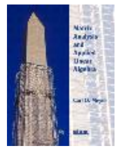
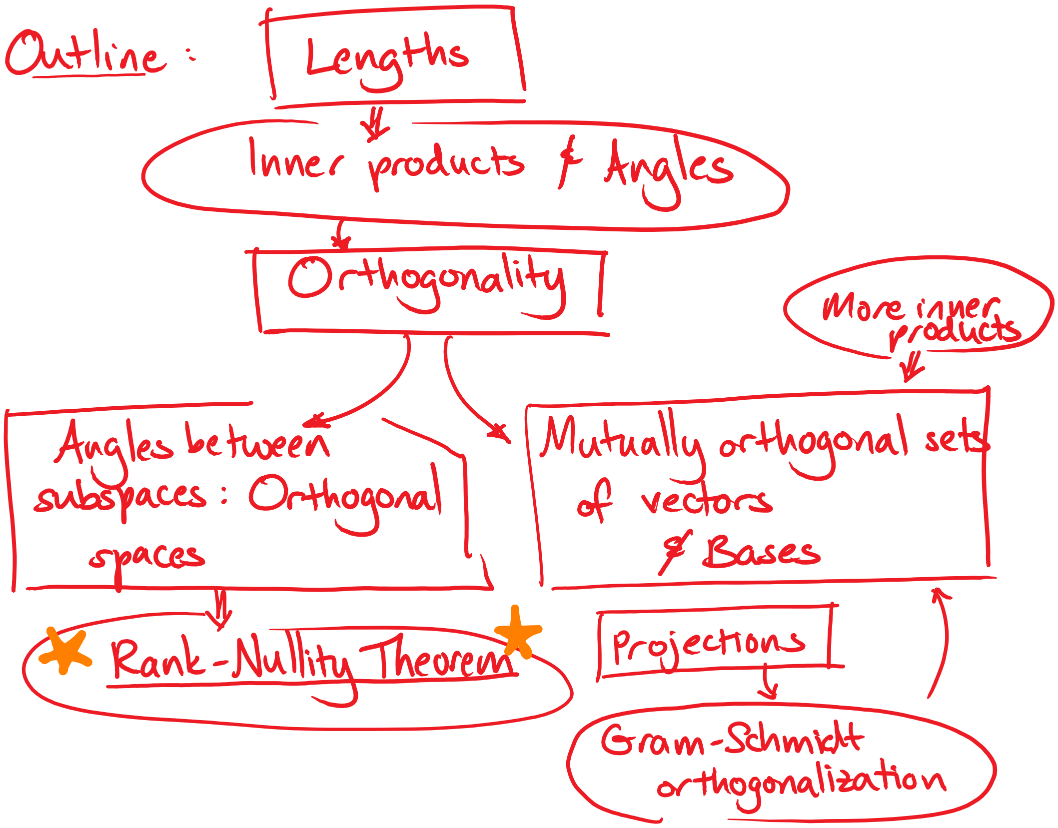
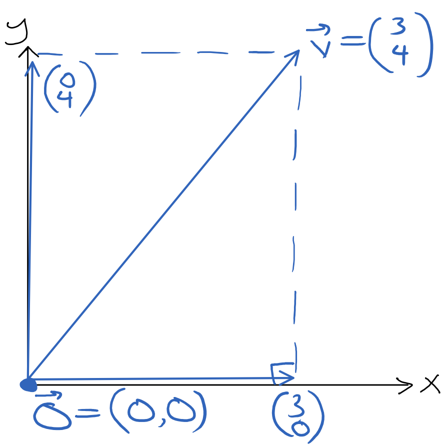
Lecture 8: Orthogonality

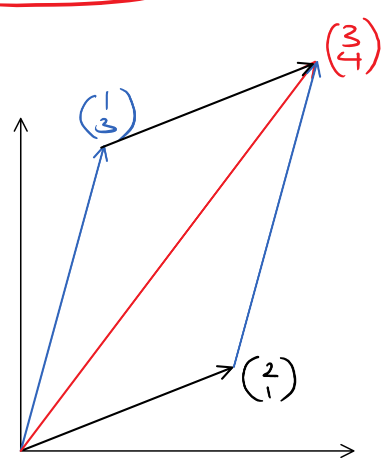
Reading:  2-3  5

Outline:



VECTORS





LENGTHS

- The **length** of a vector $\vec{v} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) is

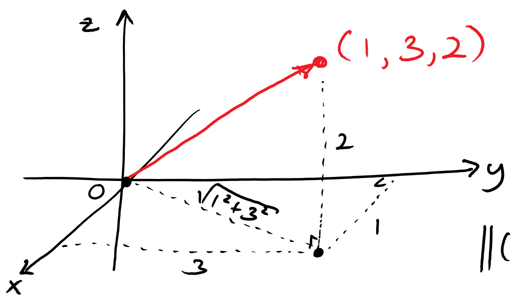
$$\|v\| = \sqrt{|v_1|^2 + \cdots + |v_n|^2}.$$

Example:

(*Note: We will define)

Example:

(*Note: We will define other lengths later.)



