MAT223

Linear Algebra

Module 1

Sets

- unordered collection of distinct objects
- $\{1,2,3\}$ → "the set containing the elements 1, 2, and 3"
- elements: things in a set
    - $3 \in \{1,2,3\}$
    - $4 \notin \{1,2,3\}$
- can contain mixture of objects, including other sets
    - $\{1, 2, a, \{-70, \infty\}, x\}$
- $\varnothing$ or $\{\}$: empty set

Operation on Sets

- if set A contains all the elements of set B,
    - B is a subset of A → $B \subseteq A$
    - A is a superset of B

- $\{1,2,3\} \subseteq \{1,2,3,4\}$
- $\{1,2,3\} \subseteq \{1,2,3\}$
- sets A and B are equal, $A = B$, if $A \subseteq B$ and $B \subseteq A$

Set-builder Notation

- if X is a set, we can define a subset:
    - $Y = \{a \in X : \text{some rules involving } a\}$
    - "Y is a set of a in X such that some rules involving a is true"
- Union: $X \cup Y = \{a : a \in X \text{ or } a \in Y\}$
- Intersection: $X \cap Y = \{a : a \in X \text{ and } a \in Y\}$
    - i.e. $A = \{1,2,3\}$ and $B = \{-1,0,1,2\}$
        - $A \cap B = \{1,2\}$
        - $A \cup B = \{-1,0,1,2,3\}$
- sets and unions are associative, it does not matter where to place parentheses to an expression involving just unions or just intersections

- $(A \cup B) \cup C = A \cup (B \cup C)$
- $(A \cup B) \cap C \neq A \cup (B \cap C)$

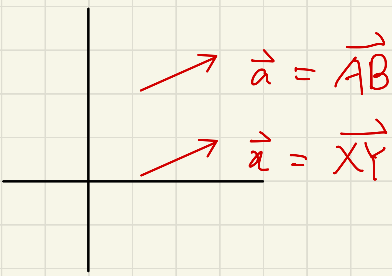
- common sets:
  - $\emptyset = \{\}$, the empty set
  - $\mathbb{N} = \{0, 1, 2, 3, \dots\} = \{\text{natural numbers}\}$
  - $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\} = \{\text{integers}\}$
  - $\mathbb{Q} = \{\text{rational numbers}\}$
  - $\mathbb{R} = \{\text{real numbers}\}$
  - $\mathbb{R}^n = \{\text{vectors in n-dimensional Euclidean space}\}$

Vectors and Scalars

- a <u>scalar</u> number models a relationship between quantities
  - i.e. 6 times as much flour as sugar
- a <u>vector</u> models a relationship between points
  - i.e. 2km east and 4km north

Vector Notation

- represented by arrow over letter
    - i.e. $\vec{a}$
- $\|\vec{a}\|$ is the magnitude of vector $\vec{a}$
    - also called the <u>norm</u> or <u>length</u>
- directed line segments always start somewhere, but a vector models a displacement and has no sense of "origin"



- $\vec{a} = \vec{x}$ despite the fact that $A \neq X$

- a vector is not the same as a line segment and a vector itself has no origin

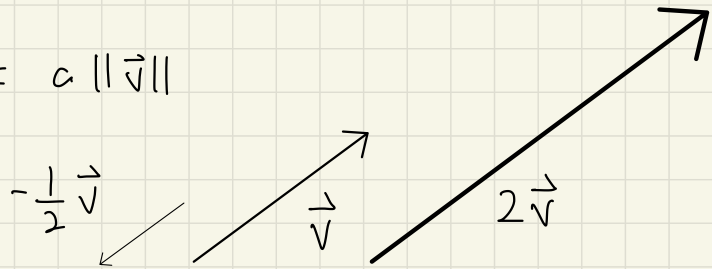
<u>Vectors and Points</u>

- a point specifies an absolute position
- a vector specifies a displacement
- given a point P, one associates P with the vector $\vec{p} = \overrightarrow{OP}$, where O is the origin

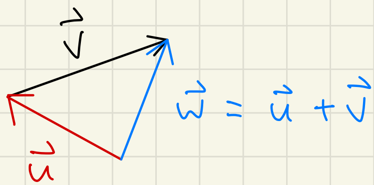
- points and vectors can be treated interchangeably

Vector Arithmetic

- Scalar multiplication: for a vector $\vec{v}$ and a scalar $a > 0$, the vector $\vec{w} = a\vec{v}$ is the vector pointing to the same direction as $\vec{v}$ but with length scaled by a
    - $\|\vec{w}\| = a\|\vec{v}\|$



- for two vectors $\vec{u}$ and $\vec{v}$, the sum $\vec{w} = \vec{u} + \vec{v}$ represents the displacement vector



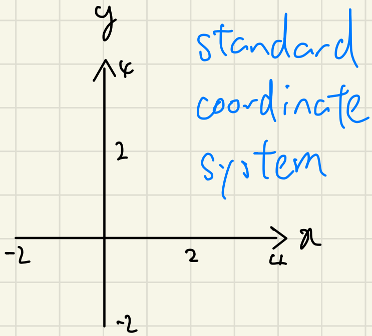
- you add vectors tip to tail and scale vectors by changing their length

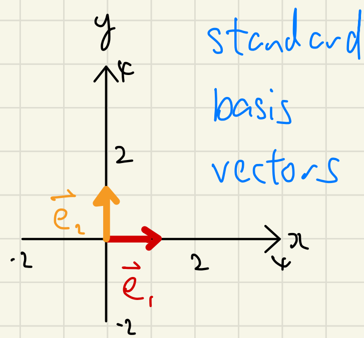
- zero vector : $\vec{0}$, vector with no magnitude
    - does not have a well-defined direction

    - point in every direction / no direction
- laws (for vectors $\vec{u}$, $\vec{v}$, $\vec{w}$; scalars α, β)
    - associativity
        - $(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w})$
        - $(\alpha\beta)\vec{v} = \alpha(\beta\vec{v})$
    - commutativity
        - $\vec{u}+\vec{v} = \vec{v}+\vec{u}$
    - distributivity
        - $\alpha(\vec{u}+\vec{v}) = \vec{u} + \alpha\vec{v}$
        - $(\alpha+\beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$
- linear combination
    - a <u>linear combination</u> of the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is a vector

      $\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_n\vec{v}_n$
    - the scalars $a_1, a_2, \ldots, a_n$ are <u>coefficients</u>

<u>Coordinates and the Standard Basis</u>





- the vector $\vec{e}_1$ always point one unit in the direction of $+x$ axis
- $\vec{e}_2$ always point 1 unit in the direction of $+y$ axis
- using the standard basis, every point (or vector) in the plane can be represented as a linear combination
    - if point $P = (\alpha, \beta)$, then $\vec{OP} = \alpha \vec{e}_1 + \beta \vec{e}_2$
- every vector in $\mathbb{R}^2$ can be written uniquely as a linear combination of the standard basis vectors
    - $\mathbb{R}^2$: the standard flat Euclidean plane
- for a vector $\vec{w} = \alpha \vec{e}_1 + \beta \vec{e}_2$, the pair (α, β)

is the standard coordinates of $\vec{w}$

- notations
  - (α, β)
  - $\langle \alpha, \beta \rangle$
  - $[\alpha \quad \beta]$
  - $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$
- i.e. $\vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \rightarrow$ " $\vec{v} = \alpha \vec{e}_1 + \beta \vec{e}_2$ "

Solving Problems with Coordinates

- if $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$
- then $\vec{u} = \vec{v} \iff \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \iff a = x$ and $b = y$
- further $\vec{u} + \vec{v} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a+x \\ b+y \end{bmatrix}$

  and $t\vec{v} = t \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ta \\ tb \end{bmatrix}$ for any scalar t

Higher Dimensions

- $\mathbb{R}^3$ has standard basis vectors $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$, which each point 1 unit along the x, y, and z axes, respectively

- notations for 3D standard basis:
  - $\hat{x}$ $\hat{y}$ $\hat{z}$
  - $\hat{i}$ $\hat{j}$ $\hat{k}$
  - i j k
  - $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$
- $\mathbb{R}^n$: n-dimensional Euclidean space
  - standard basis: $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$
- every vector in $\mathbb{R}^n$ can be written uniquely as a linear combination of the standard basis

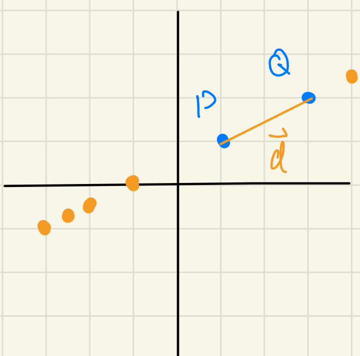
Module 2

Lines

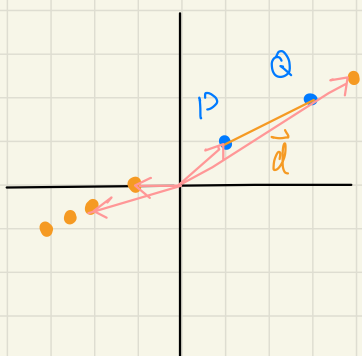
- a line ℓ passes through points P, Q, $P, Q \in \mathbb{R}^3$
  - let $\vec{d} = \overrightarrow{PQ}$, the set of points (or vectors) $\vec{x}$ can be expressed as

    $\vec{x} = t\vec{d} + P$ for $t \in \mathbb{R}$

    - set of all points that starts at P and displacing some multiple of $\vec{d}$



the line as the set of points



the line as the set of vectors rooted at the origin

- $\ell = \{\vec{x} : \vec{x} = t\vec{d} + P \text{ for some } t \in \mathbb{R}\}$
    - not "for all $t \in \mathbb{R}$"
- vector form of a line:
    - Let ℓ be a line and $\vec{d}$ and $\vec{p}$ be vectors. If $\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}$, the vector equation $\vec{x} = t\vec{d} + \vec{p}$ is ℓ expressed in vector form
    - $\vec{d}$ is the direction vector for ℓ
- $\begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ is line passing through $\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ with $\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ as a direction vector
    - t : parameter variable
- vector form is a shorthand for a set, do not add extra words!
    - $\vec{x} = t\vec{d} + \vec{p}$ works
    - $\vec{x} = t\vec{d} + \vec{p}$ where $t \in \mathbb{R}$ works
    - $\vec{x} = t\vec{d} + \vec{p}$ for some $t \in \mathbb{R}$ doesn't

- $\vec{x} = t\vec{d} + \vec{p}$ for all $t \in \mathbb{R}$ doesn't
- $\ell = t\vec{d} + \vec{p}$ doesn't work
- the ones that don't work do not specify ℓ in vector form

- downside of vector form: not unique
  - the same line can be modelled by different equations
  - the line being the same does not require the vectors to be the same
- determine if ℓ_1 and ℓ_2,

$$\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{x} = t\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

are the same line

  - if $\vec{x} \in \ell_1$, then $\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ for some $t \in \mathbb{R}$
  - if $\vec{x} \in \ell_2$, then $\vec{x} = s\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ for some $s \in \mathbb{R}$

- thus if $t\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}2\\1\end{bmatrix}=s\begin{bmatrix}2\\2\end{bmatrix}+\begin{bmatrix}4\\3\end{bmatrix}$

always has a solution, $\ell_1=\ell_2$

$\rightarrow$ equation becomes $(s+1-\frac{t}{2})\begin{bmatrix}2\\2\end{bmatrix}$

- since for every s and for every t the above equation has a solution, $\ell_1=\ell_2$.

