Recall

A and M <u>similar</u> if $\exists P$ invertible so that $M = P^{-1} A P$.

A is <u>diagonalizable</u> if it is to D (diagonal matrix)

If A and M <u>similar</u> $(M = P^{-1} A P)$ then : (Prop. 5.6)

(1) $\chi_A(\lambda) = \chi_M(\lambda)$

(2) $\sigma(A) = \sigma(M)$

(3) v is an eigenvector of M for λ

$\Leftrightarrow$ Pv is an eigenvector of A for λ

A is diagonalizable $\Leftrightarrow$ has an eigenbasis. (Thm 5.7)

A has n distinct eigenvalues $\Rightarrow$ A is diagonalizable (Cor 5.9)

Let $\mathcal{B} = \{b_1 \dots b_n\}$ be an eigenbasis of A

and λ_i the eval. for the evec. b_i of A

$\Rightarrow$ $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ for $P = P_{\mathcal{B}} = (b_1 \dots b_n)$

D unique up to order of the λ_i

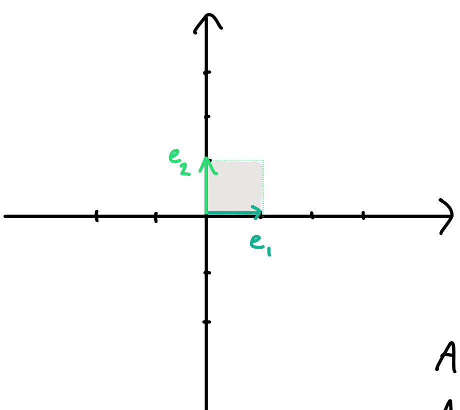
Geometric interpretation:

Recall that changing basis geometrically means choosing new axes for our coordinate system.

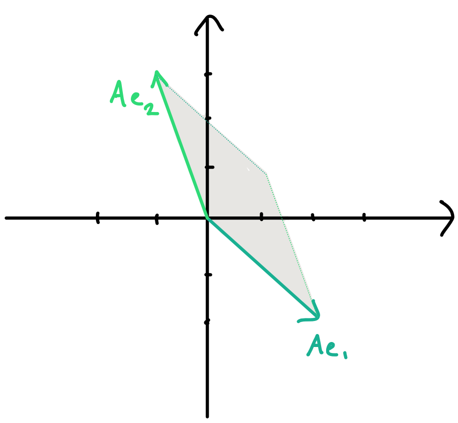
Changing from any basis (e.g. standard) to an eigenbasis $\mathcal{B}$ for a matrix A means choosing axes that are given by eigenvectors of A. In this coordinate system, the map $f(x) = Ax$ is a "stretch map" that in the direction of each axis stretches by the eigenvalue of A whose eigenvector defined the axis.

Ex. $A = \begin{pmatrix} 2 & -1 \\ -2 & 3 \end{pmatrix} \Rightarrow \lambda_1 = 1, \lambda_2 = 4$

$\Rightarrow b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}$ and $D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$



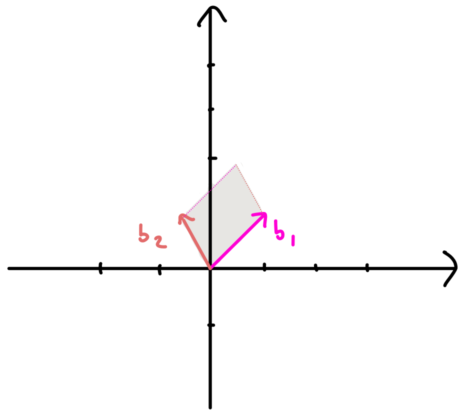
$\xrightarrow{A}$



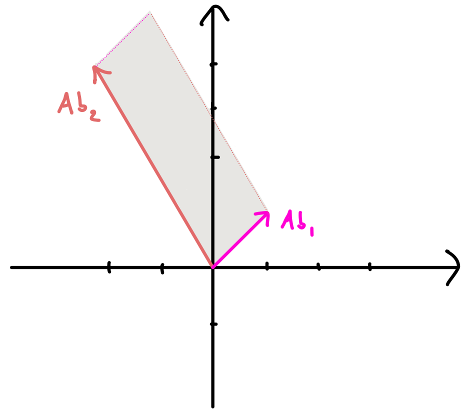
the unit square goes to some arbitrary parallelogram

$Ae_1 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$

$Ae_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$



$\xrightarrow{A}$



parallelograms with same side directions

$Ab_1 = 1 \cdot b_1$

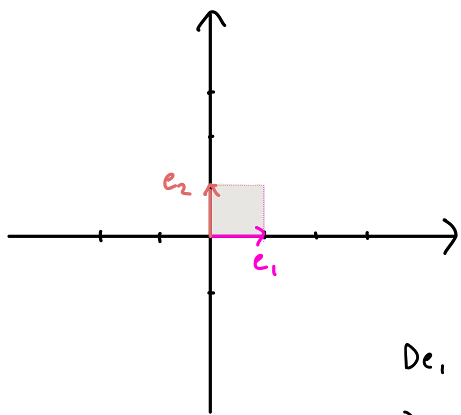
$Ab = 4 \cdot b_2$

Change of basis

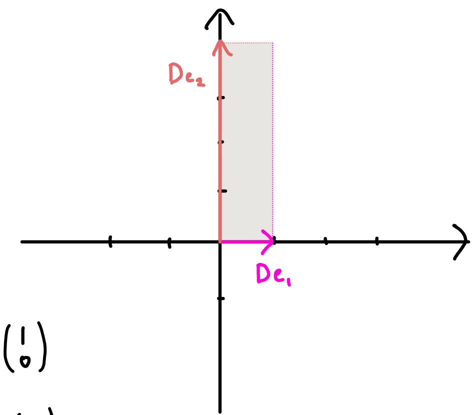
$b_1 \rightsquigarrow [b_1]_B = e_1$

$b_2 \rightsquigarrow [b_2]_B = e_2$

i.e. we interpret the straight lines given by b_1 and b_2 as our new coordinate axes.