$$\|(1,3,2)\| = \sqrt{1^2+3^2+2^2} = \sqrt{14}$$

Observe: • For $v \in \mathbb{R}^n$, $\|v\|^2 = v^T v = (\text{---}\ v^T\ \text{---})\begin{pmatrix} | \\ v \\ | \end{pmatrix}$

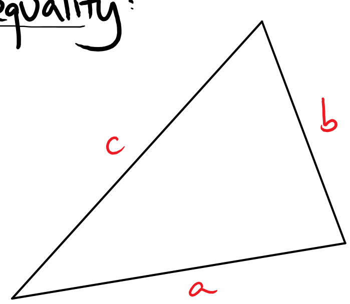
• Scaling: $\|\alpha \cdot \vec{v}\| = |\alpha| \cdot \|\vec{v}\|$

$\Rightarrow$ $\dfrac{v}{\|v\|}$ is a **unit vector**

• For $v \in \mathbb{C}^n$, $\|v\|^2 = \bar{v}^T v$,

eg. $\|(1,i)\|^2 = 1^2 + |i|^2 = 2$

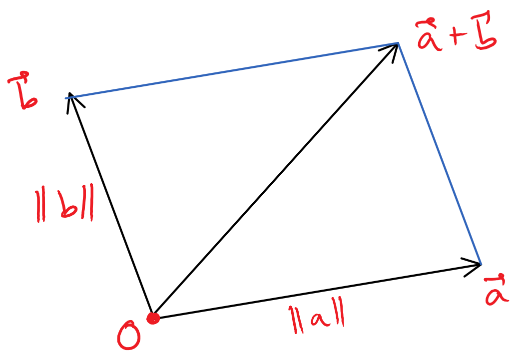
Triangle inequality:



$a+b \geq c$

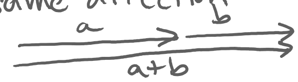
$a+c \geq b$

$b+c \geq a$



$$\boxed{\|\vec{a}+\vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|}$$

with equality if and only if a and b point in the same direction



Inner Products & Angles

• The inner product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is

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

The inner product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is

$$\boxed{\begin{aligned}\vec{u}\cdot\vec{v} &= u^T v \\ &= \sum_{i=1}^{n} u_i v_i\end{aligned}} \in \mathbb{R}$$

For two complex vectors $\vec{u}, \vec{v} \in \mathbb{C}^n$,

$$\vec{u}\cdot\vec{v} = \sum_i \bar{u}_i v_i \in \mathbb{C}.$$

(Other notation: dot product, scalar product, $\langle u|v\rangle$, (u,v).)

Observe:
* $\vec{u}\cdot\vec{v} = (\vec{v}\cdot\vec{u})^*$ ← complex conjugate
* $\vec{u}\cdot\vec{u} = \|\vec{u}\|^2$
* It is "bilinear":

$$\vec{u}\cdot(\vec{v}+\vec{w}) = \vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w} \qquad (\vec{u}+\vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w}+\vec{v}\cdot\vec{w}$$

$$\vec{u}\cdot(\alpha\vec{v}) = \alpha(\vec{u}\cdot\vec{v}) \qquad (\alpha\vec{u})\cdot\vec{v} = \alpha^*(\vec{u}\cdot\vec{v})$$

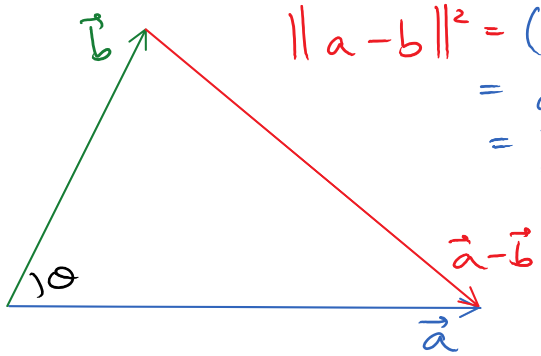
• The angle between $\vec{u}, \vec{v} \neq \vec{0}$ is

$$\boxed{\cos^{-1}\left(\frac{\vec{u}}{\|u\|}\cdot\frac{\vec{v}}{\|v\|}\right)} \in [0,\pi]$$

They are orthogonal if $\vec{u}\cdot\vec{v} = 0$ (angle $= \pi/2$)

AKA perpendicular, $\vec{u}\perp\vec{v}$

Observe: This definition agrees with what we know from geometry:

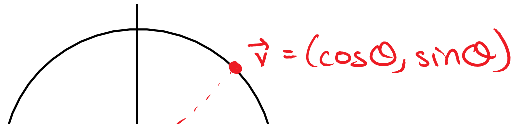


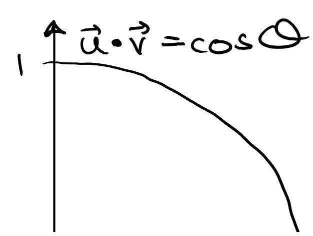
$$\begin{aligned}\|a-b\|^2 &= (a-b)\cdot(a-b) \\ &= a\cdot a + b\cdot b - 2a\cdot b \\ &= \|a\|^2 + \|b\|^2 - 2\|a\|\cdot\|b\|\cdot\cos\theta\end{aligned}$$