Vector Form in Higher Dimensions

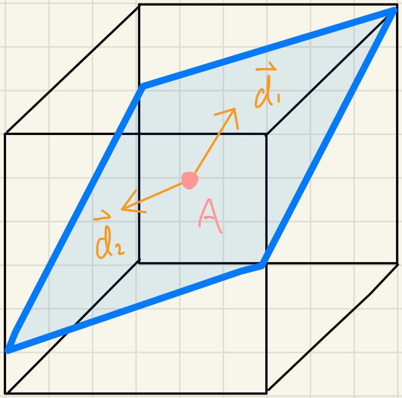
- non-parallel lines do not have to intersect in $\mathbb{R}^3$ and above
- if the lines are neither parallel nor intersecting, they are skew
- determine if lines $\vec{x}=t(1,3,-2)+(1,2,1)$ and $\vec{x}=t(0,2,3)+(0,3,9)$ intersect
  - $t(1,3,-2)+(1,2,1)=s(0,2,3)+(0,3,9)$

$$\begin{cases} t+1=0 \\ 3t+2=2s+3 \\ -2t+1=3s+9 \end{cases}$$

- above system is overdetermined
- $t=1$ and $s=2$ is consistent through all equations
- plug t into first equation or s into second equation to obtain intersection point

Planes

- to define a plane, we need 3 points that are not on the same line
- Let $A, B, C \in \mathbb{R}^3$ be three non-collinear points and let P be the plane that passes through A, B, C
- planes have direction vectors
  - for P, $\vec{d}_1 = \overrightarrow{AB}$ and $\vec{d}_2 = \overrightarrow{AC}$ are direction vectors
- plane P in vector form:
  - $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\vec{d}_1 + s\vec{d}_2 + A$



- plane P in vector form:
  - $\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p}$

    for some vectors $\vec{d}_1$ and $\vec{d}_2$ and point $\vec{p}$
  - $P = \{\vec{x} : \vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p}$ for some $t, s \in \mathbb{R}\}$
  - $\vec{d}_1, \vec{d}_2$: direction vectors for P
- example: describe the plane $P \subseteq \mathbb{R}^3$ w/ equation $z = 2x + y + 3$ in vector form
  - find 3 points by guess and check:

$$A = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$$

-

$$\vec{d}_1 = B - A = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \vec{d}_2 = C - A = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

- since the vectors are not parallel,

$$x = t\vec{d}_1 + s\vec{d}_2 + A = t\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

- example: find the intersection between P_1 and P_2,

$$\vec{x} = t\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ and } \vec{x} = t\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} + s\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

- $$a\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = c\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} + d\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

$$\rightarrow \begin{cases} a - b + c - d = -1 \\ a \quad\quad\quad - 2d = -2 \\ \quad b - 2c - d = 0 \end{cases} \quad \text{(underdetermined)}$$

- let a be r: (for every $r \in \mathbb{R}$)

$b = \frac{r}{2} - 1$; $c = -1$; $d = \frac{r}{2} + 1$

- substitute parameters into original equation:

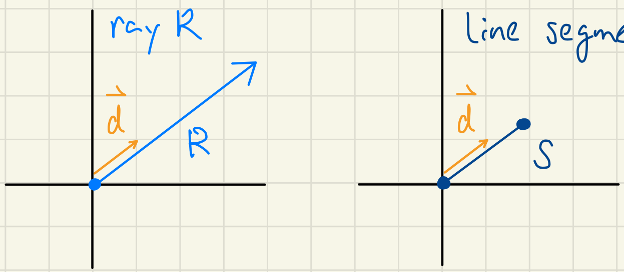
$$\vec{x} = r\begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$ which is $P_1 \cap P_2$ in vector form

Restricted Linear Combinations

- $\vec{x} = t\vec{d} + \vec{p}$ for line ℓ means

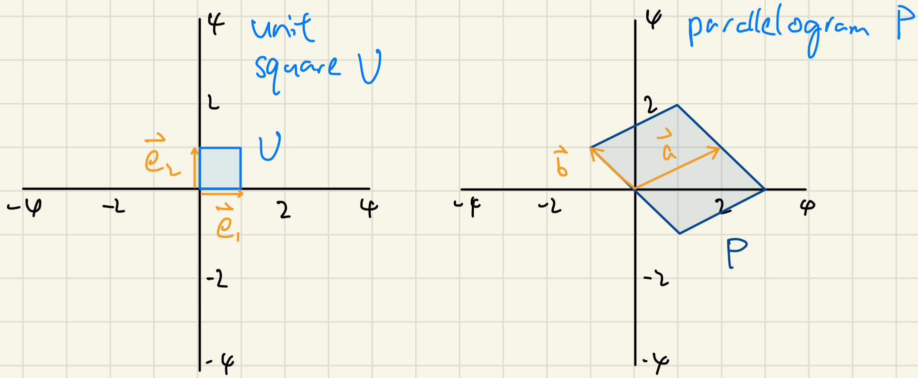
  $\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p}$ for some $t \in \mathbb{R}\}$

- add restrictions on t to only describe a part of ℓ
  - $R = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p}$ for some $t \geq 0\}$
  - $S = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p}$ for some $t \in [0,2]\}$



- add restrictions to the vector form of a plane to make polygons
  - let $\vec{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
  - $U = \{\vec{x} : \vec{x} = t\vec{e}_1 + s\vec{e}_2$ for some $t, s \in [0,1]\}$

- $P = \{ \vec{x} : \vec{x} = t\vec{a} + s\vec{b}$ for some $t \in [0,1]$ and $s \in [-1,1] \}$



- let $\vec{w} = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n$
  - the vector $\vec{w}$ is a non-negative linear combination of $\vec{v}_1 + \vec{v}_2 + \dots + \vec{v}_n$ if $a_1, a_2, \dots, a_n \geq 0$
  - the vector $\vec{w}$ is a convex linear combination of $\vec{v}_1 + \vec{v}_2 + \dots + \vec{v}_n$ if $a_1, a_2, \dots, a_n \geq 0$ and $a_1 + a_2 + \dots + a_n = 1$
- non-negative linear combinations: only displacing "forward"
- convex linear combinations: weighted average of vectors

- the average of $\vec{v}_1, \dots, \vec{v}_n$ would be the convex linear combination w/ coefficients

  $a_i = \frac{1}{n}$

- a convex linear combination of 2 vectors gives a point on the line segment connecting them

- example: let $\vec{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and define

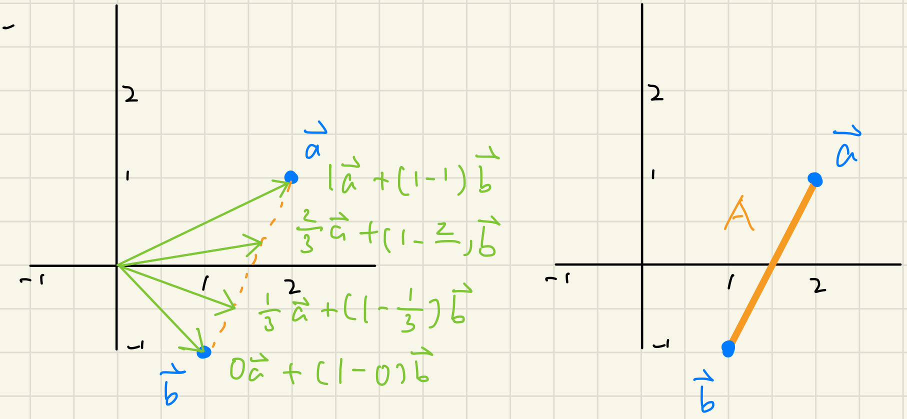
  $A = \{ \vec{x} : \vec{x}$ is a convex linear combination of $\vec{a}$ and $\vec{b} \}$

  $= \{ \vec{x} : \vec{x} = \alpha \vec{a} + (1-\alpha)\vec{b}$ for some $\alpha \in [0,1] \}$

  - $\vec{x} = \alpha\vec{a} - \alpha\vec{b} + \vec{b} = \alpha(\vec{a} - \vec{b}) + \vec{b}$

    - vector form of line that passes through $\vec{b}$ w/ direction $\vec{a} - \vec{b}$, with restriction $\alpha \in [0,1]$

    - A is just part of the line that connects $\vec{a}$ and $\vec{b}$



- since A is an infinite collection of vectors, it's better to draw it in dots rather than lines from the origin

Module 3

Span

- The span of a set of vectors V is the set of all linear combinations of vectors in V
  - i.e. span $V = \{\vec{v} : \vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n v_n$ for some $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in V$ and scalars $a_1, a_2, \dots, a_n\}$
  - span $\{\} = \{\vec{0}\}$
- span $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\} = \mathbb{R}^2$

- example: let $\vec{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. find span $\{\vec{u}, \vec{v}\}$
  - span $\{\vec{u}, \vec{v}\} = \{\vec{x} : \vec{x} = \alpha\vec{u} + \beta\vec{v}$ for some $\alpha, \beta \in \mathbb{R}\}$
  - $\begin{bmatrix} x \\ y \end{bmatrix} = \alpha \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

- $\begin{cases} x = -\alpha + \beta \\ y = 2\alpha - 2\beta \end{cases}$

- $y = -2x$

- therefore, if $\begin{bmatrix} x \\ y \end{bmatrix}$ make the above system consistent, we must have

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ -2t \end{bmatrix} = t\vec{v} \quad \text{for some } t$$

- $\text{Span}\{\vec{u}, \vec{v}\} = \{\vec{x} : \vec{x} = t\vec{v} \text{ for some } t\}$

$= \text{span}\{\vec{v}\}$

  - which is the line through origin w/ direction $\vec{v}$

- example: let $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and

$c = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, show that $\mathbb{R}^3 = \text{span}\{\vec{a}, \vec{b}, \vec{c}\}$

- $$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

- $$\begin{cases} x = a_1 + a_3 \\ y = 2a_1 + a_2 + a_3 \\ z = a_1 + 2a_3 \end{cases}$$

- $a_1 = 2x - z$, $a_2 = -3x + y + z$, $a_3 = -x + z$
- the above is always a solution (no matter the values of $x, y,$ and z)
- $\boxed{\therefore \text{span}\{\vec{a}, \vec{b}, \vec{c}\} = \mathbb{R}^3}$

Representing Lines & Planes as Spans

- if line ℓ in vector form is

  $\vec{x} = t\vec{d} + \vec{0}$

  the line ℓ passes through the origin

  $\ell = \{\vec{x} : \vec{x} = t\vec{d} \text{ for some } t \in \mathbb{R}\}$

  $= \text{span}\{\vec{d}\}$

- if P is a plane given in vector form by

  $\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{0}$

then

$P = \{\vec{x} : \vec{x} = t\vec{d}_1 + s\vec{d}_2 \text{ for some } t, s \in \mathbb{R}\}$

$= \text{span}\{\vec{d}_1, \vec{d}_2\}$

- if the $\vec{p}$ in vector form is $\vec{0}$, then that vector form defines a span
    - every line / plane through the origin can be written as a span
- if $X = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$, we know $\vec{0} = 0\vec{v}_1 + \dots + 0\vec{v}_n \in X$
    - every span passes through the origin
- spans exactly describe points, lines, planes, and volumes through the origin
- example: the line $\ell_1 \subseteq \mathbb{R}^2$ is described by the equation $x + 2y = 0$ and the line $\ell_2 \subseteq \mathbb{R}^2$ is described by the equation $4x - 2y = 6$. If possible, describe ℓ_1 and ℓ_2 using spans
    - ℓ_1 can be expressed in vector form by

$$x = t\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \vec{0}$$

and so $l_1 = \text{span}\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}$

- l_2 does not pass through $\vec{0}$, and so l_2 cannot be written as a span

Set Addition

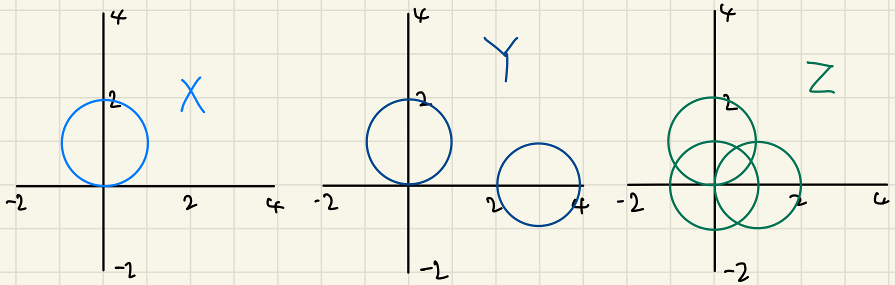
- if A and B are sets of vectors, then the set sum of A and B, denoted $A+B$, is

$$A+B = \{\vec{x} : \vec{x} = \vec{a} + \vec{b} \text{ for some } \vec{a} \in A \text{ and } \vec{b} \in B\}$$

- set sums are different from regular sums
- Let $C = \{\vec{x} \in \mathbb{R}^2 : \|\vec{x}\| = 1\}$, unit circle centered at the origin
  - $X = C + \{\vec{e}_2\}$
  - $Y = C + \{3\vec{e}_1, \vec{e}_2\}$
  - $Z = C + \{\vec{0}, \vec{e}_1, \vec{e}_2\}$
- rewriting,

- $X = \{\vec{x} + \vec{e}_2 : \|\vec{x}\| = 1\}$
- $Y = \{\vec{x} + \vec{v} : \|\vec{x}\| = 1 \text{ and } \vec{v} = 3\vec{e}_1 \text{ or } \vec{v} = \vec{e}_2\}$

  $= (C + \{3\vec{e}_1\}) \cup (C + \{\vec{e}_2\})$
  - Y is the union of two translated copies of C
- Z is the union of three translated copies of C

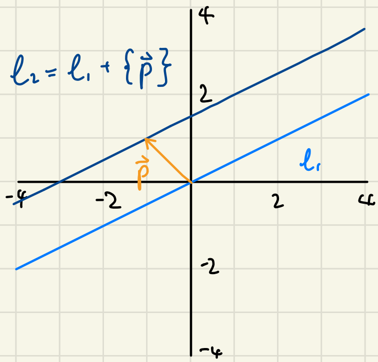


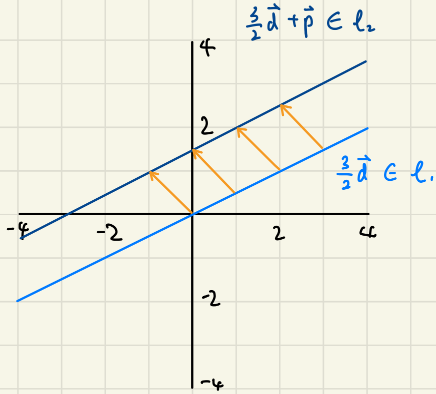
Translated Spans

- lines ℓ_1 and ℓ_2 have vector forms $\vec{x} = t\vec{d}$ and $\vec{x} = t\vec{d} + \vec{p}$, respectively, where $\vec{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{p} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
  - these lines differ from each other by a translation

- every point in ℓ_2 can be obtained by adding $\vec{p}$ to a corresponding point in ℓ_1

