}^n$ can be expanded in β.

$$\vec{w} = x_1 \vec{v}_1 + x_2 \vec{v}_2 + \cdots + x_n \vec{v}_n \rightsquigarrow [\vec{w}]_\beta = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

$$\Rightarrow A\vec{w} = A(x_1 \vec{v}_1 + x_2 \vec{v}_2 + \cdots + x_n \vec{v}_n)$$
$$= x_1 A\vec{v}_1 + x_2 A\vec{v}_2 + \cdots + x_n A\vec{v}_n$$
$$= x_1 \lambda_1 \vec{v}_1 + x_2 \lambda_2 \vec{v}_2 + \cdots + x_n \lambda_n \vec{v}_n.$$

$$\therefore \quad [\vec{w}]_\beta = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \Rightarrow [Aw]_\beta = \begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_n \end{bmatrix}}_{D} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Let $P = [\vec{v}_1 \; \vec{v}_2 \; \ldots \; \vec{v}_n]$ (invertible since $\{\vec{v}_1, \ldots, \vec{v}_n\}$ a basis).

$$AP = A[\vec{v}_1 \; \vec{v}_2 \; \ldots \; \vec{v}_n]$$
$$= [A\vec{v}_1 \; A\vec{v}_2 \; \ldots \; A\vec{v}_n]$$
$$= [\lambda_1 \vec{v}_1 \; \lambda_2 \vec{v}_2 \; \ldots \; \lambda_n \vec{v}_n].$$
$$\overset{?}{=} [\vec{v}_1 \; \vec{v}_2 \; \ldots \; \vec{v}_n] \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \; \lambda_n \end{bmatrix}$$
$$= \quad P \quad D$$

$$\Rightarrow \quad AP = PD.$$

$$\therefore \quad \boxed{A = PDP^{-1}}$$

That is, we can write A in term of nicer matrices P and D.

$$A^2 = PDP^{-1}PDP^{-1}$$
$$= PD^2P^{-1}$$
$$\Rightarrow \quad A^k = PD^kP^{-1}.$$

Def: $n \times n$ matrices A and B are called <u>similar</u> if there is an invertible $n \times n$ matrix P with the property

$$A = PBP^{-1}.$$

Remark: 1) If A has all distinct eigenvalues, it is similar to a diagonal matrix.

Thm: If A and B are similar, then they have the same eigenvalues, and the eigenspaces for A have the same dimensions as the eigenspaces for B.

if A & B similar $\Rightarrow \exists P$ such that $A = PBP^{-1}$
or $P^{-1}AP = B$

if $\vec{v}$ is an eigenvector with eigenvalue λ,

$$A\vec{v} = \lambda\vec{v}.$$
$$\Rightarrow \quad PBP^{-1}\vec{v} = \lambda\vec{v}.$$
$$B(P^{-1}\vec{v}) = \lambda(P^{-1}\vec{v}) \Rightarrow$$

$P^{-1}\vec{v}$ is an eigenvector of B, with eigenvalue λ.

An $n \times n$ matrix A is <u>diagonalizable</u> if

$$A = P \underset{\substack{\uparrow \\ \text{diagonal}}}{D} P^{-1}$$

This equivalent to $\mathbb{R}^n$ having a basis of eigenvectors for A.
This happens when A has n distinct eigenvalues.

Characteristic polynomial:

$$p_A(\lambda) = \det(A - \lambda I) \quad \text{of deg } n.$$

$p_A(\lambda)$ must have all real roots for A to be diagonalizable.

- Algebraic multiplicity of λ = degree of this root in $\cancel{p_A}$.
- Geometric multiplicity of λ = $\dim \mathrm{Nul}(A - \lambda I)$.

E.g. $A = \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 5 \\ 0 & 0 & 5 \end{bmatrix}$.

Then $p_A(\lambda) = p_B(\lambda) = (5-\lambda)^{2}(4-\lambda)^{1}$

alg. mult. of 5 is 2.
alg. mult. of 4 is 1.

$\dim \mathrm{Nul}(A - 5I) = 2$.
$\dim \mathrm{Nul}(B - 5I) = 1$.