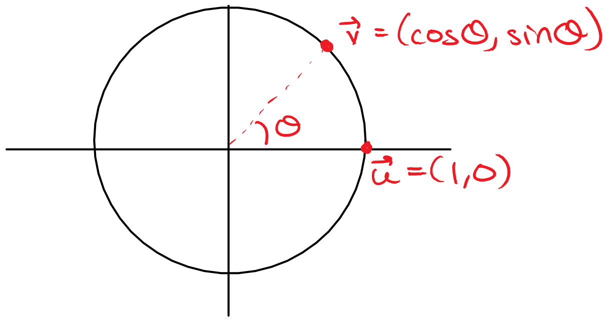
• Cauchy-Shwarz inequality:

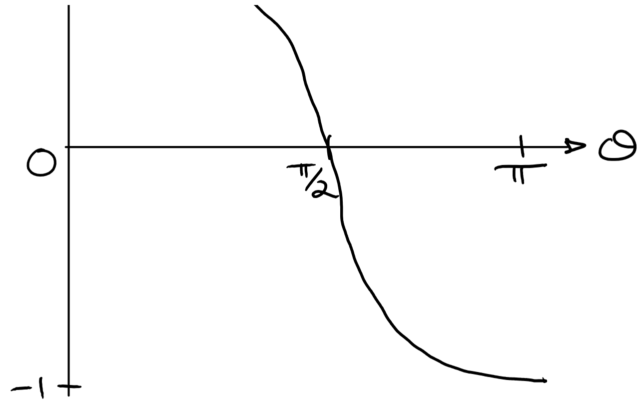
$$|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\times\|\vec{v}\|$$

Example:







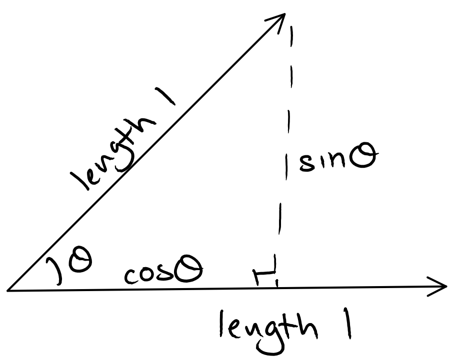


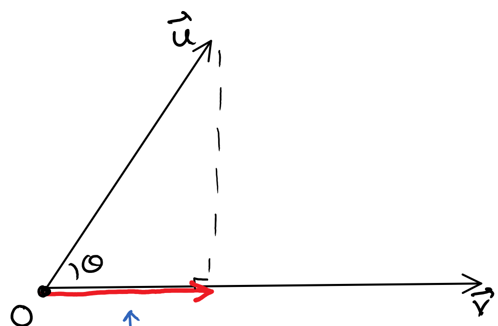
$\vec{u}\cdot\vec{v} = \|\vec{u}\|\cdot\|\vec{v}\|$ if they point in the same direction

$\vec{u}\cdot\vec{v} = -\|\vec{u}\|\cdot\|\vec{v}\|$ if they point in opposite directions

otherwise, $|\vec{u}\cdot\vec{v}| < \|\vec{u}\|\cdot\|\vec{v}\|$ strictly

Observe: The magnitude $|\vec{u}\cdot\vec{v}|$ measures the "overlap" between $\vec{u}$ and $\vec{v}$:





projection of $\vec{u}$ in the direction of $\vec{v}$

$$= \frac{\vec{v}}{\|\vec{v}\|}\cdot\|\vec{u}\|\cos\theta$$

$$= \boxed{\frac{(\vec{u}\cdot\vec{v})}{\|\vec{v}\|^2}\vec{v}}$$

– the projection is 0 if $\vec{u}\cdot\vec{v} = 0$, or $\theta = \frac{\pi}{2}$

$\Rightarrow$ If $\|\vec{v}\| = 1$,

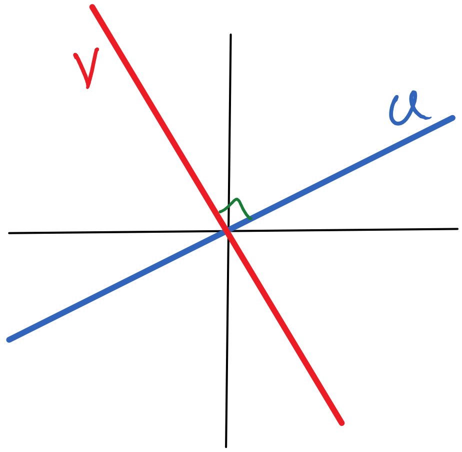
projection of $\vec{u}$ in the direction of $\vec{v}$ $= (\vec{u}\cdot\vec{v})\vec{v}$