- $\ell_2 = \ell_1 + \{\vec{p}\}$





- "$\ell_2 = \ell_1 + \vec{p}$" is incorrect
  - $\vec{p}$ is not a set

- example: $\ell_2 \subseteq \mathbb{R}$ is a line described by the equation $4x - 2y = 6$. Describe ℓ_2 as a translated span

  - ℓ_2 in vector form: $\vec{x} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

  therefore, $\ell_2 = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\} + \left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$

- if Q is described in vector form by

$\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p}$

then, $Q = \text{span}\{\vec{d}_1, \vec{d}_2\} + \{\vec{p}\}$

- all lines and planes, whether through the origin or not, can be expressed as translated spans

Linear Independence & Linear Dependence

- Let $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

- since $\vec{w} = \vec{u} + \vec{v}$, $\vec{w} \in \text{span}\{\vec{u}, \vec{v}\}$
    - $\text{span}\{\vec{u}, \vec{v}\}$ is the xy-plane in $\mathbb{R}^3$ and $\vec{w}$ lies on that plane

- Since $\vec{w}$ is a linear combination of $\vec{u}$ and $\vec{v}$, we can't get anywhere new by taking linear combinations of $\vec{u}, \vec{v}, \vec{w}$ compared to linear combinations of just $\vec{u}$ and $\vec{v}$

- $\text{span}\{\vec{u}, \vec{v}\} = \text{span}\{\vec{u}, \vec{v}, \vec{w}\}$

- $\vec{w}$ is a redundant vector, it is not needed for the span

- when a set contains a redundant vector, we call the set linearly dependent
- geometric: vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly dependent if for at least one i,
  $v_i \in \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_n\}$
  otherwise, they are called linearly independent
- geometric definition of linear independence:
  the vectors $\vec{v}_1, \dots, \vec{v}_n$ are linearly dependent if you can remove at least 1 vector w/o changing the span
- example: let $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\vec{c} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, $d = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$

  determine whether $\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\}$ is linearly independent or dependent
  - by inspection, $\vec{c} = 2\vec{b}$
    therefore, $\text{span}\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\} = \text{span}\{\vec{a}, \vec{b}, \vec{d}\}$
    and so $\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\}$ is linearly dependent
- example: planes P and Q in vector form are

$$\vec{x} = t\begin{bmatrix}1\\2\\1\end{bmatrix} + s\begin{bmatrix}2\\2\\1\end{bmatrix} \text{ and } \vec{x} = t\begin{bmatrix}3\\4\\2\end{bmatrix} + s\begin{bmatrix}2\\2\\1\end{bmatrix},$$

respectively. Determine if P and Q are on the same plane

- let $\vec{a}_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix}2\\2\\1\end{bmatrix}$;

  $\vec{b}_1 = \begin{bmatrix}3\\4\\2\end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix}2\\2\\1\end{bmatrix}$

  be the direction vectors for P and Q, respectively
- $P = Q$ if every direction vector of Q is a linear combination of the direction vectors for P
- $P = Q$ if $\{\vec{a}_1, \vec{a}_2, \vec{b}_1\}$ and $\{\vec{a}_1, \vec{a}_2, \vec{b}_2\}$ are both linearly dependent sets
- since $\vec{a}_2 = \vec{b}_2$, $\{\vec{a}_1, \vec{a}_2, \vec{b}_1\}$ is

linearly dependent

- since $\{\vec{a}_1, \vec{a}_2\}$ is linearly independent, we only need to check if b_1 can be written as a linear combination of $\vec{a}_1$ and $\vec{a}_2$

- $$\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

  $\vec{b}_1 = \vec{a}_1 + \vec{a}_2$, and so $\{\vec{a}_1, \vec{a}_2, \vec{b}_1\}$ is linearly dependent, therefore, $P = Q$

- suppose $\vec{u}, \vec{v}, \vec{w}$ satisfy $\vec{w} = \vec{u} + \vec{v}$, $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly dependent since $\vec{w} \in \text{span}\{\vec{u}, \vec{v}\}$
- equation can be rearranged to $\vec{0} = \vec{u} + \vec{v} - \vec{w}$
  - the right side of the equation has non-zero coefficients, which makes the linear combination non-trivial

- the linear combination $a_1\vec{v}_1 + \ldots + a_n\vec{v}_n$ is called trivial if $a_1 = \ldots = a_n = 0$; if at least one $a_i \neq 0$, the linear combination is called non-trivial

- we can always write $\vec{0}$ as a linear combination of vectors if we let all the coefficients be 0
- only when the vectors are linearly dependent can we write $\vec{0}$ as a non-trivial linear combination of vectors
- algebraic: the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly dependent if there is a non-trivial combination of $\vec{v}_1, \dots, \vec{v}_n$ that equals the zero vector
    - otherwise they are linearly independent
- example: let $\vec{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$. Use the algebraic definition of linear independence to determine whether $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent or dependent
    - determine if there is a non-trivial solution to

$$x\vec{u} + y\vec{v} + z\vec{w} = \vec{0}$$

$$\begin{cases} x + 2y + 4z = 0 \\ 2x + 3y + 5z = 0 \end{cases}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$$

- $(x,y,z) = (-2,3,1)$ is a solution to this system. Therefore $\{\vec{u},\vec{v},\vec{w}\}$ is linearly dependent
- the geometric and algebraic definitions of linear independence are equivalent

Linear Independence and Unique Solutions

- the linearly dependent vectors

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

satisfy the non-trivial relationship $\vec{u} + \vec{v} - \vec{w} = \vec{0}$

- the above equation can be used to generate others
  - i.e. $17\vec{u} + 17\vec{v} - 17\vec{w} = \vec{0}$
- if the equation $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = 0$ has a non-trivial solution, it has infinitely many

non-trivial solutions

- conversely, if the above equation has infinitely many solutions, one of them has to be non-trivial

- homogeneous equations: equations where one side is $\vec{0}$
- a system of linear equations or a vector equation in the variables $a_1 \dots a_n$ is called homogeneous if it takes the form $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n = \vec{0}$, where the right side of the equation is $\vec{0}$
- the vectors $\vec{v}_1, \dots, \vec{v}_n$ are linearly independent if and only if the homogeneous equation $a_1\vec{v}_1 + \dots + a_n\vec{v}_n = \vec{0}$ has a unique solution
- to decide if the vectors $\vec{a}, \vec{b}, \vec{c}$ are linearly independent, we could either:
  - find a non-trivial solution to $x\vec{a} + y\vec{b} + z\vec{c} = \vec{0}$
  - show that $x\vec{a} + y\vec{b} + z\vec{c} = 0$ has more than 1 solution

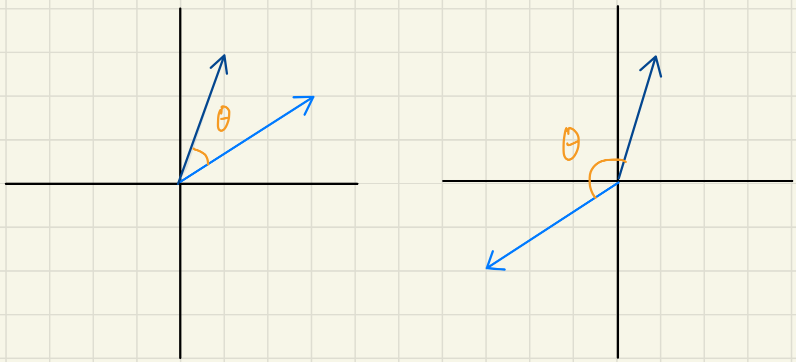
Linear Independence and Vector Form

- the equation $\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2$ represents a plane in vector form whenever $\vec{d}_1$ and $\vec{d}_2$ are non-zero, non-parallel vectors
    - in other words, the equation above represents a plane whenever $\{\vec{d}_1, \vec{d}_2\}$ is linearly independent
- the equation $\vec{x} = t\vec{d}$ represents a line in vector form when $\vec{d} \neq \vec{0}$
    - $\{\vec{d}\}$ is linearly independent when $\vec{d} \neq 0$
- this reasoning generalizes to volumes
- the equation $\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2 + t_3\vec{d}_3$ represents a volume in vector form when $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ is linearly independent
    - if $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ is linearly dependent, the above equation would at best represent a plane
- when writing an object in vector form, the direction vectors must always be linearly independent

Module 4

Dot Product

- Let $\vec{a}$ and $\vec{b}$ be vectors rooted at the same point and let θ denote the smaller of the 2 angles between them (so $0 \leq \theta \leq \pi$)
    - the dot product of $\vec{a}$ and $\vec{b}$ is defined to be $\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta$
        - this is the geometric definition of the dot product



- the dot product is also called the scalar product because the result is a scalar
- algebraically, we can define the dot product

in terms of coordinates

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

- this is the algebraic definition of the dot product

- Ex. Find the angle between the vectors $\vec{v} = (1, 2, 3)$ and $\vec{w} = (1, 1, -2)$

- from the algebraic definition of the dot product, we know

$$\vec{v} \cdot \vec{w} = 1(1) + 2(1) + 3(-2) = -3$$

- from the geometric definition, we know

$$\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|\vec{w}\| \cos\theta = \sqrt{14}\sqrt{6} \cos\theta = 2\sqrt{21} \cos\theta$$

- equating the two definitions of $\vec{v} \cdot \vec{w}$,

$$\cos\theta = \frac{-3}{2\sqrt{21}}$$

- and so $\theta = \arccos\left(\frac{3}{2\sqrt{21}}\right)$

- since the angle between $\vec{a}$ and itself is 0, the geometric definition of dot product tells us

$$\vec{a}\cdot\vec{a} = \|\vec{a}\|\,\|\vec{a}\|\cos 0 = \|\vec{a}\|^2$$

in other words, $\|\vec{a}\| = \sqrt{\vec{a}\cdot\vec{a}}$

- For any vectors $\vec{a}$, $\vec{b}$, $\vec{c}$ and any scalar k,
  - $(\vec{a}+\vec{b})\cdot\vec{c} = \vec{a}\cdot\vec{c} + \vec{b}\cdot\vec{c}$
  - $\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}$
  - $(k\vec{a})\cdot\vec{b} = k(\vec{a}\cdot\vec{b}) = \vec{a}\cdot(k\vec{b})$
  - $\vec{a}\cdot\vec{b} = \vec{b}\cdot\vec{a}$

Orthogonality

- for vectors $\vec{a}$ and $\vec{b}$, the relationship $\vec{a}\cdot\vec{b} = 0$ can hold for 2 reasons:

  1. $\vec{a} = \vec{0}$ or $\vec{b} = \vec{0}$ (or both)
  2. $\vec{a}$ and $\vec{b}$ meet at $90°$

- the dot product can be used to tell if the two vectors are perpendicular
- two vectors $\vec{u}$ and $\vec{v}$ are orthogonal to each other if $\vec{u}\cdot\vec{v} = 0$

- the word orthogonal is synonymous w/ the word perpendicular
- the definition orthogonal encapsulates both the idea of 2 vectors forming a right angle and the idea of one of them being $\vec{0}$
- the vector $\vec{u}$ points in the direction of the vector $\vec{v}$ if $k\vec{u} = \vec{v}$ for some scalar k
- the vector $\vec{u}$ points in the positive direction of $\vec{v}$ if $k\vec{u} = \vec{v}$ for some positive scalar k
- Ex. let $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\vec{c} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Which vector out of $\vec{a}$, $\vec{b}$, and $\vec{c}$ has a direction closest to the direction of $\vec{v}$?

- by equating the geometric and algebraic definitions of the dot product,

$$\cos\theta = \frac{\vec{p} \cdot \vec{q}}{\|\vec{p}\|\,\|\vec{q}\|}$$

- let α, β, γ be the angles between $\vec{v}$ and

$\vec{a}, \vec{b}, \vec{c}$, respectively

$$\cos\alpha = \frac{3+8}{5\sqrt{5}} \approx 0.9839$$

$$\cos\beta = \frac{9+12}{5\sqrt{18}} \approx 0.9899$$

$$\cos\gamma = \frac{6+4}{5\sqrt{5}} \approx 0.8944$$

- since $\cos\beta$ is the closest to 1, $\vec{b}$ has a direction closest to that of $\vec{v}$

Normal Form of Lines and Planes

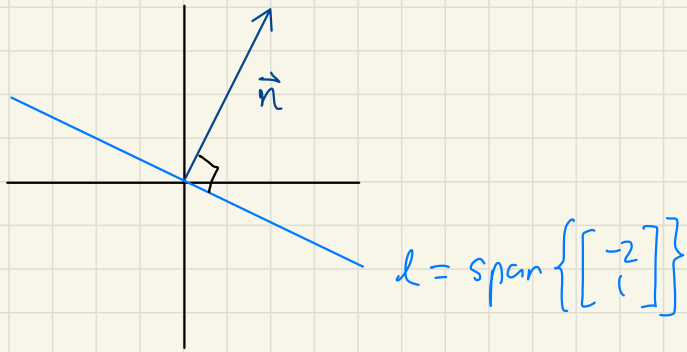
- let $\vec{n} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. if a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is orthogonal to $\vec{n}$, then $\vec{n} \cdot \vec{v} = v_1 + 2v_2 = \vec{0}$, and so $v_1 = -2v_2$
  - $\vec{v}$ is orthogonal to $\vec{n}$ when

$$\vec{v} \in \text{span}\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$$