$\xrightarrow{D}$



the unit square goes to rectangle alined with axes

$$De_1 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$De_2 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$

<u>We learn:</u> if a matrix $A \in \mathbb{R}^{n \times n}$ is diagonalizable, then this means that the linear map $f(x) = Ax$ expressed in a suitable basis (namely the eigenbasis of A) will fix each coordinate axis (as a set).

On each coordinate axis, the map is a stretch by a factor λ_i, where where λ_i is an eigenvalue of A.

In consequence, the unit square (or n-cube in $\mathbb{R}^n$) is mapped to a rectangle with side lengths $\lambda_1 \dots \lambda_n$ (evals with multiplicity).

Lastly, on popular demand:

Change of bases revisited

Let $\mathcal{B} = \{b_1 \dots b_n\}$ be a basis of $\mathbb{R}^n$

and $\{e_1 \dots e_n\}$ the standard basis.

Let $P_{\mathcal{B}} = (b_1 \dots b_n)$ be the matrix with columns $b_1 \dots b_n$.

$$\Rightarrow \quad P_{\mathcal{B}}\, e_i = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = b_i \quad \Rightarrow \quad \boxed{P_{\mathcal{B}}\, e_i = b_i} \quad \text{(Eq. 1)}$$

Now let $v \in \mathbb{R}^n$.

Recall: $[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ means $\lambda_1 b_1 + \dots + \lambda_n b_n = v$

Compute: $P_{\mathcal{B}} [v]_{\mathcal{B}} = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 b_1 + \dots + \lambda_n b_n = v$

$$\Rightarrow \quad \boxed{P_{\mathcal{B}} [v]_{\mathcal{B}} = v} \quad \text{(Eq. 2)}$$

That is why we call $P_{\text{Standard} \leftarrow \mathcal{B}} := P_{\mathcal{B}}$

the change-of-basis matrix from $\mathcal{B}$ to standard.

However, it is still true that $P_{\mathcal{B}}$ maps e_i to b_i

(and not b_i to e_i).

Now let $B = \{b_1 \dots b_n\}$ and $S = \{s_1 \dots s_n\}$ be two bases of $\mathbb{R}^n$.

Define $P_{S\leftarrow B} := P_S^{-1} P_B$ (or equivalently $P_{S\leftarrow B} := ([b_1]_C, \dots, [b_n]_C)$)

So $$P_{S\leftarrow B}[v]_B = P_S^{-1}\big(\underbrace{P_B [v]_B}_{= v \text{ (Eq. 2) for } B}\big) = P_S^{-1} v \underset{\text{(Eq. 2) for } S}{=} [v]_S$$

$$\Rightarrow \boxed{P_{S\leftarrow B}[v]_B = [v]_S}$$

We call $P_{S\leftarrow B}$ the change-of-basis matrix from B to S.

Warning:

Equation (Eq. 1) does not carry forward to this situation!

In general: $P_{S\leftarrow B}(s_i) \neq b_i$

("Reason": $P_{S\leftarrow B}(s_i) = P_S^{-1} P \; s_i = ??$)

e.g. $B = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$, $S = \left\{ \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a counterexample.

6. Orthogonality and least squares

6.1 The standard inner product

Def. Let $u, v \in \mathbb{R}^n$ be two vectors. Their standard inner product is:

$$u \cdot v := u^T v \in \mathbb{R}$$

Notation

We will refer to the standard inner product simply as the inner product for short.

(Also called: the Euclidean scalar product or dot product)

Rmk $u \cdot v = u^T v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$

Ex.

- $u = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \; v = \begin{pmatrix} 5 \\ 1 \\ -1 \\ 0 \end{pmatrix} \Rightarrow u \cdot v = 1 \cdot 5 + 2 \cdot 1 + 3 \cdot (-1) + 4 \cdot 0 = 8$

- $e_i \cdot e_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

Calculus rules for the inner product: Let $u, v, w \in \mathbb{R}^n$

(i) $u \cdot v = v \cdot u$

(ii) $(u+v) \cdot w = u \cdot w + v \cdot w$

(iii) $(\lambda u) \cdot v = u \cdot (\lambda v) = \lambda (u \cdot v) \quad \forall \lambda \in \mathbb{R}$

(iv) $u \cdot u \geq 0$

(v) $\underbrace{u \cdot u}_{= u_1^2 + \dots + u_n^2} = 0$ if and only if $u = 0$

From the calculus rules, we can derive the following fact about the inner product of a linear combination of vectors with another vector:

$$(\lambda_1 v_1 + \dots + \lambda_n v_n) \cdot u = \lambda_1 v_1 \cdot u + \dots + \lambda_n v_n \cdot u$$