Orthogonal vectors

$$\vec{u} \perp \vec{v} \iff \vec{u}\cdot\vec{v} = 0$$

$\vec{u}$ and $\vec{v}$ are orthogonal

ORTHOGONAL SUBSPACES



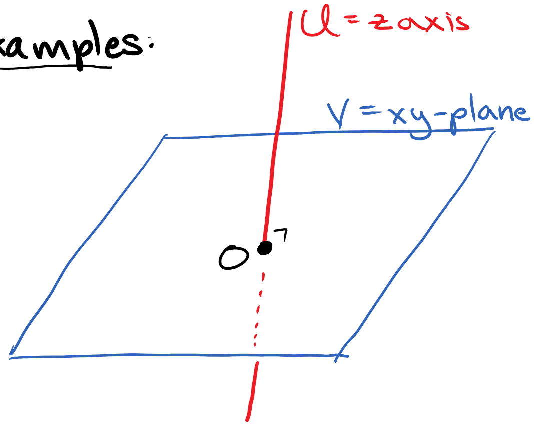
subspaces

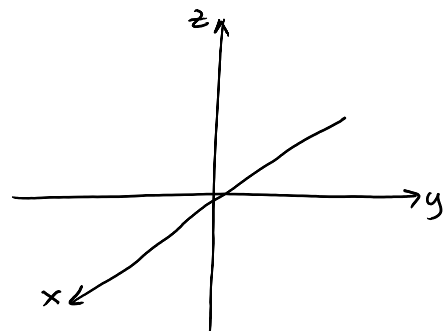
$\mathcal{U} \perp V$

$\Updownarrow$

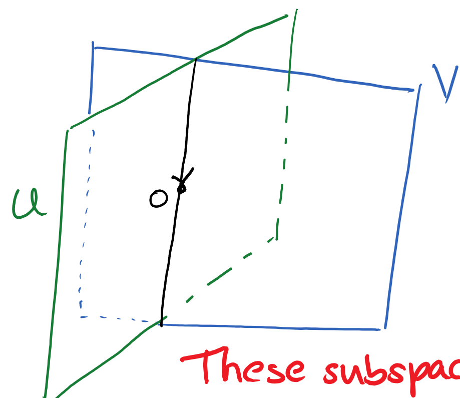
Every vector in $\mathcal{U}$ is orthogonal to every vector in V.

Examples:





coordinate axes are pairwise orthogonal

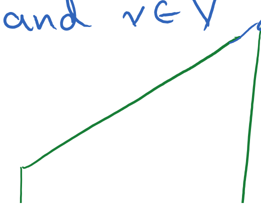


These subspaces are NOT orthogonal!!!

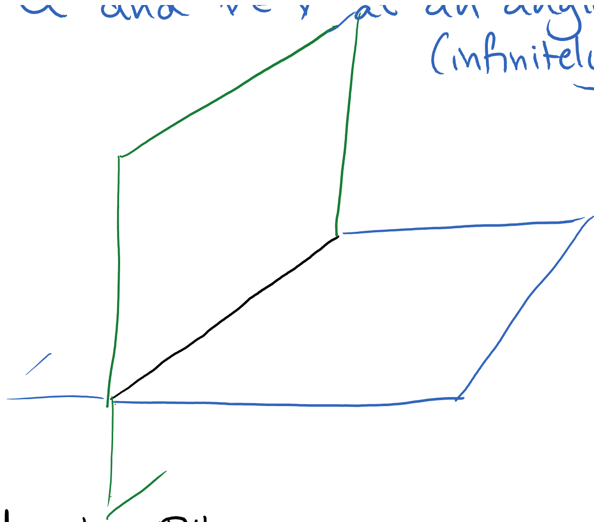
— because they intersect nontrivially

for $u \in \mathcal{U} \cap V$, u makes angle 0 with itself

In fact, for any angle θ, there are vectors $u \in \mathcal{U}$ and $v \in V$ at an angle θ to each other.
(infinitely many angles!)



(infinitely many angles!)

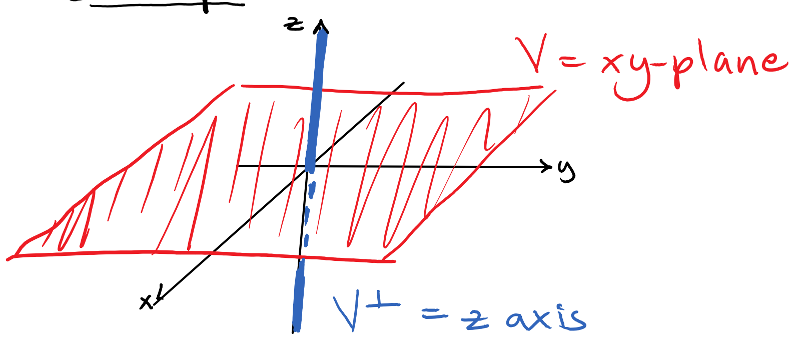


Example: In $\mathbb{R}^4$,

$$\mathrm{Span}\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right) \perp \mathrm{Span}\left(\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right)$$

> **Definition:** For a subspace V,
>
> $V^\perp = \{$ all vectors orthogonal to $V \}$
>
> "orthogonal complement of V", "V perp"

Example:



- $V \cap V^\perp = \{\vec{0}\}$
- $(V^\perp)^\perp = V$
- x and z axes are orthogonal, but **not** complements