- the set of all vectors orthogonal to $\vec{n}$

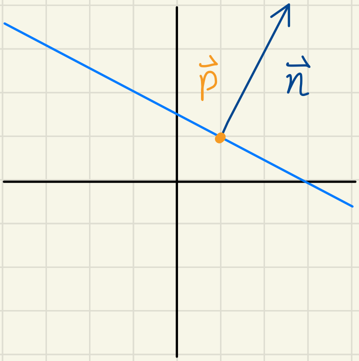
forms a line $\ell = \text{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\}$

- in this case, we call $\vec{n}$ a normal vector of ℓ



- a normal vector to a line (or plane / hyperplane) is a non-zero vector that is orthogonal to all direction vectors for the line (or plane / hyperplane)

- let $\vec{n} = \begin{bmatrix}1\\2\end{bmatrix}$ as before and fix $\vec{p} = \begin{bmatrix}1\\1\end{bmatrix}$. If we draw the set of all vectors orthogonal to $\vec{n}$ but root all the vectors at $\vec{p}$, we get a line passes through $\vec{p}$

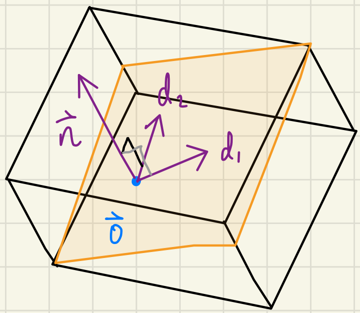


$\ell_2 = \operatorname{span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- $\ell_2 = \operatorname{span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \ell + \{\vec{p}\}$
  - ℓ translated by $\vec{p}$
- for every $\vec{v} \in \ell_2$, $\vec{n} \cdot (\vec{v} - \vec{p}) = 0$
- when a line is expressed as above, it is expressed in normal form
- a line $\ell \subseteq \mathbb{R}^2$ is expressed in normal form if there exist vectors $\vec{n} \neq \vec{0}$ and $\vec{p}$ so that ℓ is the solution to the equation $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$
  - the above equation is called the normal form of ℓ
- if a line $\ell \in \mathbb{R}^2$ is expressed in normal form as $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$, then $\vec{n}$ is

necessarily a normal vector for ℓ

- fix a non-zero vector $\vec{n} \in \mathbb{R}^3$ and let $Q \subseteq \mathbb{R}^3$ be the set of vectors orthogonal to $\vec{n}$. Q is a plane through the origin, and $\vec{n}$ is a normal vector of the plane Q



- Q is the set of solutions to $\vec{n} \cdot \vec{x} = 0$
- for any $p \in \mathbb{R}^3$, the translated plane $Q + \{\vec{p}\}$ is the solution set to $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$
- Ex. Find vector form and normal form of the plane P passing through the points $A = (1,0,0)$, $B = (0,1,0)$, $C = (0,0,1)$
  - to find the vector form of P, we need 1 point on the plane and 2 direction vectors

$$\vec{d}_1 = \overrightarrow{AB} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{d}_2 = \overrightarrow{AC} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

- using point A, we express P in vector form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

- to find the normal vector

$\vec{n} \cdot \vec{d}_1 = 0 \qquad 0 = -a + b$

$\vec{n} \cdot \vec{d}_2 = 0 \qquad 0 = -a + c$

$\rightarrow b - c = 0 \rightarrow b = c = a$

$\vec{n} = (1, 1, 1)$

- we express P in normal form

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = 0$$

- in $\mathbb{R}^2$, only lines have normal form; in $\mathbb{R}^3$, only planes have normal form

- we call objects in $\mathbb{R}^n$ which have a normal form hyperplanes
- the set $X \subseteq \mathbb{R}^n$ is called a hyperplane if there exists $\vec{n} \neq \vec{0}$ and $\vec{p}$ so that X is the set of solutions to the equation
  $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$

Hyperplanes and Linear Equations

- suppose $\vec{n}, \vec{p} \in \mathbb{R}^3$ and $\vec{n} \neq \vec{0}$, solutions to $\vec{n} \cdot (\vec{x} - \vec{p})$ define a plane P
  - but, $\vec{n} \cdot (\vec{x} - p) = 0$ if and only if
    $\vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{p} = \alpha$
  - since $\vec{n}$ and $\vec{p}$ are fixed, α is a constant
  - expanding using coordinates,
    $\vec{n} \cdot (\vec{x} - \vec{p}) = \vec{n} \cdot \vec{x} - \alpha$
    $= n_x x + n_y y + n_z z - \alpha = 0$
    and so P is the set of solutions to
    $n_x x + n_y y + n_z z - \alpha = 0$
  - the above equation is called

Scalar form of a plane

- use the row reduction algorithm to write the complete solution

- Ex. Let $Q \subseteq \mathbb{R}^3$ be the plane passing through $\vec{p}$ and w/ normal vector $\vec{n}$ where

$$\vec{p} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{n} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Write Q in vector form.

- Q is the set of solutions to $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$

  In scalar form, this equation becomes

  $$\vec{n} \cdot (\vec{x} - \vec{p}) = \vec{n} \cdot \vec{x} - \vec{n} \cdot \vec{p}$$

  $$= x + y + z - 2 = 0$$

- Q is the set of all solutions to

  $x + y + z = 2$

- Using the row reduction algorithm,

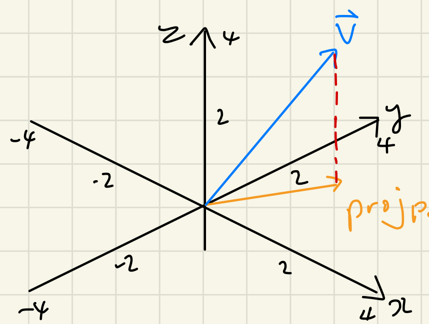
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Module 5

Projection

- Let X be a set. The projection of the vector $\vec{v}$ onto X, written $\text{proj}_X \vec{v}$, is the closest point in X to $\vec{v}$
- Let $P_{xy} \subseteq \mathbb{R}^3$ be the xy-plane in $\mathbb{R}^3$ and

let $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Intuitively, $\text{proj}_{P_{xy}} \vec{v}$ is the "shadow" that $\vec{v}$ would cast on P_{xy} if the sun were directly over-head.



- continuing, let $\ell_y \subseteq \mathbb{R}^3$ be the y-axis in $\mathbb{R}^3$
  - by definition, every vector in ℓ_y takes the

form $\vec{u}_t = \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix}$ for some $t \in \mathbb{R}$.

- the distance between $\vec{u}_t$ and $\vec{v}$ is

$$\|\vec{u}_t - \vec{v}\| = \left\| \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|$$

$$= \sqrt{1^2 + (t-2)^2 + 3^2}$$

- since $(t-2)^2$ is always positive, the distance between $\vec{u}_t$ and $\vec{v}$ would be minimized when $t=2$

- $\text{proj}_{\ell_y} \vec{v} = \vec{u}_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$

- Ex. let $\ell \subseteq \mathbb{R}^2$ be the line given in vector form by $\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ and let $\vec{v} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$.

Find $\text{proj}_\ell \vec{v}$.

- Let $\vec{u}_t = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} \in \ell$

The distance between $\vec{v}$ and $\vec{u}_t$:

$$\|\vec{u}_t - \vec{v}\| = \left\| \left(t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} \right) - \begin{bmatrix}-1\\-1\end{bmatrix} \right\|$$

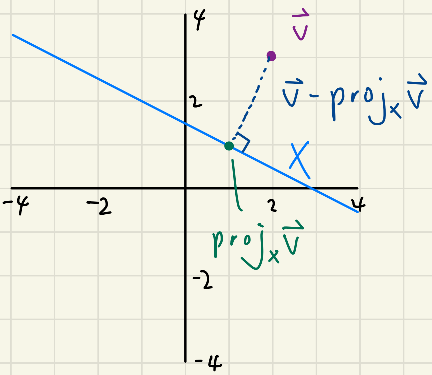
$$= \left\| \begin{bmatrix}t+4\\t-1\end{bmatrix} \right\| = \sqrt{2t^2+6t+17}$$

- $2t^2+6t-17$ is minimized when $t = -\frac{3}{2}$,

the closest point in ℓ to $\vec{v}$ is $\vec{u}_{-3/2}$

- $\text{proj}_\ell \vec{v} = -\frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} = \begin{bmatrix}3/2\\-7/2\end{bmatrix}$

- the vector from the projection to the original point is a normal vector for the line or plane



- if X is a line or plane and $\vec{v} \notin X$ is a vector, then $\vec{v} - \text{proj}_X \vec{v}$ is a normal vector for X.

- Ex. Let $\ell \subseteq \mathbb{R}^2$ be the line given in vector form by $\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -2 \end{bmatrix}$, and let $\vec{v} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

Find $\text{proj}_\ell \vec{v}$.

  - since $\vec{v} - \text{proj}_\ell \vec{v}$ is a normal vector to ℓ,

    $\vec{v} - \text{proj}_\ell \vec{v}$ is orthogonal to $\vec{d} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  - let $\begin{bmatrix} x \\ y \end{bmatrix} = \text{proj}_\ell \vec{v}$ for some $x, y \in \mathbb{R}$

  - $(\vec{v} - \text{proj}_\ell \vec{v}) \cdot \vec{d} = \left(\begin{bmatrix} -1 \\ -1 \end{bmatrix} - \begin{bmatrix} x \\ y \end{bmatrix}\right) \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

    $= \begin{bmatrix} -1-x \\ -1-y \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -2-x-y = 0$

  - $x + y = -2$

  - since $\text{proj}_\ell \vec{v} \in \ell$,

$$\text{proj}_\ell \vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} t+3 \\ t-2 \end{bmatrix}$$

$x - t = 3, \quad y - t = 2$

$$\begin{cases} \qquad x + y = -2 \\ -t + x \qquad = 3 \\ -t \qquad + y = 2 \end{cases}$$

- $x = \frac{3}{2}, \; y = -\frac{7}{2}$, we don't care about t

- $\text{proj}_\ell \vec{v} = \begin{bmatrix} 3/2 \\ -7/2 \end{bmatrix}$

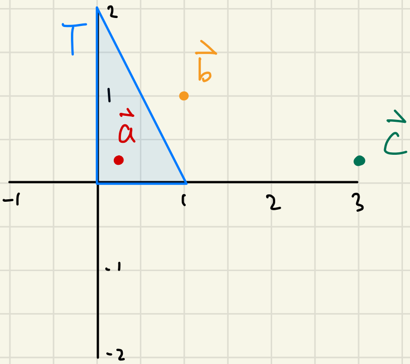
Projections Onto Other Sets

- Ex. Let $T \subseteq \mathbb{R}^2$ be the filled in triangle with

vertices $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, and let

$\vec{a} = \begin{bmatrix} 1/4 \\ 1/4 \end{bmatrix}, \; \vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; \vec{c} = \begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$

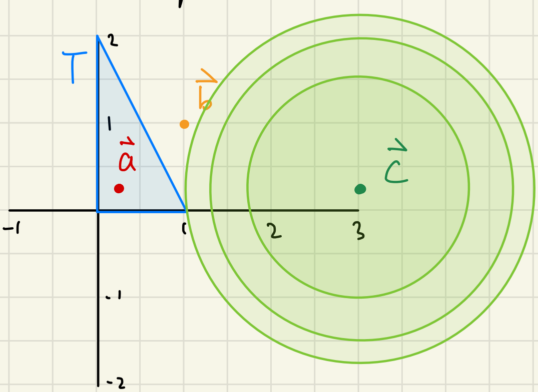
Find $\text{proj}_T \vec{a}, \; \text{proj}_T \vec{b}, \; \text{proj}_T \vec{c}$

- 

- since $\vec{a} \in T$, $\text{proj}_T \vec{a} = \vec{a}$
- since $\vec{b}$ is closest to the hypotenuse of T, $\text{proj}_T \vec{b}$ is the same as the projection of $\vec{b}$ onto the line $y = -2x + 2$

$$\text{proj}_T \vec{b} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$$

- drawing concentric circles centered at $\vec{c}$, the lower-right corner of T is the closest point in T to $\vec{c}$



and so, $\text{proj}_{U_T} \vec{c} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

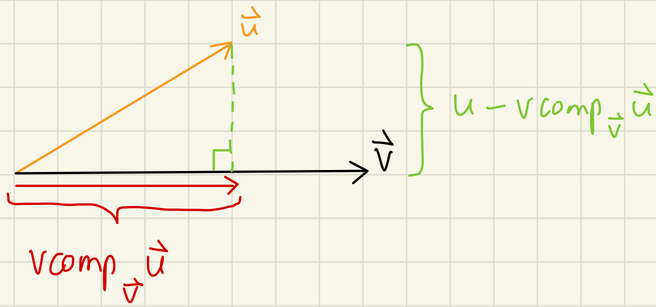
Subtleties of Projections

- if $\vec{v}$ is equidistant from two closest points in X, or if X is an open set (i.e. an open interval in $\mathbb{R}^1$), there will not be a closest point in X to $\vec{v}$
    - in both cases, $\text{proj}_X \vec{v}$ is undefined

Vector Components

- suppose $\vec{v} \neq \vec{0}$ and $\vec{u}$ are vectors, we can decompose $\vec{u}$ into the sum of two vectors:
    - one in the direction of $\vec{v}$
    - one orthogonal to $\vec{v}$
    - the tool that does this is the vector component
- let $\vec{u}$ and $\vec{v} \neq 0$ be vectors, the vector component of $\vec{u}$ in the $\vec{v}$ direction, written $\text{vcomp}_{\vec{v}} \vec{u}$, is the vector in the direction of $\vec{v}$ so that

$\vec{u} - \text{vcomp}_{\vec{v}}\vec{u}$ is orthogonal to $\vec{v}$



- $\vec{u} = \text{vcomp}_{\vec{v}}\vec{u} + (\vec{u} - \text{vcomp}_{\vec{v}}\vec{u})$ is a decomposition of $\vec{u}$ into the sum of two vectors:
  - one ($\text{vcomp}_{\vec{v}}\vec{u}$) is in the direction of $\vec{v}$
  - one ($\vec{u} - \text{vcomp}_{\vec{v}}\vec{u}$) is orthogonal to $\vec{v}$
- Ex. Find the vector component of

$\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in the direction of $\vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  - since $\text{vcomp}_{\vec{b}}\vec{a}$ is a vector in the direction of $\vec{b}$,

    $\text{vcomp}_{\vec{b}}\vec{a} = k\vec{b}$ for some $k \in \mathbb{R}$
  - since $\vec{a} - \text{vcomp}_{\vec{b}}\vec{a}$ is orthogonal to $\vec{b}$,

    $(\vec{a} - \text{vcomp}_{\vec{b}}\vec{a}) \cdot \vec{b} = 0$
  - combining the above equations,

$$(\vec{a} - \text{vcomp}_{\vec{b}}\,\vec{a}) \cdot \vec{b} = \Bigg(\underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{a} - \underbrace{\begin{bmatrix} k \\ k \end{bmatrix}}_{k\vec{b}}\Bigg) \cdot \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\vec{b}}$$

$$= (1-k) + (2-k) = 3 - 2k = 0$$

and so $k = \frac{3}{2}$.