Thm: If A has a repeated eigenvalue λ, then
~~dim Nul~~ $\dim \mathrm{Nul}(A - \lambda I) \leq$ multiplicity of λ.
The matrix A is diagonalizable if and only if p_A has only real roots and
$\dim \mathrm{Nul}(A - \lambda I) =$ mult. of λ for <u>each</u> λ.

Inner products.

Def: The inner product ~~and~~ (or dot product) on $\mathbb{R}^n$ is defined by

$$\vec{u}, \vec{v} \in \mathbb{R}^n, \quad \vec{u} \cdot \vec{v} = \langle \vec{u}, \vec{v} \rangle$$

$$= \vec{u}^T \vec{v}.$$

scalar

$$= u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

E.g. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = (1)\cdot(-1) + 0(1) + 1(0) = -1.$

Remark: $\underset{\text{scalar}}{\vec{u}^T \vec{v}} \neq \underset{\text{matrix}}{\vec{u}\,\vec{v}^T}.$

Properties:

1) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}.$
2) $(\vec{u}_1 + \vec{u}_2) \cdot \vec{v} = \vec{u}_1 \cdot \vec{v} + \vec{u}_2 \cdot \vec{v}.$
3) $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}).$
4) $\vec{u} \cdot \vec{u} = u_1^2 + u_2^2 + \dots + u_n^2 > 0$ unless $\vec{u} = \vec{0}.$

The length of a vector is defined to be

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}.$$

the distance between two vectors $\vec{u}, \vec{v}$ is

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|.$$

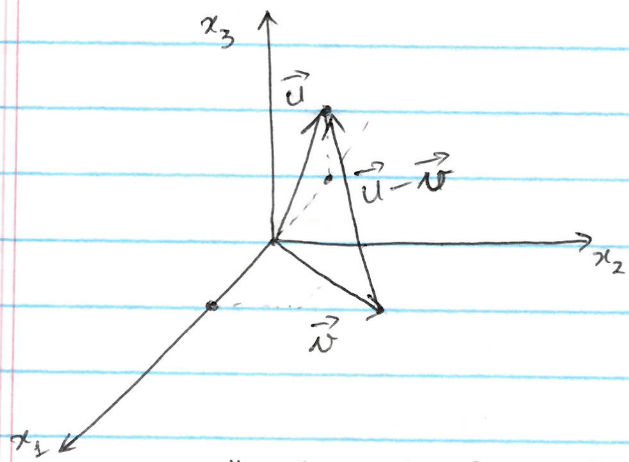
E.g. $\vec{u} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ $\quad \vec{v} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$.

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \left\| \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \right\|$$

$$= \sqrt{0^2 + (-2)^2 + 1^2}$$

$$= \sqrt{5}.$$

How does $\|\vec{u} - \vec{v}\|$ relate to $\|\vec{u}\|$, $\|\vec{v}\|$?



$$\|\vec{u} - \vec{v}\|^2 = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$$
$$= \vec{u} \cdot (\vec{u} - \vec{v}) + (-\vec{v}) \cdot (\vec{u} - \vec{v})$$
$$= \vec{u} \cdot \vec{u} - \vec{u} \cdot \vec{v}$$
$$- \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v}$$
$$= \|u\|^2 - 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2.$$

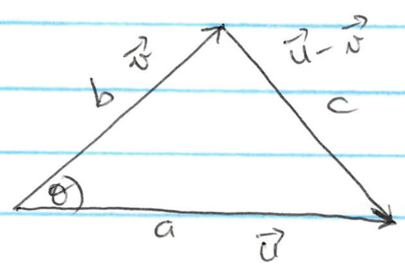
$$\therefore \quad \|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u} - \vec{v}\|^2 + 2\vec{u} \cdot \vec{v}.$$

If $\vec{u} \cdot \vec{v} = 0$ or $\vec{u} \perp \vec{v}$, then

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u} - \vec{v}\|^2$$

Pythagorean theorem.