Def. The length (or norm) of a vector $u \in \mathbb{R}^n$ is:

$$\|u\| := \sqrt{u \cdot u} = \sqrt{u_1^2 + \dots + u_n^2}$$

In particular, we have:
- $\|\lambda u\| = |\lambda| \, \|u\| \quad \forall \lambda \in \mathbb{R}$
- $\|u\|^2 = u \cdot u$

Given a vector $u \in \mathbb{R}^n$, $u \neq 0$, we can normalize it, that is: we consider the vector $\frac{u}{\|u\|}$ which has the same direction as u but has length $= 1$:

$$\left\| \frac{u}{\|u\|} \right\| = \Big\| \underbrace{\frac{1}{\|u\|}}_{=\lambda} u \Big\| = \underbrace{\left| \frac{1}{\|u\|} \right|}_{= \frac{1}{\|u\|} \text{ since } \frac{1}{\|u\|} \geq 0} \cdot \|u\| = \frac{1}{\|u\|} \|u\| = 1$$

__Def.__ The distance of two vectors $u, v \in \mathbb{R}^n$ is $\|v - u\|$.

__Def.__ Two vectors $u, v \in \mathbb{R}^n$ are called orthogonal if $u \cdot v = 0$.

(Tipp: draw / compute some examples in $\mathbb{R}^2$ or $\mathbb{R}^3$ to convince yourself that the concepts of length and orthogonality as defined here actually make sense.)

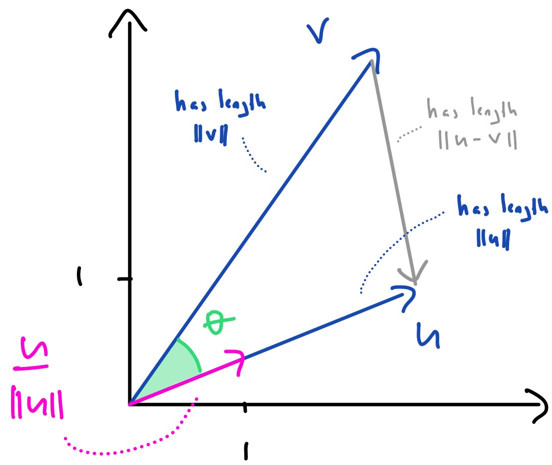
__Def.__ For two non-zero vectors $u, v \in \mathbb{R}^n$ the angle θ between u and v is defined by :

$$\cos(\theta) = \frac{u \cdot v}{\|u\| \, \|v\|}$$

(In order for this definition to be valid, $\frac{u \cdot v}{\|u\| \, \|v\|}$ must be in the interval $[-1, 1]$ because cos only takes values in $[-1, 1]$.

This is guaranteed by the Cauchy-Schwarz inequality which says that : $\| u \cdot v \| \leq \|u\| \, \|v\|$ (we will prove it later))

Geometric interpretation :



<u>Thm 6.1</u> $u, v \in \mathbb{R}^n$ are orthogonal if and only if

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2$$

<u>Rmk</u> : Draw a picture for Thm 6.1 and you will see that $\boxed{\Rightarrow}$ is the pythagorean theorem !

<u>Proof</u>

$$\|u+v\|^2 \underset{\text{Def. of norm}}{=} (u+v)\cdot(u+v) \underset{\text{Calculus}}{=} u\cdot u + u\cdot v + \underbrace{v\cdot u}_{\substack{= u\cdot v \\ \text{calculus}}} + v\cdot v$$

$$= u\cdot u + 2(u\cdot v) + v\cdot v = \|u\|^2 + \|v\|^2 + 2(u\cdot v)$$

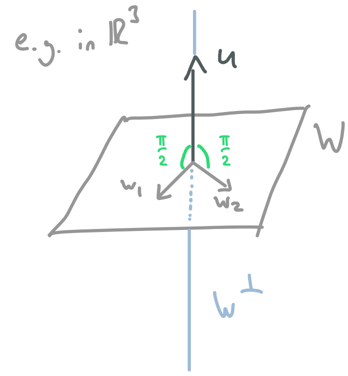
We know : $u\cdot v = 0$ if and only if u, v orthogonal

Hence $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ if and only if u, v orthogonal

$\square$

Orthogonal complements

Def. Let W be a subspace of $\mathbb{R}^n$. A vector $u \in \mathbb{R}^n$ is called orthogonal to W if $u \cdot w = 0$ for all $w \in W$



We call the set of all vectors u that are orthogonal to W, the orthogonal complement of W:

$$W^\perp := \{u \in \mathbb{R}^n : u \cdot w = 0 \ \forall w \in W\}$$

Lemma 6.2 (i) $W^\perp$ is a subspace of $\mathbb{R}^n$

(ii) If B spans W, then we can write $W^\perp$ as

$$W^\perp := \{u \in \mathbb{R}^n : u \cdot b = 0 \ \forall b \in B\}$$

Hint: compare each statement in the lemma to the above picture.

Proof See exercises. → HW 12

Another easy-to-observe but useful fact: $(W^\perp)^\perp = W$.