- $\text{vcomp}_{\vec{b}}\,\vec{a} = k\vec{b} = \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix}$

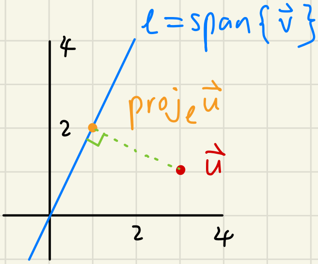
- $\text{vcomp}_{\vec{v}}\,\vec{u}$ is a vector in the direction of $\vec{v}$, so
  - $\text{vcomp}_{\vec{v}}\,\vec{u} = k\vec{v}$
- $\vec{u} - \text{vcomp}_{\vec{v}}\,\vec{u}$ is orthogonal to $\vec{v}$, and so
  - $\vec{v} \cdot (\vec{u} - \text{vcomp}_{\vec{v}}\,\vec{u}) = \vec{v} \cdot (\vec{u} - k\vec{v})$

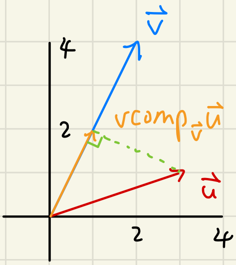
    $= \vec{v} \cdot \vec{u} - k\vec{v} \cdot \vec{v} = 0$
- because $\vec{v} \neq \vec{0}$, we know $\vec{v} \cdot \vec{v} \neq 0$, therefore,
  - $k = \dfrac{\vec{v} \cdot \vec{u}}{\vec{v} \cdot \vec{v}}$
- $\text{vcomp}_{\vec{v}}\,\vec{u} = \left(\dfrac{\vec{v} \cdot \vec{u}}{\vec{v} \cdot \vec{v}}\right)\vec{v}$

The Relationship Between Vector Components and Projections

- let $\vec{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, let $\ell = \text{span}\{\vec{v}\}$

  - drawing ℓ, $\vec{u}$, and $\vec{v}$, we see that $\text{proj}_\ell \vec{u}$ satisfies all the properties of $\text{vcomp}_{\vec{v}} \vec{u}$





- since $\ell = \text{span}\{\vec{v}\}$ and $\text{proj}_\ell \vec{u} \in \ell$, $\text{proj}_\ell \vec{u}$ is in the direction of $\vec{v}$
- $\vec{u} - \text{proj}_\ell \vec{u}$ is a normal vector for ℓ and is therefore orthogonal to its direction vector $\vec{v}$
- for vectors $\vec{u}$ and $\vec{v} \neq 0$, we have

$$\text{proj}_{\text{span}\{\vec{v}\}} \vec{u} = \text{vcomp}_{\vec{v}} \vec{u}$$

- Ex. compute the projection of

$$\vec{a} = \begin{bmatrix} 3 \\ 7 \end{bmatrix} \text{ onto } \ell = \text{span}\left\{\begin{bmatrix} 1 \\ -4 \end{bmatrix}\right\}$$

- let $\vec{b} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$

- since $\ell = \text{span}\{\vec{b}\}$ and $\vec{b} \neq \vec{0}$,

$$\text{proj}_\ell \vec{a} = \text{vcomp}_{\vec{b}} \vec{a} = \left(\frac{\vec{b} \cdot \vec{a}}{\vec{b} \cdot \vec{b}} \right) \vec{b}$$

$$= \frac{3-28}{1+16} \begin{bmatrix} 1 \\ -4 \end{bmatrix} = \begin{bmatrix} -25/17 \\ 100/17 \end{bmatrix}$$

- vector components are equal to projections only in the case when projecting onto a span
  - in general, projections and vector compositions are unrelated

- Ex. let $\vec{a} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, and let ℓ be the line given in vector form by $\vec{x} = t\vec{b} + \vec{a}$.

  Show that $\text{proj}_\ell \vec{a} \neq \text{vcomp}_{\vec{b}} \vec{a}$

  - $\text{proj}_\ell \vec{a}$ is the closest point in ℓ to $\vec{a}$

  - since $\vec{a} \in \ell$, $\text{proj}_\ell \vec{a} = \vec{a} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$

- $vcomp_{\vec{b}}\,\vec{a}$ is already computed in the last example
- $vcomp_{\vec{b}}\,\vec{a} = \begin{bmatrix} -25/17 \\ 100/17 \end{bmatrix} \neq \begin{bmatrix} 3 \\ 7 \end{bmatrix} = proj_{\ell}\,\vec{a}$

- When projecting onto the span of a single vector, we can use vector components as a computational shortcut
  - if the set isn't a span, we cannot

Module 6

Subspaces

- since a linear combination of linear combinations is still a linear combination, a span is closed with respect to linear combinations
  - i.e. by taking linear combinations of vectors in a span, we cannot escape the span
  - sets having the above property are called subspaces
- A non-empty subset $V \subseteq \mathbb{R}^n$ is called a subspace if for all $\vec{u}, \vec{v} \in V$ and all scalars k we have:
  1. $\vec{u} + \vec{v} \in V$; and
  2. $k\vec{u} \in V$
- property 1 is called being closed with respect to vector addition
- property 2 is called being closed with respect to scalar multiplication

- subspaces generalize the idea of flat spaces through the origin, including lines, planes, volumes, etc.

- Ex. Let $V \subseteq \mathbb{R}^2$ be the complete solution to $x + 2y = 0$. Show that V is a subspace.

  - let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be in V,

    let k be a scalar

  - by definition, we have

    $u_1 + 2u_2 = 0$

    $v_1 + 2v_2 = 0$

  - show that $\vec{u} + \vec{v} \in V$

    - $\vec{u} + \vec{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$

    - $(u_1 + v_1) + 2(u_2 + v_2)$

      $= (u_1 + 2u_2) + (v_1 + 2v_2) = 0 + 0 = 0$

  - show that $k\vec{u} \in V$

    - $k\vec{u} = \begin{bmatrix} ku_1 \\ ku_2 \end{bmatrix}$

- $(k\vec{u})_1 + 2(k\vec{u}_2) = k(u_1 + 2u_2) = k0 = 0$

- since $\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ satisfies $x + 2y = 0$,

  we conclude that $\vec{0} \in V$ and so V is non-empty

- Thus, by definition, V is a subspace

- Ex. Let $W \subseteq \mathbb{R}^2$ be the line expressed in vector form as $\vec{x} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  Determine whether W is a subspace

  - W is not a subspace.

  - $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in W$, but $0\vec{v} = \vec{0} \notin W$

  - therefore, W is not closed under scalar multiplication and so it cannot be a subspace.

- Every subspace is a span and every span is a subspace.

  - $V \subseteq \mathbb{R}^n$ is a subspace if and only if

$V = \text{span}X$ for some set X

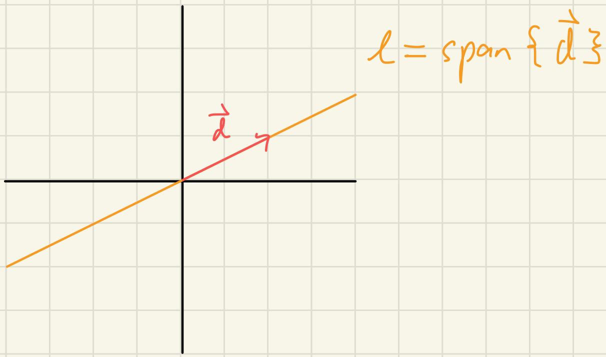
- spans provide a constructive definition of lines/planes/volumes/etc. through the origin
  - when describing a line/plane/volume/etc. through the origin as a span, we're saying "this is a line/plane/etc. through the origin because every point in it is a linear combination of these specific vectors."
- subspaces provide a categorical definition of lines/planes/etc. through the origin
  - when describing a line/plane/etc. through the origin as a subspace, we're saying "this is a line/plane/etc. through the origin because these properties are satisfied."
- spans and subspaces are two different ways of talking about the same objects: points/lines/volumes/etc. through the origin

Special Subspaces

- When thinking about $\mathbb{R}^n$, two special subspaces are always available
    - $\mathbb{R}^n$: it is non-empty, and linear combinations of $\mathbb{R}^n$ remain in $\mathbb{R}^n$
    - $\{\vec{0}\}$: trivial subspace
- $\{\vec{0}\}$ is non-empty since $\vec{0} \in \{\vec{0}\}$
- $\vec{0} + \vec{0} = \vec{0} \in \{\vec{0}\}$
- $\alpha \vec{0} = \vec{0} \in \{\vec{0}\}$

Bases

- Let $\vec{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and consider $\ell = \text{span}\{\vec{d}\}$



- ℓ is defined as the span of $\{\vec{d}\}$

- We can also define $\ell = \text{span}\{\vec{d}, -2\vec{d}, \frac{1}{2}\vec{d}\}$
    - The simplest descriptions of a line involve the span of only 1 vector
- Let $P = \text{span}\{\vec{d}_1, \vec{d}_2\}$ be the plane through the origin w/ direction vectors $\vec{d}_1$ and $\vec{d}_2$
    - The simplest way to write P as a span involve two vectors
- A basis for a subspace V is a linearly independent set of vectors, B, so that $\text{span } B = V$
- A basis for a subspace is a linearly independent set that spans the subspace

- Ex. Let $\ell = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$

Find two different bases for ℓ.

- we are looking for a set of linearly independent vectors that spans ℓ.
- $\text{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix} -2 \\ -4 \end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$

$= \ell$

- Because $\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$ is linearly independent and spans ℓ, it is a basis for ℓ.

- $\left\{\begin{bmatrix}1/2\\1\end{bmatrix}\right\}$ is another basis for ℓ

- Facts about bases

  1. Bases are not unique. Every subspace (except the trivial subspace) has multiple bases

  2. Given a basis for a subspace, every vector in the subspace can be written as a unique linear combination of vectors in that basis

  3. Any 2 bases for the same subspace have the same number of elements

Dimension

- Let V be a subspace, all bases of V have the same number of vectors in them.

- The maximum number of linearly independent vectors that can simultaneously exist in V is the dimension of V

- the dimension of a subspace V is the number of elements in a basis of V

- Ex. Find the dimension of $\mathbb{R}^2$
    - since $\{\vec{e}_1, \vec{e}_2\}$ is a basis for $\mathbb{R}^2$, $\mathbb{R}^2$ is two dimensional.

- Ex. Let $\ell = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$.

  Find the dimension of the subspace ℓ.
    - $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$ and $\left\{\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$ are bases for ℓ
    - ℓ is a one-dimensional space

- Ex. Let $A = \{(x_1, x_2, x_3, x_4):$ $x_1 + 2x_2 - x_3 = 0$ and $x_1 + 6x_4 = 0\}$

  Find the basis for and the dimension of A.
    - A is the complete solution to the system

$$\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ x_1 + 6x_4 = 0 \end{cases}$$

in vector form,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 0 \\ 1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -6 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

- Therefore, $A = \text{span} \left\{ \begin{bmatrix} 0 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}$

- A is two-dimensional.