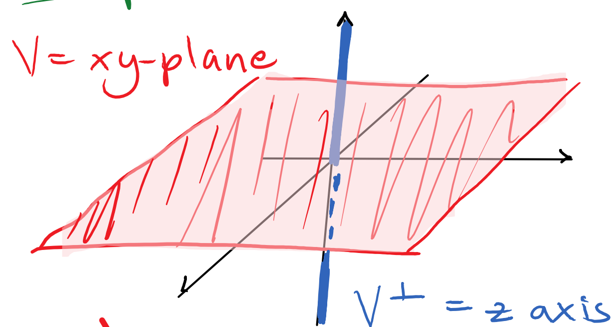
> **Theorem:** (Orthogonal complementary subspaces)
> Let V be a subspace of $\mathbb{R}^n$. Then,
> - $\dim(V^\perp) = n - \dim(V)$
> - $\mathrm{Span}(V \cup V^\perp) = \mathbb{R}^n$

- $\boxed{\mathrm{Span}(V \cup V^{\perp}) = \mathbb{R}^n}$

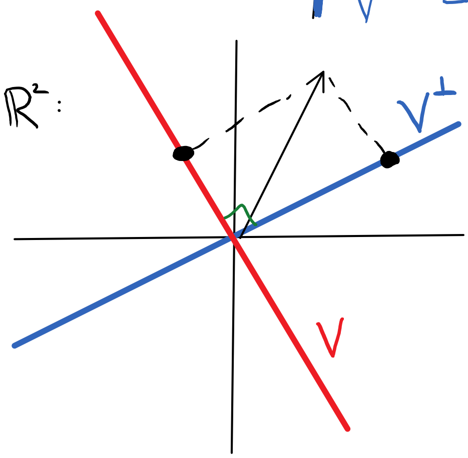
Example: $3 = 2 + 1$

This means that any vector can be split

$$(a,b,c) = \underset{\substack{\cap \\ V}}{(a,b,0)} + \underset{\substack{\cap \\ V^{\perp}}}{(0,0,c)}$$



Exercise: Prove that this decomposition is unique.



$2 = 1 + 1$

Proof of the theorem:

Let $\vec{v}_1, \dots, \vec{v}_d$ be a basis for V.

Claim: $\vec{u}$ is orthogonal to all vectors in V

$$\Updownarrow$$

$\vec{u}$ is orthogonal to each $\vec{v}_1, \dots, \vec{v}_d$.

[Why? If $\vec{u}\cdot\vec{v}_1 = \vec{u}\cdot\vec{v}_2 = \cdots = \vec{u}\cdot\vec{v}_d = 0$,

then $\vec{u}\cdot(c_1\vec{v}_1 + \cdots + c_d\vec{v}_d) = \sum_j c_j(\vec{u}\cdot\vec{v}_j)$
$= \vec{0}$

$\Rightarrow \vec{u}$ is orthogonal to anything in $\mathrm{Span}\{\vec{v}_1, \dots, \vec{v}_d\} = V$. ✓]

Thus,

$$V^{\perp} = \{\vec{u} \in \mathbb{R}^n \mid \vec{u} \text{ is orthogonal to all vectors in } V\}$$
$$= \{\vec{u} \in \mathbb{R}^n \mid \vec{u}\cdot\vec{v}_1 = 0, \dots, \vec{u}\cdot\vec{v}_d = 0\}$$

$$= \{\vec{u} \in \mathbb{R}^n \mid \underbrace{\vec{u}\cdot\vec{v}_1 = 0, \ \ldots, \ \vec{u}\cdot\vec{v}_d = 0}_{d \text{ equations}} \}$$

independent equations since
$\vec{v}_1, \ldots, \vec{v}_d$ are linearly indep. (basis)

$\Rightarrow$ $n-d$ free variables

$\Rightarrow$ $(n-d)$-dimensional solution set ✓

Next, to show $\text{Span}(V \cup V^\perp) = \mathbb{R}^n$, let $\vec{u}_1, \ldots, \vec{u}_{n-d}$ be a basis for $V^\perp$.

Consider the vectors

$$\vec{v}_1, \ldots, \vec{v}_d, \vec{u}_1, \ldots, \vec{u}_{n-d}.$$

Claim: These vectors are linearly independent.

Why?

No nontrivial combinations of the basis vectors for V alone, or for $V^\perp$ alone, can equal 0.

If a nontrivial combination of all the basis vectors

$$\begin{pmatrix}\text{comb. of basis}\\ \text{for } V\end{pmatrix} + \begin{pmatrix}\text{comb. of}\\ \text{basis for } V^\perp\end{pmatrix} = 0,$$

$$\Rightarrow \underbrace{\begin{pmatrix}\text{comb. of basis}\\ \text{for } V\end{pmatrix}}_{\in V} = \underbrace{-\begin{pmatrix}\text{comb. of}\\ \text{basis for } V^\perp\end{pmatrix}}_{\in V^\perp}$$

$\Rightarrow$ contradicts $V \cap V^\perp = \{\vec{0}\}$. ✓

But n linearly independent vectors in a space of dimension n must be a basis, spanning everything. □

How to find $V^\perp$:

Start with a basis for V.