In particular (ii) in Lemma 6.2 is very useful for finding $W^\perp$ for a given W:

Ex. Let $W = \left\{ \begin{pmatrix} x \\ x \\ z \end{pmatrix} : x, z \in \mathbb{R} \right\}$

Q: Find $W^\perp$

We can write $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$$\Rightarrow W^\perp = \left\{ u \in \mathbb{R}^3 : \underbrace{u \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 0}_{\text{Cond. I}} \text{ and } \underbrace{u \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0}_{\text{Cond II}} \right\}$$

Lemma 6.2. (ii)

Let $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in W^\perp$

By Cond. I: $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 0 \iff u_1 + u_2 = 0 \iff \underline{u_1 = -u_2}$

By Cond II: $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \iff \underline{u_3 = 0}$

linear system with unknowns u_1, u_2, u_3

$$\Rightarrow W = \left\{ \begin{pmatrix} t \\ -t \\ 0 \end{pmatrix} : t \in \mathbb{R} \right\} = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$$

Solution space of the system

Basis of W

Thm. 6.3 Let $A \in \mathbb{R}^{m \times n}$. Then:

(1) $\mathrm{Row}(A)^\perp = \mathrm{Ker}(A)$

(2) $\mathrm{Col}(A)^\perp = \mathrm{Ker}(A^T)$

For the proof, the following observation will be helpful:

Rmk. For matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$ denote

the rows of A by $a_1, \dots, a_m \in \mathbb{R}^n$ and

the columns of B by $b_1, \dots, b_k \in \mathbb{R}^n$

$$\Rightarrow AB = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \cdot \begin{pmatrix} b_1 & b_2 & \dots & b_k \end{pmatrix} = \begin{pmatrix} a_1\cdot b_1 & a_1\cdot b_2 & \dots & a_1\cdot b_k \\ a_2\cdot b_1 & a_2\cdot b_2 & \dots & a_2\cdot b_k \\ \vdots & & & \\ a_m\cdot b_1 & a_m\cdot b_2 & \dots & a_m b_k \end{pmatrix}$$

In particular for a vector $v \in \mathbb{R}^n$ $(v = B \in \mathbb{R}^{n\times 1})$

$$\Rightarrow Av = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \cdot v = \begin{pmatrix} a_1\cdot v \\ a_2\cdot v \\ \vdots \\ a_m\cdot v \end{pmatrix}$$

Proof (1) $v \in \mathrm{Row}(A)^\perp$

(*) $\begin{cases} \mathrm{Row}(A) = \mathrm{Span}\{a_1,\dots,a_n\} \\ + \text{Lemma } 6.2.(ii) \end{cases}$

$\overset{(*)}{\Leftrightarrow}$ $v\cdot a_i = 0 \quad \forall$ rows a_i of A

$\Leftrightarrow Av = 0$ by the above Remark.

$\Leftrightarrow v \in \mathrm{Ker}(A)$

(2) Recall that $\mathrm{Col}(A) = \mathrm{Row}(A^T)$

$$\Rightarrow \mathrm{Col}(A)^\perp = \mathrm{Row}(A^T)^\perp \underset{(1)}{=} \mathrm{Ker}(A^T) \qquad \square$$

Ex. Let $W = \text{Span}\left\{ \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\4\\1\\0\end{pmatrix}, \begin{pmatrix}3\\5\\2\\0\end{pmatrix} \right\}$ Q: $W^\perp = ?$

Solution (using Thm 6.3):

Define the matrix $A \in \mathbb{R}^{4\times 3}$ by $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$

$\Rightarrow W = \text{Col}(A)$

By Thm 6.3 $\Rightarrow W^\perp = \text{Ker}(A^T)$

We can find the $\text{Ker}(A^T)$ by as usual:

- Row-reduction:

$$A^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 0 \\ 3 & 4 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(At first glance we see: two free variables, so $\dim \text{Ker} A = 2$

- $x_4 = t$, $x_3 = s$ $\Rightarrow$ $x_2 = s + 2t$, $x_1 = -2s - 3t$

$$\Rightarrow W^\perp = \left\{ s\begin{pmatrix}-2\\1\\1\\0\end{pmatrix} + t\begin{pmatrix}-3\\2\\0\\1\end{pmatrix} : t, s \in \mathbb{R} \right\}$$

$$= \text{Span}\left\{ \begin{pmatrix}-2\\1\\1\\0\end{pmatrix}, \begin{pmatrix}-3\\2\\0\\1\end{pmatrix} \right\}$$

$\square$

Recall the <u>Cauchy-Schwarz inequality</u>: $|u \cdot v| \leq \|u\| \, \|v\|$

<u>Proof</u> Case 1: if $u = 0$ or $v = 0$, then $u \cdot v = 0 = \|u\| \, \|v\|$.

Case 2 $u \neq 0$ and $v \neq 0$.