- The standard basis for $\mathbb{R}^n$ is the set $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ where

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix}, \ldots$$

- That is $\vec{e}_i$ is the vector w/ a

  1 in its ith coordinate and zeros elsewhere

- The notation $\vec{e}_i$ is context specific
  - if $\vec{e}_i \in \mathbb{R}^2$, $\vec{e}_i$ must have 2 components
  - if $\vec{e}_i \in \mathbb{R}^{45}$, $\vec{e}_i$ must have 45 components

Module 7

Matrix Equations

- the system

$$\begin{cases} x+2y-2z=-15 \\ 2x+y-5z=-21 \\ x-4y+z=18 \end{cases}$$

is equivalent to the vector equation

$$\begin{bmatrix} x+2y-2z \\ 2x+y-5z \\ x-4y+z \end{bmatrix} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}$$

which we can rewrite using matrix-vector multiplication:

$$\underbrace{\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -5 \\ 1 & -4 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}$$

- The matrix A on the left is called the coefficient matrix because it is made up of the coefficients

- By using coefficient matrices, every system of linear equations can be written as a single matrix equation of the form $A\vec{x} = \vec{b}$
  - A - coefficient matrix
  - $\vec{x}$ - column vector of variables
  - $\vec{b}$ - column vector of constants
- Ex. Rewrite the below system as a single matrix equation:

$$\left\{ x - 4y + z = 5 \right.$$

  - $$\begin{bmatrix} 1 & -4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}$$

- Ex. Rewrite the below system as a single matrix equation:

$$\begin{cases} x - 4y + z = 5 \\ \quad\ \ y - z = 9 \end{cases}$$

  - $$\begin{bmatrix} 1 & -4 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$$

Interpretations of Matrix Equations

- the solution set to a system of linear equations,

  i.e.
$$\begin{cases} x + 2y - 2z = -15 \\ 2x + y - 5z = -21 \\ x - 4y + z = 18 \end{cases}$$

  can be interpreted as the intersection of 3 planes (or hyperplanes if there were more variables)
  - Each equation (i.e. row) specifies a plane, and the solution set is the intersection of all of these planes
- rewriting a system of equations in matrix form gives 2 additional ways to interpret the solution set

The Column Picture

- The system above is equivalent to

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -5 \\ 1 & -4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= x\begin{bmatrix}1\\2\\1\end{bmatrix} + y\begin{bmatrix}2\\1\\-4\end{bmatrix} + z\begin{bmatrix}-2\\-5\\1\end{bmatrix} = \begin{bmatrix}-15\\-21\\18\end{bmatrix}$$

- "what are the solutions to the system above?" is equivalent to "what coefficients allow

$$\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}2\\1\\-4\end{bmatrix}, \begin{bmatrix}-2\\-5\\1\end{bmatrix} \text{ to form } \begin{bmatrix}-15\\-21\\-18\end{bmatrix}$$

as a linear combination?"

- $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}2\\1\\-4\end{bmatrix}, \begin{bmatrix}-2\\-5\\1\end{bmatrix}$ are the columns of the coefficient matrix

The Row Picture

- Let $\vec{r}_1, \vec{r}_2, \vec{r}_3$ be the rows of the coefficient matrix for the system above. The system is equivalent to

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -5 \\ 1 & -4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \\ \hline \vec{r}_2 \\ \hline \vec{r}_3 \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vec{r}_2 \cdot \vec{x} \\ \vec{r}_3 \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}$$

- we can interpret solutions to the above system as vectors whose dot product
    - w/ $\vec{r}_1$ is -15
    - w/ $\vec{r}_2$ is -21
    - w/ $\vec{r}_3$ is 18
- This perspective is powerful when the right side of the equation is all zeros

<u>Interpreting Homogeneous Systems</u>

- consider the homogeneous system / matrix equation

$$\begin{cases} x + 2y - 2z = 0 \\ 2x + y - 5z = 0 \\ x - 4y + z = 0 \end{cases} \iff \underbrace{\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -5 \\ 1 & -4 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- column interpretation: what linear combinations of the column vectors of A give $\vec{0}$?

- let $\vec{r}_1, \vec{r}_2, \vec{r}_3$ be the rows of A. The row interpretation of the above system asks: what vectors are simultaneously orthogonal to $\vec{r}_1, \vec{r}_2, \vec{r}_3$?

- There are 3 ways to interpret solutions to a matrix equation $A\vec{x} = \vec{b}$:
  - the intersection of hyperplanes specified by rows
  - what linear combinations of the columns of A give $\vec{b}$
  - what vectors yield the entries of $\vec{b}$ when dot producted w/ the rows of A

- Ex. Find all vectors orthogonal to

$$\vec{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

  - we need to find vectors $\vec{x}$ satisfying $\vec{a} \cdot \vec{x} = 0$ and $\vec{b} \cdot \vec{x} = 0$

- $$\begin{bmatrix} \vec{a} \\ \hline \vec{b} \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{a} \cdot \vec{x} \\ \vec{b} \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- by row reducing, $\operatorname{rref}(A) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

- $$\begin{cases} x + z = 0 \\ y = 0 \end{cases}$$

- the complete solution expressed in vector form:

$$\vec{x} = t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

- Ex. Let Q be the hyperplane specified in vector form

by $\vec{x} = t \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} + r \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

Find a normal vector for Q and

write Q in normal form

- the normal vectors for Q need to be orthogonal to

$$d_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad d_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- we can find the normal vectors by solving

$$\underbrace{\begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- by reducing A, we get

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\begin{cases} x = 0 \\ y + w = 0 \\ z = 0 \end{cases}$$

- the complete solution expressed in vector form:

$$\vec{x} = t\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

- Therefore, any non-zero multiple of $\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$

is a normal vector for Q.

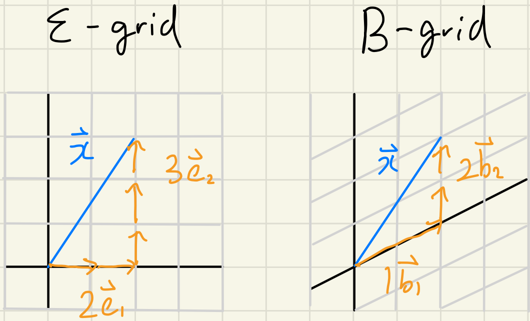
- Q can be written in normal form as

$$\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right) = 0$$

Module 8

Change of Basis

- When we write $\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, we mean $\vec{x} = 2\vec{e}_1 + 3\vec{e}_2$
    - 2 and 3 are the coordinates of the vector $\vec{x}$ with respect to the standard basis
- Subspaces have many bases
    - it is possible to represent a single vector in many different ways as coordinates with respect to different bases
- Let $\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, let $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ be the standard basis for $\mathbb{R}^2$, and let $B = \{\vec{b}_1, \vec{b}_2\}$, where $\vec{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ be another basis for $\mathbb{R}^2$. The coordinates of $\vec{x}$ w.r.t. $\mathcal{E}$ are $(2,3)$, but the coordinates of $\vec{x}$ w.r.t B are $(1,2)$



- The coordinates $(2,3)$ and $(1,2)$ represent $\vec{x}$ equally well
- Let $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis for a subspace V and let $\vec{v} \in V$. The representation of $\vec{v}$ in the B basis, notated $[\vec{v}]_B$, is the column matrix

$$[\vec{v}]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$ where $a_1 \ldots a_n$ uniquely satisfy

$$\vec{v} = a_1\vec{b}_1 + \ldots + a_n\vec{b}_n.$$

Conversely, $$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}_B = a_1\vec{b}_1 + \ldots + a_n\vec{b}_n$$

is notation for the linear combination of $\vec{b}_1, \ldots, \vec{b}_n$ w/ coefficients $a_1, \ldots, a_n$

- Ex. Let $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ be the standard basis for $\mathbb{R}^2$ and let $C = \{\vec{c}_1, \vec{c}_2\}$ where $\vec{c}_1 = \vec{e}_1 + \vec{e}_2$ and $\vec{c}_2 = 3\vec{e}_2$ be another basis for $\mathbb{R}^2$. Given that $\vec{v} = 2\vec{e}_1 - \vec{e}_2$, find $[\vec{v}]_{\mathcal{E}}$ and $[\vec{v}]_C$.

  - since $\vec{v} = 2\vec{e}_1 - \vec{e}_2$, we know $[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

  - to find $[\vec{v}]_C$, we need to write $\vec{v}$ as a linear combination of $\vec{c}_1$ and $\vec{c}_2$.

  - suppose $\vec{v} = x\vec{c}_1 + y\vec{c}_2$ for some unknown scalars x and y.

  - on the one hand, $\vec{v} = 2\vec{e}_1 - \vec{e}_2$

  - on the other hand, $\vec{v} = x\vec{c}_1 + y\vec{c}_2$

    $= x(\vec{e}_1 + \vec{e}_2) + 3y\vec{e}_2 = x\vec{e}_1 + (x+3y)\vec{e}_2$

  - combining the above two equations,

    $2\vec{e}_1 - \vec{e}_2 = x\vec{e}_1 + (x+3y)\vec{e}_2$

  - and so $(x-2)\vec{e}_1 + (x+3y+1)\vec{e}_2 = \vec{0}$

  - since $\vec{e}_1$ and $\vec{e}_2$ are linearly independent, their coefficients must equal to 0

$$\begin{cases} x = 2 \\ x + 3y = -1 \end{cases}$$

- after solving, we see $\vec{v} = 2\vec{c}_1 - \vec{c}_2$, and so

$$[\vec{v}]_C = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Notation Conventions

- given the representation-in-a-basis notation, to mean $\vec{x} = 2\vec{e}_1 + 3\vec{e}_2$, we should be writing

$$\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathcal{E}}$$, where $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$

- if a problem involves only one basis, we may write $\begin{bmatrix} x \\ y \end{bmatrix}$ to mean $\begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{E}}$, where $\mathcal{E}$ is the standard basis

- if there are multiple bases in a problem, we will always write $\begin{bmatrix} x \\ y \end{bmatrix}_X$ to specify a vector in coordinates relative to a particular basis X

True Vectors vs. Representations

- Let $\vec{x} = 2\vec{e}_1 + 3\vec{e}_2 \in \mathbb{R}^2$
    - The vector $\vec{x}$ is a real-life geometrical thing
    - We call $\vec{x}$ a true vector
- When we write the column matrix $[\vec{x}]_\varepsilon = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$,

  we are writing a list of numbers
    - the list of numbers has no meaning until we give it a meaning by assigning it a basis
- By writing $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_\varepsilon$, we declare that the numbers 2 and 3 are the coefficients of $\vec{e}_1$ and $\vec{e}_2$
- Since a list of numbers without a basis has no meaning,

  we must acknowledge $\vec{x} \neq [\vec{x}]_\varepsilon = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ since

  the left side of the equation is a true vector and the right side is a list of numbers
- Similarly, we must acknowledge $[\vec{x}]_\varepsilon \neq \begin{bmatrix} 2 \\ 3 \end{bmatrix}_\varepsilon = \vec{x}$

since the left side is a list of numbers and the right side is a true vector

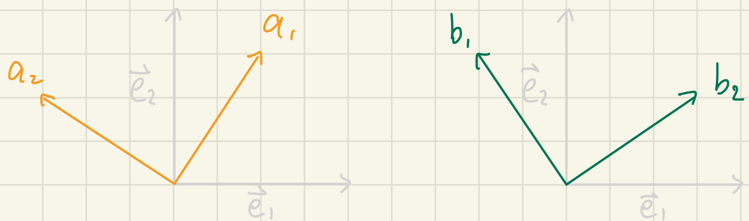
- $[\text{true vector}]_X$ = list of numbers

  $[\text{list of numbers}]_X$ = true vector

Orientation of a Basis

- Consider the ordered bases $\mathcal{E}$, A, B shown below

$A = \{\vec{a}_1, \vec{a}_2\}$ $\qquad$ $B = \{\vec{b}_1, \vec{b}_2\}$



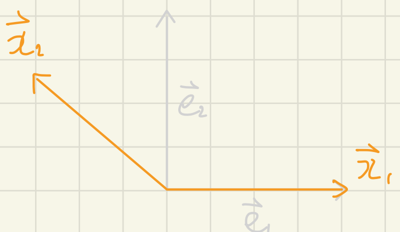
- The A basis can be rotated to get the $\mathcal{E}$ basis while maintaining the proper order of the basis vectors
    - i.e. $\vec{a}_1 \to \vec{e}_1$ and $\vec{a}_2 \to \vec{e}_2$
- It is impossible to rotate the B basis to get the $\mathcal{E}$ basis while maintaining the proper order
- $\mathcal{E}$ and A have the same orientation
- $\mathcal{E}$ and B have opposite orientations
- A and B have different orientations

- Two types of orientations of bases:
    - right-handed (positively oriented)
    - left-handed (negatively oriented)
- By convention, the standard basis is called right-handed
- Orthogonal bases: bases consisting of unit vectors that are orthogonal to each other
    - if they can be rotated to align w/ the standard basis, they are right-handed
    - else, they are left-handed
- a left hand and a right hand can never be rotated to alignment
- The ordered basis $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is right-handed or positively oriented if it can be continuously transformed to the standard basis (with $\vec{b}_i \to \vec{e}_i$) while remaining linearly independent throughout the transformation; otherwise, B is called left-handed or negatively oriented
- Continuously transformed: one basis transforms

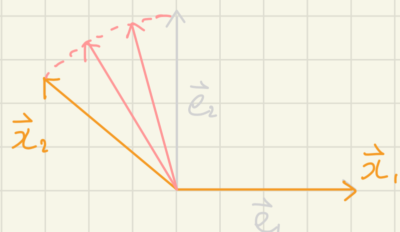
into the other smoothly and w/o jumps

- Let $X = \{\vec{x}_1, \vec{x}_2\}$, we could imagine $\vec{x}_1, \vec{x}_2$ continuously transforming to $\vec{e}_1, \vec{e}_2$ by $\vec{x}_1$ staying in place and $\vec{x}_2$ smoothly moving along the dotted line

$X = \{\vec{x}_1, \vec{x}_2\}$



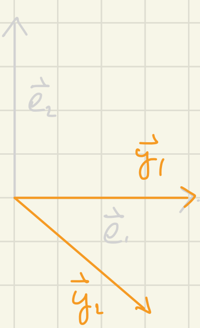
Continuous transform of X to $\mathcal{E}$



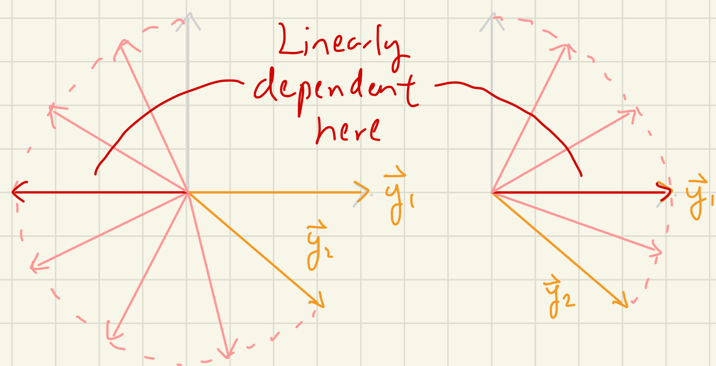
- Because at every step along this motion, the set $\vec{x}_1$ and the transformed $\vec{x}_2$ is linearly independent, X is positively oriented.