Law of Cosines.



$$a^2 + b^2 = c^2 + 2ab\cos\theta.$$
$$\|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u} - \vec{v}\|^2 + 2\|\vec{u}\|\|\vec{v}\|\cos\theta.$$

$$\Rightarrow \quad \vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta.$$

$$\theta = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|} \right).$$

- Cauchy-Schwartz inequality

$$|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\,\|\vec{v}\|.$$

Pf. Consider a function

$$f(t) = \|\vec{u} - t\vec{v}\|^2 \geq 0.$$

$$= (\vec{u} - t\vec{v})\cdot(\vec{u} - t\vec{v}).$$

$$= \dots = \|u\|^2 - 2t\vec{u}\cdot\vec{v} + \|\vec{v}\|^2 t^2.$$

$$= a_0 + a_1 t + a_2 t^2.$$

$\Rightarrow$ find minimum value of f.

$$f' = a_1 + 2a_2 t = 0.$$

$$t = -\frac{a_1}{2a_2}$$

min f occurs at $t = -\dfrac{a_1}{2a_2}$.

$$\Rightarrow \quad a_0 + a_1\left(-\frac{a_1}{2a_2}\right) + a_2\left(-\frac{a_1}{2a_2}\right)^2 \geq 0.$$

$$a_0 - \frac{a_1^2}{4a_2} \geq 0.$$

$$a_1^2 \leq 4a_0 a_2.$$

$$4(\vec{u}\cdot\vec{v})^2 \leq 4.\|u\|_2^2\,\|v\|_2^2.$$

The inner product on $\mathbb{R}^n$ encodes both lengths and angles.

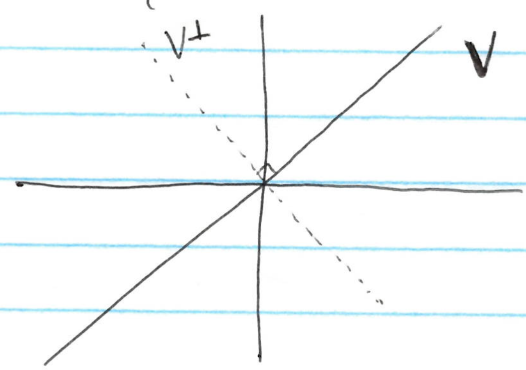
$$\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} \qquad\qquad \vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos\theta$$

length $\qquad\qquad\qquad\qquad$ angle between $\vec{u}$ & $\vec{v}$

We say that two vectors $\vec{u}$ and $\vec{v}$ are <u>orthogonal</u> if and only if $\vec{u}\cdot\vec{v} = 0$.

<u>Def:</u> If $V \subseteq \mathbb{R}^n$ (subspace), the <u>orthogonal complement</u> of V, denoted $V^\perp$, is defined to be

$$V^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x}\cdot\vec{v} = 0 \text{ for all } \vec{v} \in V\}.$$



<u>Thm:</u> For any $m\times n$ matrix A,

1) $(\text{Row } A)^\perp = \text{Nul}(A)$.
2) $(\text{Col } A)^\perp = \text{Nul}(A^T)$.

For 1), it means $\vec{v} \in \text{Row}(A)$, $\vec{u} \in \text{Nul}(A)$ $\rightarrow$ $\vec{v}\cdot\vec{u} = 0$.

For 2) it means $\vec{v} \in \text{Col}(A)$, $\vec{u} \in \text{Nul}(A^T)$ $\rightarrow$ $\vec{v}\cdot\vec{u} = 0$.

72

E.g. Consider a 2×3 matrix A.

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{matrix} \leftarrow \vec{a}^T \\ \leftarrow \vec{b}^T \end{matrix}$$

denote $\vec{a}^T$ 1st row of A

$\vec{b}^T$ 2nd row of A

$$\text{Row}(A) = \text{Col}(A^T) = \text{span}\left\{ \underset{\vec{a}}{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}}, \underset{\vec{b}}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}} \right\}.$$

$$\text{Nul}(A) = \{ \vec{x} : A\vec{x} = \vec{0} \}.$$

$$= \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

$$\begin{bmatrix} a_1 x_1 + a_2 x_2 + a_3 x_3 \\ b_1 x_1 + b_2 x_2 + b_3 x_3 \end{bmatrix} = \begin{bmatrix} \vec{a}\cdot\vec{x} \\ \vec{b}\cdot\vec{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence, $\vec{x} \in \text{Nul}(A) \iff \vec{x} \cdot (\text{each row of } A) = 0$

So for any $\vec{v} \in \text{Row}(A)$

$$\vec{v} = \alpha \vec{a} + \beta \vec{b}.$$

$$\begin{aligned} \vec{v}\cdot\vec{x} &= (\alpha\vec{a} + \beta\vec{b})\cdot\vec{x} \\ &= \alpha\vec{a}\cdot\vec{x} + \beta\vec{b}\cdot\vec{x}. \\ &= 0 + 0. \\ &= 0. \end{aligned}$$

Orthogonal sets of vectors

Recall that two vectors $\vec{u}$ and $\vec{v}$ are orthogonal if $\vec{u}\cdot\vec{v} = 0$.