$\Rightarrow V^\perp = \{\text{all vectors} \perp \text{to } V\}$

$= \{\text{all vectors} \perp \text{to all basis elements}\}$

(ie. $b_i \cdot v = 0$ for each basis elt. b_i)

$\{$all vectors $\perp$ to all basis elements$\}$

(i.e. $b_i \cdot v = 0$ for each basis elt. b_i)

Let $A = \begin{pmatrix} \text{———} \\ \text{———} \\ \text{rows a} \\ \text{basis for } V \\ \text{———} \end{pmatrix} \Rightarrow \begin{array}{l} R(A^T) = V \\ N(A) = V^\perp \end{array}$

Lecture 8: Rank-Nullity Theorem

RANK-NULLITY THEOREM

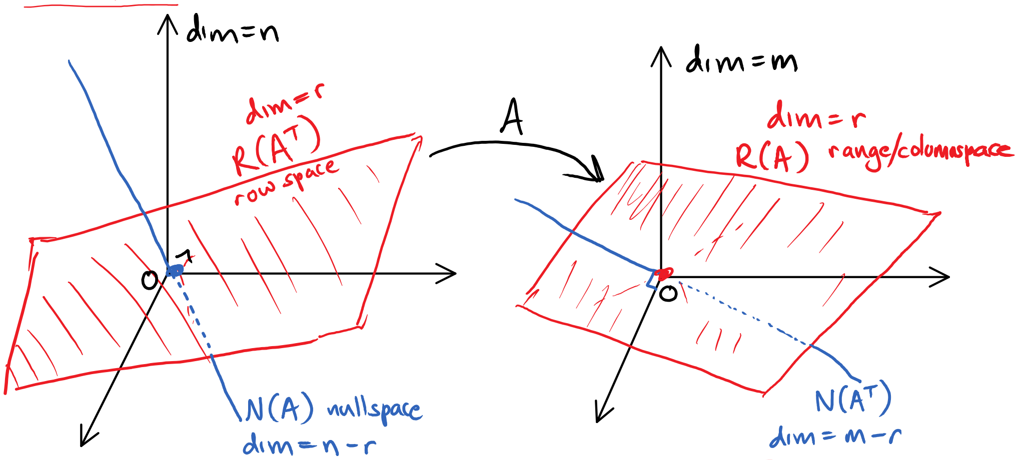
Definition: The rank of a matrix A is

$$\operatorname{rank}(A) = \dim R(A),$$

the dimension of the range.

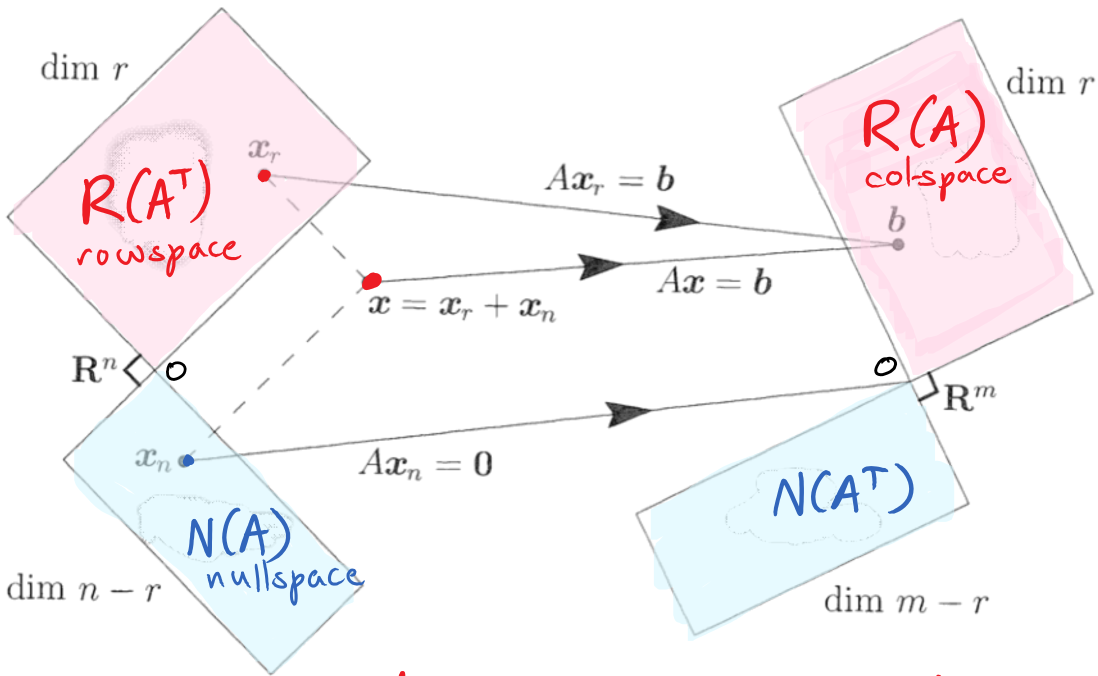
> **Rank-Nullity Theorem:**
> Let A be an $m \times n$ matrix. Then
> - $\dim R(A^T) = \dim R(A)$
> - $\dim N(A) = n - \dim R(A)$
> - $\dim N(A^T) = m - \dim R(A)$.