- Let $Y = \{\vec{y}_1, \vec{y}_2\}$.

$Y = \{\vec{y}_1, \vec{y}_2\}$



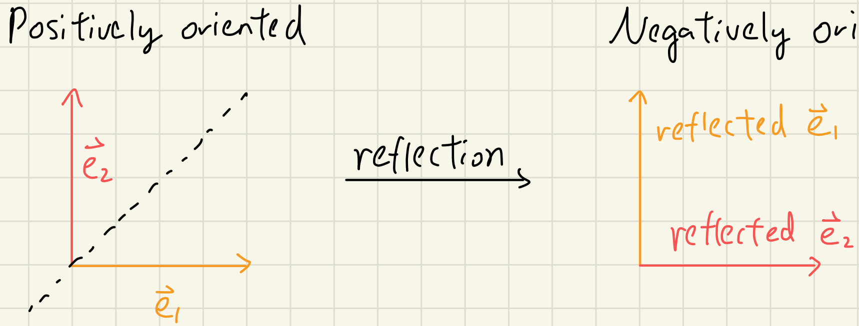
Continuous transform of Y to $\mathcal{E}$



- Somewhere along $\vec{y}_2$'s path, the set of $\vec{y}_1$ and the transformed $\vec{y}_2$ becomes linearly dependent
- There are no paths for them to stay linearly independent because Y is negatively oriented

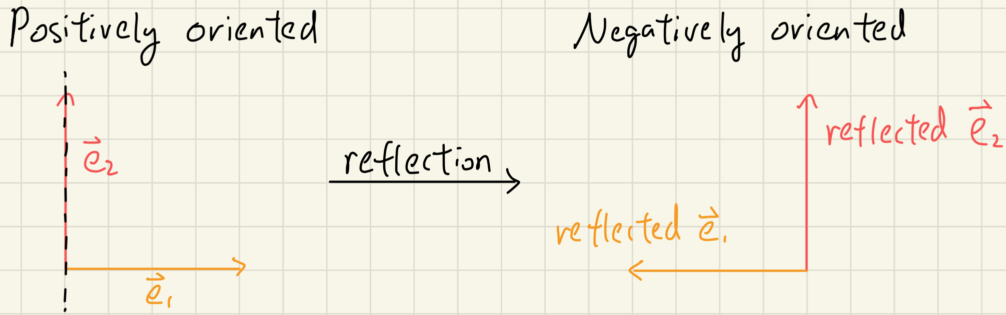
Reversing Orientation

- consider the reflection of $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ across the line $y = x$



  - this reflection sends $\{\vec{e}_1, \vec{e}_2\} \to \{\vec{e}_2, \vec{e}_1\}$

- alternatively, reflection across the line $x = 0$ sends $\{\vec{e}_1, \vec{e}_2\} \to \{-\vec{e}_1, \vec{e}_2\}$



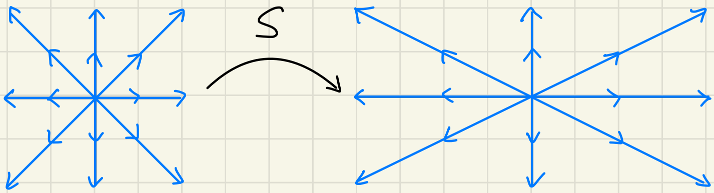
- both $\{\vec{e}_2, \vec{e}_1\}$ and $\{-\vec{e}_1, \vec{e}_2\}$, as ordered bases, are negatively oriented
- Let $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ be an ordered basis. The ordered basis obtained from B by replacing $\vec{b}_i$ w/ $-\vec{b}_i$, and the ordered basis obtained from B by swapping the order of $\vec{b}_i$ and $\vec{b}_j$ (with $i \neq j$) have the opposite orientation as B

Module 9

Transformation

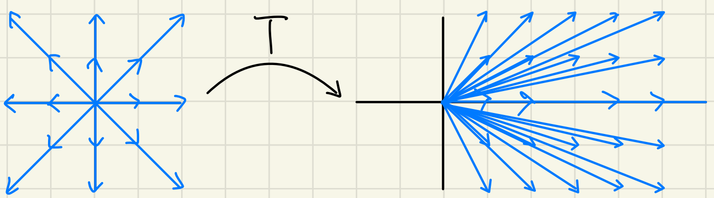
- Transformation is another word for a function
- i.e. $S: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} 2x \\ y \end{bmatrix}$

  stretches all vectors in the $\vec{e}_1$ direction by a factor of 2



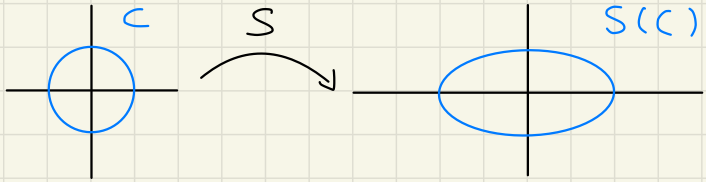
- i.e. $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} x+3 \\ y \end{bmatrix}$

  translates all vectors 3 units in the $\vec{e}_1$ direction



Images of Sets

- Let C be the unit circle, applying S to all vectors that make up C produces an ellipse

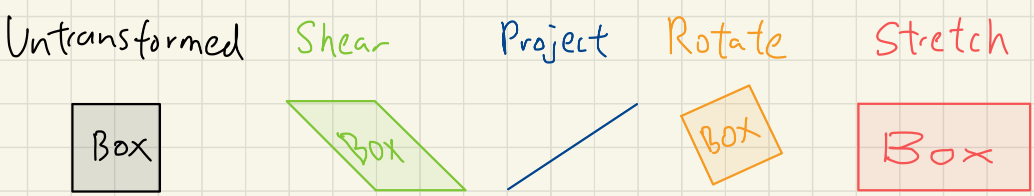


- the operation of applying a transformation to a specific set of vectors and seeing what results is called taking the image of a set
- Let $L: \mathbb{R}^n \to \mathbb{R}^m$ be a transformation and let $X \subseteq \mathbb{R}^n$ be a set. The image of the set X under L, denoted $L(X)$, is the set

$$L(X) = \{\vec{y} \in \mathbb{R}^m : \vec{y} = L(\vec{x}) \text{ for some } \vec{x} \in X\}$$

- the image of a set X under a transformation L is the set of all outputs of L when the inputs comes from X

<u>Linear Transformations</u>



- Let V and W be subspaces. A function $T: V \to W$ is called a linear transformation if

$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ and $T(a\vec{v}) = aT(\vec{v})$

for all vectors $\vec{u}, \vec{v} \in V$ and all scalars a

- The transformation T is linear if it distributes over addition and scalar multiplication
    - T distributes over linear combinations

- Ex. Let $S: \mathbb{R}^2 \to \mathbb{R}^2$ and $T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow{S} \begin{bmatrix} 2x \\ y \end{bmatrix} \text{ and } \begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow{T} \begin{bmatrix} x \\ y+4 \end{bmatrix}$$

For each of S and T, determine whether the transformation is linear.

- Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be vectors,

  let a be a scalar

- $S(\vec{u}+\vec{v}) = S\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} 2u_1+2v_1 \\ u_2+v_2 \end{bmatrix}$

$$= \begin{bmatrix} 2u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 2v_1 \\ v_2 \end{bmatrix} = S(\vec{u}) + S(\vec{v})$$

- $S(a\vec{u}) = \begin{bmatrix} 2a\vec{u}_1 \\ a\vec{u}_2 \end{bmatrix} = S\begin{bmatrix} a\vec{u}_1 \\ a\vec{u}_2 \end{bmatrix} = a\begin{bmatrix} 2u_1 \\ u_2 \end{bmatrix}$

  $= aS(\vec{u})$

- and so S satisfies all the properties of a linear transformation

- By inpection, $T(\vec{u}+\vec{v}) = \begin{bmatrix} u_1+v_1 \\ u_2+v_2+4 \end{bmatrix}$,

  $T(\vec{u}) + T(\vec{v}) = \begin{bmatrix} u_1+v_1 \\ u_2+v_2+8 \end{bmatrix}$

- Use $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as a counter-example

- $T(\vec{e}_1+\vec{e}_2) = T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$

  $T(\vec{e}_1) + T(\vec{e}_2) = \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix}$

- since at least one property of a linear transformation is violated, T cannot be a linear transformation.

Function Notation vs. Linear Transformation Notation

- $\underbrace{f : \mathbb{R} \to \mathbb{R}}_{\text{a function}}$ $\qquad$ $\underbrace{f(x)}_{f \text{ evaluated at } x}$

- $\underbrace{T : \mathbb{R}^n \to \mathbb{R}^m}_{\text{a transformation}}$ $\qquad$ $\underbrace{T(\vec{x})}_{T \text{ evaluated at } \vec{x}}$ $\qquad$ $\underbrace{T\vec{x}}_{\text{also } T \text{ evaluated at } \vec{x}}$

- "the function f" is valid,

  "the function $f(x)$" is invalid

  - $f(x)$ means "the output of f when x is the input"
  - $f(x)$ is a number, not a function

The "Look" of a Linear Transformation

- If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then
  - $T(\vec{0}) = 0$
  - T takes lines to lines (or points)
  - T takes parallel lines to parallel lines (or points)
  - T takes subspaces to subspaces

Linear Transformations and Proofs

- Start the proof w/

  "Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and let α be a scalar"

- Proof: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and let α be a scalar.

  By applying the definition of T, we see

  $T(\vec{x}+\vec{y}) =$ application(s) of definition

  $= T(\vec{x}) + T(\vec{y})$

  Similarly,

  $T(\alpha\vec{x}) =$ application(s) of definition $= \alpha T(\vec{x})$

  Since T satisfies the two properties of a linear transformation, T is a linear transformation. ■

- For a transformation that is not linear, pick one of the properties of linearity and a single example where that property fails

- Ex. Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\vec{x}) = \vec{x} + \vec{e}_1$.

  Show that T is not linear.

  - Proof: We will show that T does not distibute with respect to scalar multiplication.

    $T(2\vec{0}) = T(\vec{0}) = \vec{e}_1 \neq 2\vec{e}_1 = 2T(\vec{0})$

    Therefore, T cannot be a linear transformation. ■

Matrix Transformations

- Let $M = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$. For a vector $\vec{v} \in \mathbb{R}^2$,

  $M\vec{v}$ is another vector in $\mathbb{R}^2$.

- We can think of multiplication by M as a transformation on $\mathbb{R}^2$. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x) = M\vec{x}$
  - T is a matrix transformation
- All matrix transformations are linear
- " the matrix transformation given by $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ "
- We can specify a linear transformation using a matrix, but a matrix itself is not a linear transformation
- to specify a linear transformation using a matrix:
  - " the transformation T defined by $T(\vec{x}) = M\vec{x}$ "
  - " the transformation given by multiplication by M "
  - " the transformation induced by M "
  - " the matrix transformation given by M "

- "the linear transformation whose matrix is M"

Finding a Matrix for a Linear Transformation

- Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
- Since T inputs vectors w/ n coordinates and outputs vectors w/ m coordinates, any matrix for T must be $m \times n$
- The process of finding a matrix T:
    1. Create an $m \times n$ matrix of variables
    2. use known input-output pairs for T to set up a system of equations involving the unknown variables
    3. solve the variables

- Ex. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x+y \\ x \end{bmatrix}$.

  Find a matrix, M, for T.

    - Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

- We know $T\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}$, and $T\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$

- Since M is a matrix for T, we know

  $T\vec{x} = M\vec{x}$ for all $\vec{x}$, and so

$$M\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}a+b\\c+d\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}$$

and

$$M\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}b\\d\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$$

- The above gives us the system of equations

$$\begin{cases} a+b = 3 \\ c+d = 1 \\ b = 1 \\ d = 0 \end{cases}$$

- solving the above system tells us

$$M=\begin{bmatrix}a & b\\c & d\end{bmatrix}=\begin{bmatrix}2 & 1\\1 & 0\end{bmatrix}$$

Module 10

Composition

- Let $f: A \to B$ and $g: B \to C$. The composition of g and f, notated $g \circ f$, is the function $h: A \to C$ defined by $h(x) = g \circ f(x) = g(f(x))$
- We can understand complicated linear transformations by breaking them into the composition of simpler ones
  - i.e. define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\vec{x}) = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \vec{x}$
  - Notice that $T = S \circ R$ where

    R is the rotation counterclockwise by $45°$,

    S is the stretch in the $\vec{e}_1$ direction by a factor of 2
- Ex. Let $U: \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation given by

  $M = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & 0 \end{bmatrix}$, let $R: \mathbb{R}^2 \to \mathbb{R}^2$ be rotation

  counter clockwise by $45°$, and let $P: \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the x-axis. Write $\mathcal{U}$ as the composition (in some order) of R and P.