E.g. the standard basis vectors are all orthogonal.

$$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

E.g. $\vec{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$ $\vec{u}_2 = \begin{bmatrix}1\\-1\\0\end{bmatrix}$ $\vec{u}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$.

are all orthogonal.

Thm: If $\vec{u}_1, \dots, \vec{u}_n$ are orthogonal, they are linearly independent. all $\neq \vec{0}$

Pf. Consider $x_1\vec{u}_1 + \dots + x_n\vec{u}_n = \vec{0}$. $(*)$

~~then~~ $\cancel{(x_1\vec{u}_1 + \dots +}$

We'll show that $x_1 = x_2 = \dots = x_n = 0$.

~~Applying~~ Taking dot product both sides of $(*)$ with $\vec{u}_1$,

$$(x_1\vec{u}_1 + \dots + x_n\vec{u}_n)\cdot\vec{u}_1 = \vec{0}\cdot\vec{u}_1.$$

$$x_1\vec{u}_1\cdot\vec{u}_1 + \dots + x_n\cancelto{0}{\vec{u}_n\cdot\vec{u}_1} = 0.$$

$$x_1\vec{u}_1\cdot\vec{u}_1 = 0.$$

$$x_1 \|\vec{u}_1\|_2^2 = 0.$$

Since $\vec{u}_1 \neq 0$, $\|\vec{u}_1\| > 0$

$\Rightarrow$ $x_1 = 0$

Similarly, $x_2 = \dots = x_n = 0$.

Defi Let $V \subseteq \mathbb{R}^n$ be a subspace. A basis for V is called an <u>orthogonal basis</u> if all the basis vectors are orthogonal. (It is call an <u>orthonormal basis</u> if it is an orthogonal basis and each basis vector has length 1.)

E.g. The standard basis is an orthonormal basis.

E.g. Given a vector $\vec{u} \neq 0$ and $\|\vec{u}\| \neq 1$. (i.e. $\vec{u}$ does not have length 1.) How to get a vector of length 1 with the same diraction as $\vec{u}$?

$\Rightarrow$ Normalize it.

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

Then $$\hat{u} \cdot \hat{u} = \frac{\vec{u} \cdot \vec{u}}{\|\vec{u}\| \, \|\vec{u}\|} = \frac{\|\vec{u}\|^2}{\|\vec{u}\|^2} = 1.$$

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

E.g. $\vec{u}_1 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$ $\vec{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ $\vec{u}_3 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ is an orthogonal basis.

$\vec{u}_1 \cdot \vec{u}_1 = 2^2 + 0 + 2^2 = 8$ $\qquad$ $\vec{u}_1 \cdot \vec{u}_2 = (-2) + 2 = 0.$

$\vec{u}_2 \cdot \vec{u}_2 = (-1)^2 + 0 + 1^2 = 2$ $\qquad$ $\vec{u}_1 \cdot \vec{u}_3 = 0.$

$\vec{u}_3 \cdot \vec{u}_3 = 0 + 2^2 + 0 = 4.$ $\qquad$ $\vec{u}_2 \cdot \vec{u}_3 = 0.$

$\left\{ \hat{u}_1 = \frac{\vec{u}_1}{\sqrt{8}}, \ \hat{u}_2 = \frac{\vec{u}_2}{\sqrt{2}}, \ \hat{u}_3 = \frac{\vec{u}_3}{\sqrt{4}} \right\}$ is an orthonormal basis.

<u>Theorem</u>: If $\beta = \{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthogonal basis for a subspace V, then the coefficients of any vector $\vec{v} \in V$ in the basis β are

$$[\vec{v}]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{where} \quad c_j = \frac{\vec{v} \cdot \vec{u}_j}{\|\vec{u}_j\|^2}.$$

Pf: $\vec{v} = c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_n\vec{u}_n$.