Trick: define $z = \frac{u}{\|u\|} - \frac{v}{\|v\|}$

$$\Rightarrow \quad 0 \leq \|z\|^2 = z \cdot z = \left(\frac{u}{\|u\|} - \frac{v}{\|v\|} \right) \cdot \left(\frac{u}{\|u\|} - \frac{v}{\|v\|} \right)$$

$$= \underbrace{\frac{u \cdot u}{\|u\| \, \|u\|}}_{=1} + \underbrace{\frac{v \cdot v}{\|v\| \, \|v\|}}_{=1} - \frac{2 \, u \cdot v}{\|u\| \, \|v\|}$$

because $v \cdot v = \|v\|^2 = \|v\| \, \|v\|$

$$\Rightarrow \ 0 \leq 2 - 2 \frac{u \cdot v}{\|u\| \, \|v\|} \ \Rightarrow \ \frac{u \cdot v}{\|u\| \, \|v\|} \leq 1 \ \Rightarrow \ u \cdot v \leq \|u\| \, \|v\| \quad (*)$$

If $u \cdot v \geq 0$ then $|u \cdot v| = u \cdot v \overset{(*)}{\leq} \|u\| \, \|v\|$

If $u \cdot v \leq 0$ then $|u \cdot v| = -u \cdot v \underset{(*)}{\leq} \|-u\| \, \|v\| = \|u\| \, \|v\|$. $\square$

<u>Thm. 6.4</u> Let V be a vectorspace of dimension n and $W \subseteq V$ a subspace. Then: $\dim W + \dim W^\perp = n$

<u>Proof</u> $\longrightarrow$ HW 12

Hint: Let $b_1 \ldots b_k$ a basis of W and $M := (b_1 \ldots b_k) \in \mathbb{R}^{? \times ?}$

Check that $W^\perp = \mathrm{Ker}(M)$.

6.2. Orthogonal sets

<u>Def.</u> A set $S \subseteq \mathbb{R}^n$ is called orthogonal if $v \cdot w = 0 \quad \forall v \neq w \in S$.

<u>Ex.</u> (i) $S = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ is orthogonal since $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -2 + 2 = 0$

(ii) $S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$. Is S orthogonal?

$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = -2 + 2 + 0 = 0 \quad \rightarrow$ ok

$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 1 + 2 + 0 = 3 \neq 0 \quad \rightarrow$ not ok

$\Rightarrow$ S is <u>not</u> orthogonal.

(iii) Let $u \in \mathbb{R}^n$: $\{0, u\} \subseteq \mathbb{R}^n$ is orthogonal as $u \cdot \underset{\text{zero vector}}{0} = \underset{\text{zero in } \mathbb{R}}{0}$

(iv) $\{e_1, \dots, e_n\} \subseteq \mathbb{R}^n$ is orthogonal as $e_i \cdot e_j = 0$ for $i \neq j$

e.g. in $\mathbb{R}^3$: $e_1 \cdot e_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 + 0 + 0 = 0$

<u>Rmk</u> If $S \subseteq \mathbb{R}^n$ orthogonal and $\tilde{S} \subseteq S$ subset then $\tilde{S}$ is also orthogonal.

We are particularly interested in orthogonal bases of $\mathbb{R}^n$.

Thm. 6.5 Let $W \subset \mathbb{R}^n$ be a subspace and $B = \{b_1 \dots b_k\}$ an orthogonal basis of W. Then for each $w \in W$, the unique coefficients λ_i for which $w = \lambda_1 b_1 + \dots + \lambda_k b_k$ are:

$$\lambda_i = \frac{w \cdot b_i}{\|b_i\|^2}$$

Proof Let $w \in W$. Since B is a basis, there exist unique $\lambda_1 \dots \lambda_k$ so that $w = \lambda_1 b_1 + \dots + \lambda_k b_k$

Compute λ_i by computing $b_i \cdot w$:

$$b_i \cdot w = b_i \cdot (\lambda_1 b_1 + \dots + \lambda_k b_k)$$

$$= \lambda_1 \underbrace{b_i \cdot b_1}_{=0} + \dots + \lambda_i \underbrace{b_i \cdot b_i}_{= \|b_i\|^2} + \dots + \lambda_k \underbrace{b_i \cdot b_k}_{=0} = \lambda_i \|b_i\|^2$$

Ex. $S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}$ Q: Is $v = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ in Span(S)?

Q rephrased: Do there exist $\lambda_1, \lambda_2, \lambda_3$ so that $v = \lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 s_3$

We observe: S is orthogonal basis of Span(S).

Assume that $v \in \text{Span}(S)$

Thm 6.5 $\Rightarrow$ $v = \lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 s_3$ for $\lambda_i = \frac{v \cdot s_i}{\|s_i\|^2}$

Compute $\lambda_1 = \frac{2}{30}$, $\lambda_2 = \frac{-2}{3}$, $\lambda_3 = \frac{-2}{3}$

$$\Rightarrow \lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 s_3$$

$$= \frac{2}{30}\begin{pmatrix}1\\2\\3\\4\end{pmatrix} - \frac{2}{3}\begin{pmatrix}-1\\-1\\1\\0\end{pmatrix} - \frac{2}{3}\begin{pmatrix}-1\\0\\-1\\1\end{pmatrix} \overset{?}{=} \begin{pmatrix}2\\0\\0\\0\end{pmatrix} = v$$

First coordinate $\Rightarrow \frac{2}{30} + \frac{2}{3} + \frac{2}{3} \overset{?}{=} 2$ $\rightarrow$ no!

$\Rightarrow v \notin \mathrm{Span}(S)$.

Ex. $\mathcal{B} = \left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix} \right\}$ is an orthog. basis of $\mathbb{R}^2$.