Intuition:



$$\boxed{\dim R(A^T) = \dim R(A)}$$

$$\boxed{\begin{aligned}&\dim N(A) + \dim R(A^T)\\ &= \text{total dimension } n\end{aligned}} \qquad \boxed{\begin{aligned}&\dim R(A) + \dim N(A^T)\\ &= \text{total dimension } m\end{aligned}}$$



$N(A) = R(A^T)^\perp$ $\qquad\qquad$ $N(A^T) = R(A)^\perp$

Corollary: A is invertible $\Longleftrightarrow$ $\underbrace{m = n = \text{rank}(A)}_{\text{"full rank"}}$.

Example:

Problem: Give a basis for $\underbrace{\{\vec{x} \in \mathbb{R}^n \mid x_1 + x_2 + \cdots + x_n = 0\}}_{\text{"}V\text{"}}$.

Answer:

$V = N(\underbrace{(1 \ 1 \ 1 \ \cdots \ 1)}_{\text{"}A\text{"}})$

$\text{rank}(A) = 1$

$\Rightarrow \dim N(A) = n - \text{rank}(A)$ $\qquad$ [Rank-Nullity]

$= n - 1$

$\Rightarrow$ Any linearly independent set of $n-1$ elements of V is a basis for V.

For example,

$\{\vec{e}_1 - \vec{e}_n, \vec{e}_2 - \vec{e}_n, \vec{e}_3 - \vec{e}_n, \ldots, \vec{e}_{n-1} - \vec{e}_n\}$ is a basis. ✓

Proof of the Rank-Nullity Theorem:

1) Claim: For any $m \times n$ matrix A,

$\boxed{N(A) = R(A^T)^\perp}$ $\quad$ $(\rightarrow \dim N(A) = n - \dim R(A^T))$

1) Claim: For any $m \times n$ matrix A,

$$\boxed{\begin{aligned} N(A) &= R(A^T)^\perp \\ N(A^T) &= R(A)^\perp \end{aligned}} \qquad \left(\Rightarrow \begin{aligned} \dim N(A) &= n - \dim R(A^T) \\ \dim N(A^T) &= m - \dim R(A) \end{aligned} \right)$$

Proof: Let $\mathcal{U} = R(A^T)$.

$$A = \begin{pmatrix} \text{———} \\ \text{———} \\ \text{rows span } \mathcal{U} \\ \text{———} \end{pmatrix}$$

For $\vec{x} \in \mathbb{R}^n$, $\quad A\vec{x} = \begin{pmatrix} (\text{row } 1) \cdot x \\ \vdots \\ (\text{row } m) \cdot x \end{pmatrix} \in \mathbb{R}^m$.

Hence $A\vec{x} = \vec{0} \iff$ for all j, $(\vec{\text{row}_j}) \cdot \vec{x} = 0$

$\iff \vec{x} \perp \mathcal{U} = \text{span}\{\text{row } 1, \ldots, \text{row } m\}$.

Thus $\underbrace{N(A)}_{\text{nullspace}} = \underbrace{R(A^T)}_{\text{rowspace}}{}^\perp$. ✓

To see that $\underbrace{N(A^T)}_{\text{left nullspace}} = \underbrace{R(A)}_{\text{columnspace}}{}^\perp$, repeat the argument with A^T. ☑

2) Claim: $\dim R(A) = \dim R(A^T)$.

Proof #1 (works over $\mathbb{R}$ or $\mathbb{C}$):

Let $\vec{r}_1, \ldots, \vec{r}_d$ be a basis for the rowspace $R(A^T)$.

We claim that $A\vec{r}_1, \ldots, A\vec{r}_d$ are linearly independent.

Why?

If $c_1 A\vec{r}_1 + \cdots + c_d A\vec{r}_d = \vec{0}$

$\Rightarrow A\left(\sum_j c_j \vec{r}_j\right) = \vec{0}$

$\Rightarrow \sum_j c_j \vec{r}_j \in N(A)$

But this is also in $R(A^T)$.

To be in $N(A) \cap R(A^T)$, it must be $\vec{0}$,

implying (by lin. indep.) that all $c_j = 0$.

Hence $A\vec{r}_1, \ldots, A\vec{r}_d$ are independent, too. ✓

Therefore $\dim R(A) \geq \dim R(A^T)$.

Repeating the argument with A^T gives the opposite inequality, $\dim R(A^T) \geq \dim R(A)$. Hence they are equal. □

Proof #2 (works in general):

inequality, $\dim R(A^T) \geq \dim R(A)$. Hence they are equal. $\square$

Proof #2 (works in general):

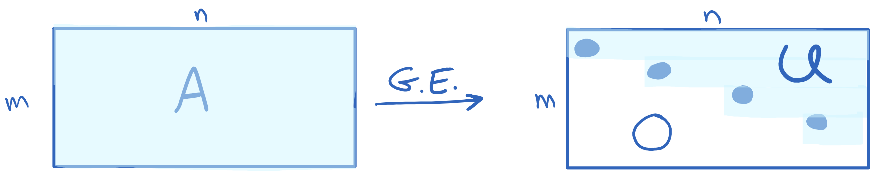
The above proof doesn't work over finite fields.