Then $\vec{v} \cdot \vec{u}_1 = c_1\vec{u}_1 \cdot \vec{u}_1 + \cancelto{0}{c_2\vec{u}_2 \cdot \vec{u}_1} + \dots + \cancelto{0}{c_n\vec{u}_n \cdot \vec{u}_1}$

$= c_1 \|\vec{u}_1\|^2$.

$$\therefore \; c_1 = \frac{\vec{v} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}.$$

<u>Orthogonality and matrices</u>.

Given an $m \times n$ matrix A, the matrix A^TA encodes the dot products of the columns of $A = [\vec{a}_1\ \vec{a}_2\ \vec{a}_3 \dots \vec{a}_n]$

$$A^TA = \begin{bmatrix} -\ \vec{a}_1^T\ - \\ \vdots \\ -\ \vec{a}_n^T\ - \end{bmatrix} [\vec{a}_1 \dots \vec{a}_n] = \begin{bmatrix} \vec{a}_1 \cdot \vec{a}_1 & \vec{a}_1 \cdot \vec{a}_2 & \dots & \vec{a}_1 \cdot \vec{a}_n \\ \vec{a}_2 \cdot a_1 & \vec{a}_2 \cdot \vec{a}_2 & \dots & \vec{a}_2 \cdot \vec{a}_n \\ \vdots & \vdots & & \vdots \\ \vec{a}_n \cdot \vec{a}_1 & \vec{a}_n \cdot \vec{a}_2 & \dots & \vec{a}_n \cdot \vec{a}_n \end{bmatrix}$$

If $\vec{a}_1, \dots, \vec{a}_n$ are <u>orthonormal</u>,

$$A^TA = I_{n\times n}.$$

E.g. ~~A~~ $A = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{3} \\ 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{3} \end{bmatrix}$.

$$A^TA = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{3} \\ 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

A collection $\vec{u}_1, \dots, \vec{u}_p$ of vectors in $\mathbb{R}^n$ is <u>orthogonal</u> if $\vec{u}_i \cdot \vec{u}_j = 0$ for $i \neq j$.

They are <u>orthonormal</u> if, in addition, $\|\vec{u}_i\| = 1$ for all $i = 1, \dots, p$.

Let $U = [\vec{u}_1 \; \vec{u}_2 \; \dots \; \vec{u}_p]$. ($n \times p$ matrix)

Then the (i,j) entry of $U^T U$ is $\vec{u}_i \cdot \vec{u}_j$.

$\{\vec{u}_1, \dots, \vec{u}_p\}$ are orthonormal if and only if

$$\underset{p\times n}{U^T} \, \underset{n\times p}{U} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \; 1 \end{bmatrix}_{p\times p} = I_p .$$

Remark: if $p < n$, this does not mean that $UU^T = I_n$!

<u>Def:</u> If the columns of an <u>$n \times n$</u> matrix U form an orthonormal basis for $\mathbb{R}^n$, we call U an <u>orthogonal matrix</u>

$$\underset{n\times n}{U^T} \, \underset{n\times n}{U} = I_n .$$

$\Rightarrow \quad U^T = U^{-1}$

Recall $\quad UU^{-1} = I$, then $\quad UU^T = I$.

i.e. the rows of U also form an O.N.B. of $\mathbb{R}^n$.

Think of such U as a linear transformation.

$$T(\vec{x}) = U\vec{x}.$$

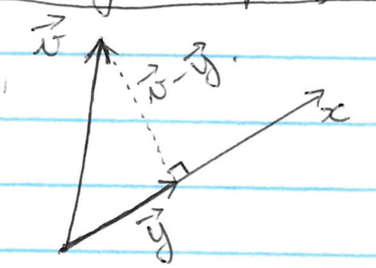
Then U preserves inner products and norm.