Q: Find $[w]_{\mathcal{B}}$ for $w = \begin{pmatrix}2\\1\end{pmatrix}$

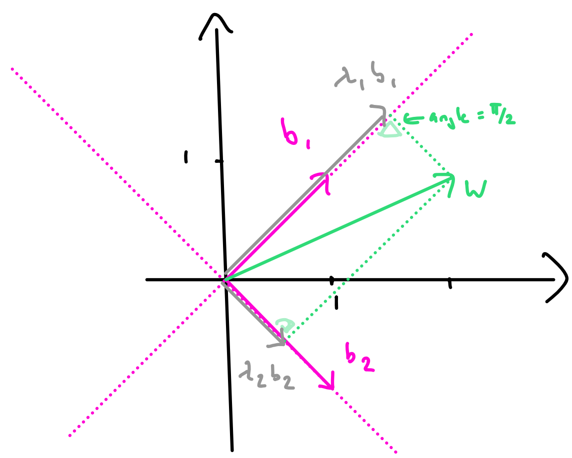
$\rightarrow$ we need to find λ_1, λ_2 so that $w = \lambda_1 \begin{pmatrix}1\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}1\\-1\end{pmatrix}$

Thm 6.5 :

$$\left.\begin{aligned} \lambda_1 &= \frac{w \cdot b_1}{\|b_1\|^2} = \frac{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}2\\1\end{pmatrix}}{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}} = \frac{3}{2} \\ \lambda_2 &= \frac{w \cdot b_2}{\|b_2\|^2} = \frac{\begin{pmatrix}1\\-1\end{pmatrix}\cdot\begin{pmatrix}2\\1\end{pmatrix}}{\begin{pmatrix}1\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\-1\end{pmatrix}} = \frac{1}{2} \end{aligned}\right\} \Rightarrow [w]_{\mathcal{B}} = \begin{pmatrix}3/2\\1/2\end{pmatrix}$$

Geometric interpretation :

$\lambda_i b_i$ is the orthogonal projection of w onto the line $L_i = \mathrm{Span}\{b_i\}$.



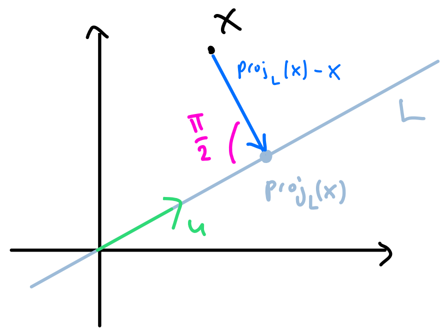
Recall that for $u \in \mathbb{R}^n$, $u \neq 0$, $L := \mathrm{Span}\{u\}$ is a one-dimensional subspace of $\mathbb{R}^n$. We call L a line.

Def. For $u \in \mathbb{R}^n$, $u \neq 0$ and $L = \mathrm{Span}\{u\}$ we define the orthogonal projection of $\mathbb{R}^n$ onto L to be the map

$$\mathrm{proj}_L : \mathbb{R}^n \to L \quad \text{by} \quad \mathrm{proj}_L(x) := \underbrace{\frac{x \cdot u}{u \cdot u}}_{\in \mathbb{R}} \underbrace{u}_{\in \mathbb{R}^n}$$

"λ_i" "b_i"

Rmk. It is called the orthog. projection because $\mathrm{proj}_L(x) - x$ is orthogonal to L.



Ex. Consider the line $L = \left\{ \begin{pmatrix} x \\ 2x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$ in $\mathbb{R}^3$

Q: Find the orthogonal projection of $p = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$ onto L.

Observe L is a subspace of $\mathbb{R}^3$ and $L = \mathrm{Span}\left\{ \overbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}^{=:u} \right\}$.

$$\Rightarrow \mathrm{proj}_L(p) = \frac{u \cdot p}{\|u\|^2} \cdot u = \frac{\begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\|^2} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \frac{9}{5} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

Orthonormal sets and matrices

Def. A set $\{b_1, \dots, b_k\} \subseteq \mathbb{R}^n$ is called orthonormal if it is orthogonal and $\|b_i\| = 1 \ \forall i = 1 \dots k$.

A set $\mathcal{B} \subseteq \mathbb{R}^n$ is an orthonormal basis (ONB) if it is a basis consisting of orthog. vectors of length 1.

Rmk. Every orthogonal set that does not contain the zero vector and every orthogonal basis always be made orthonormal by normalizing the vectors by replacing each element v of the set by $\frac{v}{\|v\|}$.

Ex. We know $\mathcal{B} = \{ \underbrace{\begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{=b_1}, \underbrace{\begin{pmatrix} -1 \\ 2 \end{pmatrix}}_{=b_2} \}$ is an orthogonal basis of $\mathbb{R}^2$ but it is not orthonormal because:

$$\left\| \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5} \neq 1$$

Normalize: $\tilde{b}_1 := \frac{b_1}{\|b_1\|} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\quad \tilde{b}_2 := \frac{b_2}{\|b_2\|} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

$\Rightarrow \tilde{\mathcal{B}} := \{ \tilde{b}_1, \tilde{b}_2 \}$ is an ONB.

Ex. $\{e_1, \dots, e_n\}$ is an ONB for $\mathbb{R}^n$.

Thm 6.6. Let $U \in \mathbb{R}^{m \times n}$ be a matrix.

but $UU^T \neq I_m$ (see Rmk on page 26)

(1) U has orthonormal columns if and only if $U^T U = I_n$

(2) If U has orthonormal columns, then also the following hold:

(a) $\|Ux\| = \|x\|$

(b) $Ux \cdot Uy = x \cdot y$

(c) $Ux \cdot Uy = 0 \iff x \cdot y = 0$

(In other words: the linear map with matrix U preserves the inner product.)