For example, if $A = \begin{pmatrix} 1 & 1 \end{pmatrix}$ over $\mathbb{F}_2$,

$$N(A) = R(A^T) = \left\{ \vec{0}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

However $\dim R(A) = \dim R(A^T)$ is still true.

Proof using Gaussian elimination:



Let $d = \#$ of pivot columns.

Since those rows span $R(A^T) = R(U^T)$, $d = \dim R(A^T)$.

Since the corresponding columns in A span $R(A)$ (see Lecture: Subspaces of a matrix), $d = \dim R(A)$. ✓

Also, # free variables $= n - d = \dim N(A)$. ✓ $\square$

Example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 4 \\ 2 & -1 & 1 \end{pmatrix} \xrightarrow{-4,\ -2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -8 & -8 \\ 0 & -5 & -5 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$R(A) = \text{Span}\{(1,4,2),(2,0,-1)\} \perp N(A^T) = \text{Span}((4,-5,8))$

$R(A^T) = \text{Span}\{(1,2,3),(0,1,1)\} \perp N(A) = \text{Span}((1,1,-1))$

$\text{rank}(A) = \dim R(A) = \dim R(A^T) = 2$

$\dim N(A) + \dim R(A^T) = 3$

$\dim N(A^T) + \dim R(A) = 3$

Example:

$$A = \begin{array}{c} \begin{array}{ccccccc} & & & & & & n \end{array} \\ \begin{pmatrix} -2 & 1 & & & & & 1 \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & 1 & \ddots & & & \\ & & & & & -2 & 1 \\ \cdot & & & & & \cdot & \cdot \end{pmatrix} \end{array}$$

discretized 2^{nd} derivative on the circle

$$n \begin{pmatrix} \ddots & & \\ & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

Recall: $N(A) = \text{Span}((1, 1, \dots, 1))$

$A = A^T \Rightarrow N(A^T) = N(A)$

$\Rightarrow R(A) = N(A^T)^\perp$ $n-1$ dimensional

$= \{\text{all } \vec{b} \in \mathbb{R}^n \mid \vec{b} \cdot (1, 1, \dots, 1) = 0\}$

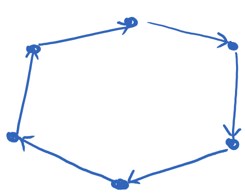
$= \{\text{all } \vec{b} \in \mathbb{R}^n \mid b_1 + b_2 + \dots + b_n = 0\}$

Hence, $\boxed{Ax = b \text{ has a solution} \iff b_1 + \dots + b_n = 0}$

This is **much** easier than using Gaussian elimination to find a basis for $R(A)$. (Essentially because $N(A^T)$ is low-dimensional.)

This example is actually very important; similar matrices arise frequently in applications.

Note that $A = -E_G E_G^T$, where G is the cycle graph and E_G is the incidence matrix.



We proved $N(E_G^T) = \text{Span}((1, 1, \dots, 1))$, which gives an alternate proof that $N(A) = \text{Span}((1, \dots, 1))$.

Exercise:

Give the *dimensions* of the four fundamental subspaces (columnspace, rowspace, nullspace, left nullspace) of the matrix

$$A = \begin{pmatrix} 500 & 5 & 6 & 3 & 0 \\ 200 & 2 & 8 & 4 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 2 & 3 & 4 & 4 & 5 \end{pmatrix}$$

Space	Dimension
$R(A)$	
$R(A^T)$	
$N(A)$	
$N(A^T)$	

Answer:

Answer:

Notice that A is 6×5

$$\Rightarrow \dim N(A) = 5 - \text{Rank}(A)$$
$$\dim N(A^T) = 6 - \text{Rank}(A).$$

To find Rank(A), notice that

row and column operations do not change the rank!

(why not?)

$$\begin{pmatrix} 500 & 5 & 6 & 3 & 0 \\ 200 & 2 & 8 & 4 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 2 & 3 & 4 & 4 & 5 \end{pmatrix} \to \begin{pmatrix} 500 & 5 & 6 & 3 & 0 \\ 200 & 2 & 8 & 4 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

$$\to \begin{pmatrix} 500 & 5 & 0 & 3 & 0 \\ 200 & 2 & 0 & 4 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \to \begin{pmatrix} 500 & 5 & 0 & 3 & 0 \\ 200 & 2 & 0 & 4 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

$$\to \begin{pmatrix} 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 500 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

$$\Rightarrow \text{Rank}(A) = \dim R(A) = \dim R(A^T) = 4$$
$$\dim N(A) = 5 - \text{Rank}(A) = 1$$
$$\dim N(A^T) = 6 - \text{Rank}(A) = 2.$$

Todo: Understand better how $R(A^T)$ is mapped onto $R(A)$.

<u>Todo</u>: Understand better how $R(A^T)$ is mapped onto $R(A)$.
(Singular-value decomposition)