- $$\begin{aligned}\langle T(\vec{x}), T(\vec{y})\rangle &= \langle U\vec{x}, U\vec{y}\rangle \\ &= (U\vec{x})^T(U\vec{y}) \\ &= \vec{x}^T(U^T U)\vec{y} \\ &= \vec{x}^T\vec{y} \\ &= \langle \vec{x}, \vec{y}\rangle.\end{aligned}$$

- $$\begin{aligned}\|T(\vec{x})\|^2 &= \langle T(\vec{x}), T(\vec{x})\rangle \\ &= \langle \vec{x}, \vec{x}\rangle \\ &= \|\vec{x}\|^2.\end{aligned}$$

* Orthogonal projections: Given a vectors $\vec{x}$ and $\vec{v}$, Find a vector $\vec{y}$ such that

1) $\vec{y} = \lambda\vec{x}$.
2) $\vec{x} \perp \vec{v} - \vec{y}$.



$$\begin{aligned}&\vec{x} \perp \vec{v} - \vec{y}. \\ &\vec{x} \cdot (\vec{v} - \vec{y}) = 0. \\ &\vec{x} \cdot (\vec{v} - \lambda\vec{x}) = 0. \\ &\vec{x} \cdot \vec{v} - \lambda\,\vec{x}\cdot\vec{x} = 0. \\ &\vec{x} \cdot \vec{v} - \lambda\|\vec{x}\|^2 = 0. \\ &\lambda = \frac{\vec{v}\cdot\vec{x}}{\|\vec{x}\|^2}.\end{aligned}$$

$$\therefore \boxed{\vec{y} = \mathrm{Proj}_{\vec{x}}(\vec{v}) = \frac{\vec{v}\cdot\vec{x}}{\|\vec{x}\|_2^2}\vec{x}}$$

orthogonal projection of $\vec{v}$ onto $\vec{x}$.

E.g. $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ $\quad \vec{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\quad \vec{w} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

$$\mathrm{Proj}_{\vec{u}}(\vec{v}) = \frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}\vec{u} = \frac{1}{\left(\sqrt{1^2+2^2+(-2)^2}\right)^2}\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}.$$

$$\mathrm{Proj}_{\vec{v}}(\vec{w}) = \vec{0}.$$

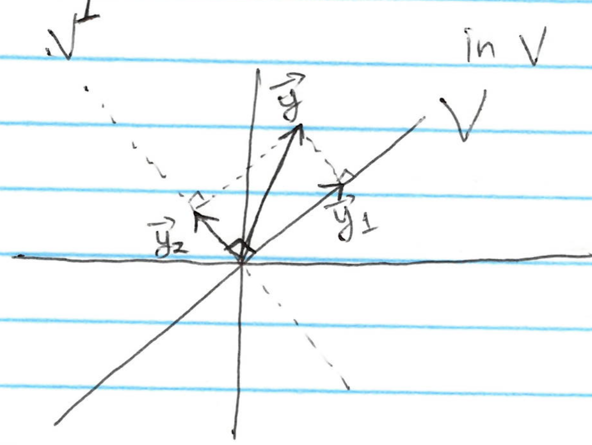
If $V \subseteq \mathbb{R}^n$ is a subspace, it has an orthogonal complement $V^\perp \subseteq \mathbb{R}^n$.

- $\dim V + \dim V^\perp = n$.
- vectors in V are linearly independent from vectors in $V^\perp$.

Thus, any vector $\vec{y} \in \mathbb{R}^n$ has a unique decomposition

$$\underset{\text{in } V}{\vec{y}_1} \qquad \vec{y} = \vec{y}_1 + \vec{y}_2 \quad \text{uniquely.}$$

$\vec{y}_1$ in V, $\vec{y}_2$ in $V^\perp$.



$$\mathrm{Proj}_V(\vec{y}) = \vec{y}_1.$$

$$\mathrm{Proj}_{V^\perp}(\vec{y}) = \vec{y}_2.$$

Q: Given a vector $\vec{v}$, how do we find $\mathrm{Proj}_V(\vec{v})$ for a given subspace V?

Note that we already saw that if $V = \mathrm{span}\{\vec{u}\}$, then $\mathrm{Proj}_V(\vec{v}) = \dfrac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|^2}\vec{u}$.