Q: Does $U^T U = I_n$ mean the same as $U^T = U^{-1}$?

$\to$ Only for square matrices!

Def. A square matrix $U \in \mathbb{R}^{n \times n}$ is called orthogonal if its columns are orthonormal.

In conclusion:

often useful in applications

$$U \in \mathbb{R}^{n \times n} \text{ is orthogonal} \iff U^T U = I_n \iff U^T = U^{-1}.$$

square matrix

Proof (of Thm.):

(1) Let $u_1 \dots u_n$ be the columns of $U \in \mathbb{R}^{m \times n}$.

Then
$$U^T U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} u_1 & \dots & u_n \end{pmatrix} = \begin{pmatrix} u_1 \cdot u_1, & u_1 \cdot u_2 & \dots & u_1 \cdot u_n \\ \vdots & \vdots & & \vdots \\ u_n u_1, & u_n \cdot u_2 & \dots & u_n \cdot u_n \end{pmatrix}$$

So $U^T U = I_n \iff \begin{cases} u_i \cdot u_i = 1 & \text{for all } i \\ u_i \cdot u_j = 0 & \text{for all } i \neq j \end{cases}$

$u_i \cdot u_i = 1$ means the columns of U are vectors of length 1

$u_i \cdot u_j = 0$ means the columns of U are an orthog. set of vectors

So together they mean that the columns of U are an orthonormal set of vectors

(2) Now assume the u_i are orthonormal and hence also $U^T U = I_n$

def. of inner product $\cdot$

(a) $\|Ux\|^2 = Ux \cdot Ux = (Ux)^T Ux = x^T \underbrace{U^T U}_{= I_n \text{ by (1)}} x$

$= x^T x = x \cdot x = \|x\|^2$

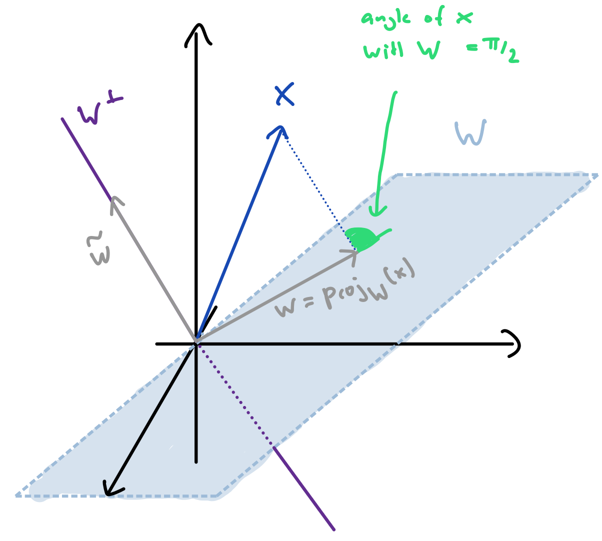
calculus for transpose

(b) Same trick as above. Try it yourselves

(c) Follows from (b). $\square$

6.3. Orthogonal projections

Above, we have learned how to compute the orthogonal projection of a vector $x \in \mathbb{R}^n$ onto a line L.



Q: Orthog. proj. of x onto a larger subspace? e.g. a plane?

Idea: Decompose $x \in \mathbb{R}^n$ into two parts:

$w \in W$ and $\tilde{w} \in W^\perp$ s.o. that $x = w + \tilde{w}$

Thm. 6.7 (Orthogonal projection theorem)

Let W be a subspace of $\mathbb{R}^n$. Then every $x \in \mathbb{R}^n$ can be uniquely written in the form $\boxed{x = w + \tilde{w}}$ with $w \in W$ and $\tilde{w} \in W^\perp$.

In case $B = \{b_1, \dots, b_k\}$ is an orthog. basis of W then $w \in W$ and $\tilde{w} \in W^\perp$ are:

$$w = \frac{x \cdot b_1}{b_1 \cdot b_1} b_1 + \dots + \frac{x \cdot b_k}{b_k \cdot b_k} \cdot b_k \quad \text{and} \quad \tilde{w} = x - w$$

Proof. *Uniqueness*: Assume $x = w_1 + \tilde{w}_1 = w_2 + \tilde{w}_2$ with $w_1, w_2 \in W$ and $\tilde{w}_1, \tilde{w}_2 \in W^\perp$.

$$\Rightarrow \quad 0 = x - x = (w_1 + \tilde{w}_1) - (w_2 + \tilde{w}_2) = w_1 - w_2 + \tilde{w}_2 - \tilde{w}_1$$

$$\Rightarrow \quad \underbrace{w_1 - w_2}_{\in W} = \underbrace{\tilde{w}_2 - \tilde{w}_1}_{\in W^\perp} \quad \Rightarrow \quad w_1 - w_2 = \tilde{w}_2 - \tilde{w}_1 = 0$$

$W \cap W^\perp = \{0\}$

$$\Rightarrow \quad w_1 = w_2 \text{ and } \tilde{w}_1 = \tilde{w}_2 .$$

Existence: In the case $W = \{0\}$ existence is trivial:

$$x = \underbrace{0}_{\in W} + \underbrace{x}_{\in W^\perp = \mathbb{R}^n} \qquad \Rightarrow w = 0,\ \tilde{w} = x .$$

So from now on assume $W \neq \{0\}$.