Theorem: Let $\{\vec{u}_1, \dots, \vec{u}_p\}$ be an orthogonal basis for V. Then

$$\mathrm{Proj}_V(\vec{v}) = \underbrace{\frac{\vec{v}\cdot\vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1}_{} + \dots + \underbrace{\frac{\vec{v}\cdot\vec{u}_p}{\|\vec{u}_p\|^2}\vec{u}_p}_{}$$

$$= \mathrm{Proj}_{\vec{u}_1}(\vec{v}) + \dots + \mathrm{Proj}_{\vec{u}_p}(\vec{v}).$$

This allows us to compute the matrix of Proj_V: start with an orthonormal basis for V.

$$\mathrm{Proj}_V(\vec{v}) = \left(\frac{\vec{v}\cdot\vec{u}_1}{\|\vec{u}_1\|^2}\right)\vec{u}_1 + \dots + \left(\frac{\vec{v}\cdot\vec{u}_p}{\|\vec{u}\|_p}\right)\vec{u}_p$$

$$= (\vec{v}\cdot\vec{u}_1)\vec{u}_1 + \dots + (\vec{v}\cdot\vec{u}_p)\vec{u}_p.$$

$$= [\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_p]\begin{bmatrix} \vec{v}\cdot\vec{u}_1 \\ \vec{v}\cdot\vec{u}_2 \\ \vdots \\ \vec{v}\cdot\vec{u}_p \end{bmatrix}$$

$$= \underbrace{[\vec{u} \ \vec{u}_2 \ \dots \ \vec{u}_p]}_{}\begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vdots \\ \vec{u}_p^T \end{bmatrix}\vec{v}.$$

$$= \quad U \quad U^T \ \vec{v}.$$

Thm: Let $V \subseteq \mathbb{R}^n$ be a subspace, and fix an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_p\}$ for V. Let $U = [\vec{u}_1 \dots \vec{u}_p]$

$$\mathrm{Proj}_V(\vec{v}) = UU^T\vec{v}.$$

E.g. Compute the matrix of the orthogonal projection in $\mathbb{R}^3$ onto the subspace $\text{span}\left\{\begin{bmatrix}1\\-2\\2\end{bmatrix}\right\} = V$.

First, need to find an O.N.B. for V.
Since $V = \text{span}\{\vec{u}\}$ $\Rightarrow$ normalize $\vec{u}$.

$$\hat{u} = \frac{1}{3}\begin{bmatrix}1\\-2\\2\end{bmatrix}.$$

$$\text{Proj}_V = U U^T = \frac{1}{3}\begin{bmatrix}1\\-2\\2\end{bmatrix} \frac{1}{3}\begin{bmatrix}1 & -2 & 2\end{bmatrix}$$

$$= \frac{1}{9}\begin{bmatrix}1 & -2 & 2\\-2 & 4 & -4\\2 & -4 & 4\end{bmatrix}.$$

Thm: $\text{Proj}_V(\vec{v})$ is the point in V that is closest to $\vec{v}$. i.e. it is the best approximation of $\vec{v}$ in V.

E.g. $\left\{\vec{u}_1 = \begin{bmatrix}2\\5\\-1\end{bmatrix}, \vec{u}_2 = \begin{bmatrix}-2\\1\\1\end{bmatrix}\right\}$ is an orthogonal basis for V.
Find the closest point in V to $\vec{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$.

Sol. $$\text{Proj}_V(\vec{v}) = \frac{2(1)+5(2)+(-1)(3)}{2^2+5^2+(-1)^2}\begin{bmatrix}2\\5\\-1\end{bmatrix} + \frac{(-2)(1)+1(2)+1(3)}{(-2)^2+1^2+1^2}\begin{bmatrix}-2\\1\\1\end{bmatrix}$$

But what if we aren't given an orthogonal basis?

E.g. $\text{Nul}\left(\underbrace{\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & -1 & 0 & 3 \end{bmatrix}}_{A}\right) = \text{span}\left\{ \underbrace{\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{\vec{v}_1}, \underbrace{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\vec{v}_2} \right\}$.

$\{\vec{v}_1, \vec{v}_2\}$ is a basis for Nul(A).

But they are not orthogonal.

$$\vec{v}_1 \cdot \vec{v}_2 = (-1)(-2) + (-1)(1) + 0 + 0 = 1 \neq 0.$$

Start from $\{\vec{v}_1, \vec{v}_2\}$, produce $\{\vec{u}_1, \vec{u}_2\}$ that is orthogonal.

- $\vec{u}_1 = \vec{v}_1$
- $\vec{u}_2 = \cancel{\text{proj}_{\ldots}}\ \vec{v}_2 - \text{proj}_{\vec{u}_1}(\vec{v}_2) \quad \in \text{span}\{\vec{v}_1, \vec{v}_2\}$.

$$= \vec{v}_2 - \frac{v_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1.$$

$$= \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -5/3 \\ 4/3 \\ -1/3 \\ 1 \end{bmatrix}.$$

* <u>The Gram-Schmidt Orthogonalization process.</u>

Given a basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ of a subspace V, we can find a new basis $\Theta = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$ with the properties

1) Θ is orthogonal.
2) $\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$ for $1 \leq k \leq p$.

$$\vec{u}_1 = \vec{v}_1$$

$$\vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1}(\vec{v}_2) = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1.$$

$$\vec{u}_3 = \vec{v}_3 - \operatorname{proj}_{\{\vec{u}_1, \vec{u}_2\}}(\vec{v}_3) = \vec{v}_3 - \operatorname{proj}_{\vec{u}_1}(\vec{v}_3) - \operatorname{proj}_{\vec{u}_2}(\vec{v}_3)$$

$$= \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2.$$

$$\vdots$$

$$\vec{u}_p = \vec{v}_p - \frac{\vec{v}_p \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 - \cdots - \frac{\vec{v}_p \cdot \vec{u}_{p-1}}{\|\vec{u}_{p-1}\|^2}\vec{u}_{p-1}.$$

The Gram-Schmidt process can also produce an orthonormal basis $\hat{\theta} = \{\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_p\}$ by normalizing vectors $\vec{u}_1, \ldots, \vec{u}_p$. That is,

$$\hat{\theta} = \left\{ \frac{\vec{u}_1}{\|\vec{u}_1\|}, \ldots, \frac{\vec{u}_p}{\|\vec{u}_p\|} \right\}.$$

E.g. Find an orthonormal basis for Col A where

$$A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ -2 & 0 & -2 & 0 \\ 2 & 4 & 6 & 9 \end{bmatrix}$$

$\vec{v}_1$ $\vec{v}_2$ $\vec{v}_3$ form a basis for Col(A)

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}. \rightsquigarrow \hat{u}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} - \frac{9}{9}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \rightsquigarrow \hat{u}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 = \cdots = \begin{bmatrix} -2 \\ -1/2 \\ 1/2 \end{bmatrix} \rightsquigarrow \hat{u}_3 = \frac{\sqrt{2}}{3}\begin{bmatrix} -2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

The triangular pattern of the G-S process can be summarized as follows.

$$\underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \end{bmatrix}}_{A_{n\times p}} = \underbrace{\begin{bmatrix} \hat{u}_1 & \hat{u}_2 & \dots & \hat{u}_p \end{bmatrix}}_{Q_{n\times p}} \underbrace{\begin{bmatrix} \langle \hat{u}_1, \vec{v}_1 \rangle & \langle \hat{u}_1, \vec{v}_2 \rangle & \dots & \langle \hat{u}_1, \vec{v}_p \rangle \\ 0 & \langle \hat{u}_2, \vec{v}_2 \rangle & \dots & \langle \hat{u}_2, \vec{v}_p \rangle \\ 0 & 0 & \ddots & \\ \vdots & \vdots & & \\ 0 & 0 & & \langle \hat{u}_p, \vec{v}_p \rangle \end{bmatrix}}_{R_{p\times p}}$$

E.g. $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ -2 & -1 \end{bmatrix}$ with columns $\vec{v}_1$, $\vec{v}_2$. Find an O.N.B for Col(A) using G-S.

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \Rightarrow \hat{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} - \frac{3}{9}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ -1/3 \end{bmatrix}$$

$\hat{u}_2 = \vec{u}_2$ as $\|\vec{u}_2\| = 1$.

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \hat{u}_1 & \hat{u}_2 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 2 & 0 \\ 1 & 1 \\ -2 & -1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 2/3 & -2/3 \\ 1/3 & 2/3 \\ -2/3 & -1/3 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix}}_{R}$$