Let $B = \{b_1, \ldots, b_k\}$ be a basis for W. $(\dim W = k)$

Thm. 6.4 $\Rightarrow \dim W^\perp = n - k$

Let $Q = \{b_{k+1} \ldots b_n\}$ be a basis of $W^\perp$

Claim: $B \cup Q = \{b_1 \ldots b_n\}$ is a basis of $\mathbb{R}^n$

Let us first use the claim to prove the theorem and later prove the claim:

Let $x \in \mathbb{R}^n \overset{\text{Claim}}{\Rightarrow}$ then there exist $\lambda_1 \dots \lambda_n \in \mathbb{R}$ s.t. that:

$$x = \underbrace{\lambda_1 b_1 + \dots + \lambda_k b_k}_{=: w \in W} + \underbrace{\lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n}_{=: \tilde{w} \in W^\perp}$$

Moreover, we know from Thm. 6.5 that if $B = \{b_1 \dots b_k\}$ is orthogonal then we can compute $\lambda_1 \dots \lambda_k$ by: $\lambda_i = \frac{b_i \cdot w}{b_i \cdot b_i} = \frac{b_i \cdot x}{b_i \cdot b_i}$ for $i \le k$.

This completes the proof of the theorem and we are left to prove the claim:

Since $\#(B \cup Q) = n$ it suffices to verify linear independence of $B \cup Q = \{b_1, \dots, b_k, b_{k+1}, \dots, b_n\}$

Let $\lambda_1 \dots \lambda_n \in \mathbb{R}$ such that

$$\lambda_1 b_1 + \dots + \lambda_k b_k + \lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n = 0$$

$$\Rightarrow \underbrace{\lambda_1 b_1 + \dots + \lambda_k b_k}_{\in W} = \underbrace{-(\lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n)}_{\in W^\perp}$$

$\Rightarrow$ both sides must be zero

Since B is lin. indep $\Rightarrow \lambda_1 = \dots = \lambda_k = 0$

Since Q is lin. indep $\Rightarrow \lambda_{k+1} = \dots = \lambda_n = 0$ $\square$

Rmk The uniqueness of w and $\tilde{w}$ (Thm 6.7), tells us that the formulas

$$w = \frac{x \cdot b_1}{b_1 \cdot b_1} b_1 + \dots + \frac{x \cdot b_k}{b_k \cdot b_k} \cdot b_k \quad \text{and} \quad \tilde{w} = x - w$$

Will yield the same w and $\tilde{w}$ regardless of our choice of orthogonal basis $\mathcal{B} = \{b_1, \dots, b_k\}$.

Def. For a subspace $W \subseteq \mathbb{R}^n$ we define the orthogonal projection $\text{proj}_W : \mathbb{R}^n \longrightarrow W$ by

$$\text{proj}_W(x) := \frac{x \cdot b_1}{b_1 \cdot b_1} b_1 + \dots + \frac{x \cdot b_k}{b_k \cdot b_k} \cdot b_k \quad \left(\Leftrightarrow \text{proj}_W(x) = w \right)$$

where $\mathcal{B} = \{b_1, \dots b_n\}$ is an orthogonal basis of W.

Ex. $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$, $W = \text{Span}\{\mathcal{B}\}$

Q: Compute $\text{proj}_W(x)$.

One can check: $\mathcal{B}$ is orthogonal (but not orthonormal)

$$\text{proj}_W(x) = \frac{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \frac{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}}{\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$= \frac{1}{6}(2x_1+x_2+x_3)\begin{pmatrix}2\\1\\1\end{pmatrix} + \frac{1}{2}(x_2-x_3)\begin{pmatrix}0\\1\\-1\end{pmatrix}$$

$$= \begin{pmatrix} \frac{2}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3 \\ \frac{1}{3}x_1 + \left(\frac{1}{6}+\frac{1}{2}\right)x_2 + \left(\frac{1}{6}-\frac{1}{2}\right)x_3 \\ \frac{1}{3}x_1 + \left(\frac{1}{6}-\frac{1}{2}\right)x_2 + \left(\frac{1}{6}+\frac{1}{2}\right)x_3 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

Rmk In response to a student's question, a few more words about the condition $U^TU = I_n$ from Thm. 6.6:

The example $U = \begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$ shows that $UU^T \neq I_m$ in general.

In fact, UU^T can never equal I_m if $n \neq m$.

Reason: The columns $u_1, \dots, u_n$ of U are orthonormal.

Hence, they are in particular orthogonal and therefore independent.

$\Rightarrow$ they are n indep. vectors in $\mathbb{R}^m$

$\Rightarrow$ $n \leq m$

$\Rightarrow$ U has less (or equally many) columns than rows.

Also, since $\dim \mathrm{Col}(A) = \dim \mathrm{Row}(A)$: only n of the m rows of A are indep.

We can read $U^TU = I_n$ as:

The linear transformation $\mathbb{R}^m \to \mathbb{R}^n$ maps the columns of U to the columns of I_n.
↳ $e_1, \dots, e_n$

This is only possible since the columns of U are indep.

By the same argument: If it were true that $UU^T = I_m$ this would imply that that the columns of U^T are indep.

But they are not since they are the rows of U. $\rightarrow UU^T \neq I_m$