- $\mathcal{U}(\vec{e}_1) = \begin{bmatrix} \sqrt{2}/2 \\ 0 \end{bmatrix}$, $\mathcal{U}(\vec{e}_2) = \begin{bmatrix} -\sqrt{2}/2 \\ 0 \end{bmatrix}$

- $R \circ P(\vec{e}_1) = R(P(\vec{e}_1)) = R\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$

$R \circ P(\vec{e}_2) = R(P(\vec{e}_2)) = R\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- $P \circ R(\vec{e}_1) = P(R(\vec{e}_1)) = P\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ 0 \end{bmatrix}$

$P \circ R(\vec{e}_2) = P(R(\vec{e}_2)) = P\begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ 0 \end{bmatrix}$

Since $P \circ R$ agrees w/ $\mathcal{U}$ on the standard basis (i.e. $P \circ R$ and $\mathcal{U}$ output the same vectors when $\vec{e}_1$ and $\vec{e}_2$ are input), they must agree for all vectors. Therefore $\mathcal{U} = P \circ R$

Compositions and Matrix Products

- Let $A : \mathbb{R}^2 \to \mathbb{R}^2$ and $B : \mathbb{R}^2 \to \mathbb{R}^2$ be matrix transformations w/ matrices

$$M_A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \text{ and } M_B = \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix}$$

- Define $T = \mathcal{A} \circ \mathcal{B}$. Since $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, we know T has a matrix M_T, which is 2×2. We can find M_T by the usual methods
  - $T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathcal{A}\left(\mathcal{B}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \mathcal{A}\begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -5 \\ -4 \end{bmatrix}$

    $T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathcal{A}\left(\mathcal{B}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \mathcal{A}\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$
  - $M_T = \begin{bmatrix} -5 & -1 \\ -4 & 0 \end{bmatrix}$
- We could have used M_A and M_B to find M_T
  - By definition, $\mathcal{A}\vec{x} = M_A\vec{x}$ and $\mathcal{B}\vec{x} = M_B\vec{x}$
  - Since $\mathcal{A}$ and $\mathcal{B}$ are matrix transformations,

    $\mathcal{A}(\mathcal{B}\vec{x}) = M_A(M_B\vec{x})$
  - Since matrix multiplication is associative,

    $M_A(M_B\vec{x}) = (M_A M_B)\vec{x}$

- Thus $M_A M_B$ must be a matrix for $A \circ B = T$
- Computing the matrix product,

$$M_A M_B = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ -4 & 0 \end{bmatrix} = M_T$$

- If $P : \mathbb{R}^a \to \mathbb{R}^b$ and $Q : \mathbb{R}^c \to \mathbb{R}^a$ are matrix transformations w/ matrices M_P and M_Q, then $P \circ Q$ is a matrix transformation whose matrix is given by the matrix product $M_P M_Q$
- The order of matrix multiplication matters
  - because the order of function composition matters

Module 11

Range

- The range (or image) of a linear transformation $T: V \to W$ is the set of vectors that T can output. That is,

  $\text{range}(T) = \{\vec{y} \in W : \vec{y} = T\vec{x} \text{ for some } \vec{x} \in V\}$
- range : the set of all outputs
- range of a linear transformation is the image of the entire domain w.r.t. that linear transformation
- range of a linear transformation is always a subspace
  - Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\text{range}(T) \subseteq \mathbb{R}^m$ is a subspace
- For a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, the rank of T, denoted $\text{rank}(T)$, is the dimension of the range of T
- A rank 0 transformation must send all vectors to $\vec{0}$, a rank 1 transformation must send all vectors to a line, etc.

- Ex. Let P be the plane given by $x+y+z=0$, and let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be the projection onto P.

  Find range(T) and rank(T)

    - since T is a projection onto P, range$(T) \subseteq P$
    - because $T(\vec{p}) = \vec{p}$ for all $\vec{p} \in P$, $P \subseteq$ range(T)
    - and so range$(T) = P$
    - since P is a plane,

      $\dim(P) = 2 = \dim(\text{range}(T)) = \text{rank}(T)$

<u>Null Space</u>

- The null space (or kernel) of a linear transformation $T: V \to W$ is a set of vectors that get mapped to the zero vector under T. That is,

  $\text{null}(T) = \{\vec{x} \in V : T\vec{x} = 0\}$

- null space of a linear transformation is always a subspace
    - Let $T \in \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

      Then null(T) is a subspace.

- For a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, the nullity of T, denoted nullity(T), is the dimension of the null space of T.

- Ex. Let P be the plane given by $x+y+z=0$, and let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be the projection onto P. Find null (T) and nullity (T)
    - Since T is a projection onto P (and because P passes through $\vec{0}$), we know every normal vector for P will get sent to $\vec{0}$ when T is applied.
        - Besides $\vec{0}$ itself, these are the only vectors that get sent to $\vec{0}$
    - null (T) = {normal vectors} $\cup$ $\{\vec{0}\}$

$$= \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$$

    - since null(T) is a line, we know nullity (T) = 1

<u>Fundamental Subspaces of a Matrix</u>

- Associated w/ any matrix M are 3 fundamental subspaces:
    - the row space of M, denoted row(M), is the

span of the rows of M

- the column space of M, denoted $\operatorname{col}(M)$, is the span of the columns of M
- the null space of M, denoted $\operatorname{null}(M)$, is the set of solutions to $M\vec{x} = \vec{0}$

- Ex. Find the null space of $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$

  - to find the null space of M, we need to solve the homogeneous matrix equation $M\vec{x} = \vec{0}$.

    - $\operatorname{rref}(M) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

  - the complete solution expressed in vector form:

    - $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$

  - and so $\operatorname{null}(M) = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right\}$

- Ex. Let $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$. Find a basis for the row space and the column space of M.

  - For the column space, we need to pick a basis for $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -2 \end{bmatrix} \right\}$

  - putting these vectors as columns in a matrix and row reducing, we see

$$\operatorname{rref}\left(\begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

  - the first and second columns are the only pivot columns, and so the first and second original vectors form a maximal linearly independent subset. Thus,

$$\operatorname{col}(M) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right\} = \mathbb{R}^2$$

and a basis is $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right\}$

- For the row space, we need to pick a basis for span $\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$

- repeating a similar procedure, we see

$$\text{rref}\left(\begin{bmatrix} 1 & 2 \\ 2 & -2 \\ 5 & -2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and so $\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$ is linearly independent.

Therefore, $\text{row}(M) = \text{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$

and a basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$

- The operation of swapping rows and columns is called the transpose

- Let M be an $n \times m$ matrix defined by

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

The transpose of M, notated M^T, is the $m \times n$ matrix produced by swapping the rows and columns of M. That is

$$M^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}$$

- $\text{col}(M) = \text{row}(M^T)$, $\text{row}(M) = \text{col}(M^T)$
- For a matrix A, the dimension of the row space equals the dimension of the column space

Equations, Null Spaces, and Geometry

- Let $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$

  - the complete solution to $M\vec{x} = \vec{0}$ (i.e. the null space of M) can be expressed in vector form as $\vec{x} = t \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$

  - Similarly, the complete solution to $M\vec{x} = \vec{b}$ where $\vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ can be expressed in vector form as $\vec{x} = t \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

  - $\{$solutions to $M\vec{x} = \vec{b}\}$

    $= \{$solutions to $M\vec{x} = \vec{0}\} + \{\vec{p}\}$

    where $\vec{p} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

- The solution set to $M\vec{x} = \vec{b}$ is $\text{null}(M) + \{\vec{p}\}$
- We know the solution set to $M\vec{x} = \vec{b}$ is a line (which doesn't pass through the origin) and may therefore be written as a translated span $\text{span}\{\vec{d}\} + \{\vec{p}\}$
- Because $\vec{p} \in \text{span}\{\vec{d}\} + \{\vec{p}\}$, we call $\vec{p}$ a particular solution to $M\vec{x} + \vec{b}$
- We can show that for any matrix A, and any vector $\vec{b}$, the set of all solutions to $A\vec{x} = \vec{b}$ (provided there are any) can be expressed as $V + \{\vec{p}\}$ where V is a subspace and $\vec{p}$ is a particular solution
    - $V = \text{null}(A)$
- Let A be a matrix, $\vec{b}$ be a vector, and let $\vec{p}$ be a particular solution to $A\vec{x} = \vec{b}$. Then, the set of all solutions to $A\vec{x} = \vec{b}$ is $\text{null}(A) + \{\vec{p}\}$
- To write a complete solution to $A\vec{x} = \vec{b}$, all we need is

the null space of A and a

particular solution to $A\vec{x} + \vec{b}$

- Let $P \subseteq \mathbb{R}^3$ be the plane w/ equation $x + 2y + 2z = 0$
    - Matrix equation: $\begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{0}$
    - Normal form: $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{0}$
    - $P = \text{null}(\begin{bmatrix} 1 & 2 & 2 \end{bmatrix})$
    - every non-zero vector in $\text{row}(\begin{bmatrix} 1 & 2 & 2 \end{bmatrix})$ is a normal vector for P
    - $\text{null}(\begin{bmatrix} 1 & 2 & 2 \end{bmatrix})$ is orthogonal to $\text{row}(\begin{bmatrix} 1 & 2 & 2 \end{bmatrix})$
- Let M be a matrix and let $\vec{r}_1, \ldots, \vec{r}_n$ be the rows of M. By definition, $M\vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vdots \\ \vec{r}_n \cdot \vec{x} \end{bmatrix}$,

and so solutions to $M\vec{x} = \vec{0}$ are the vectors which are orthogonal to every row of M

- row(M) consists of all vectors orthogonal to everything in null(M)

- Ex. Let $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix}$.

Find the set of all vectors orthogonal to both $\vec{a}$ and $\vec{b}$.

- Let $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$ be the matrix whose rows are $\vec{a}$ and $\vec{b}$

- Since null(M) consists of all vectors orthogonal to row(M), the set we are looking for is null(M).

- By row reduction, $\text{null}(M) = \text{span}\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right\}$

Transformations and Matrices

- Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let M be its corresponding matrix
  - T is a function that inputs and outputs vectors
  - M is a box of numbers, which has no meaning by itself
  - the expression $M\vec{x}$ doesn't make sense, because $\vec{x}$ is a vector
  - the expression $M[\vec{x}]_\varepsilon$ makes sense since $[\vec{x}]_\varepsilon$ is a list of numbers
  - $T(x) \neq M[\vec{x}]_\varepsilon$ because the left side is a vector and the right side is a list of numbers
- The relationship between a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ and its matrix M is

  $[T(\vec{x})]_\varepsilon = M[\vec{x}]_\varepsilon$
- If we have a matrix M, by picking a basis, we can define a linear transformation by:
  - first taking the input vector and rewriting it

in the basis

- next multiplying it by the matrix
- and finally taking the list of numbers and using them as coefficients for a linear combination involving the basis vectors

- Let M be an $n \times m$ matrix. We say M induces a linear transformation $T_M : \mathbb{R}^m \to \mathbb{R}^n$ defined by $[T_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}}$, where $\mathcal{E}$ is the standard basis for $\mathbb{R}^m$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^n$

- We can write things like "$M\vec{x}$" when we're only considering a single basis

- Ex. Let T be the transformation induced by the matrix $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$, and let $\vec{v} = 3\vec{e}_1 - 3\vec{e}_3$.

  Compute $T(\vec{v})$.

  - Since T is induced by $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$,

$$[T_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix} [\vec{v}]_{\mathcal{E}}$$

- since $\vec{v} = 3\vec{e}_1 - 3\vec{e}_3$, $[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}$

- Therefore,

$$[T_M \vec{v}]_{\mathcal{E}'} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} -12 \\ 12 \end{bmatrix}$$

- In other words, $T(\vec{v}) = \begin{bmatrix} -12 \\ 12 \end{bmatrix}_{\mathcal{E}'} = -12\vec{e}_1 + 12\vec{e}_2$

- Using induced transformations, we can extend linear-transformation definitions to matrix definitions
  - We can define the rank and nullity of a matrix
- Let M be a matrix. The rank of M, denoted $\text{rank}(M)$, is the rank of the linear transformation induced by M

- Let M be a matrix. The nullity of M, denoted nullity(M), is the nullity of the linear transformation induced by M

Range vs. Column Space & Null Space vs. Null Space

- Let $M = [C_1 \; C_2 \; \cdots \; C_m]$ be a $m \times m$ matrix w/ columns $C_1, \ldots, C_m$, and let T be the transformation induced by M.
  - The column space of M is the set of all linear combinations of the columns of M
  - Since C_i is a list of numbers, $[C_i]_{\mathcal{E}}$ is a true vector, and $\text{col}(M) = \text{span}\{[C_1]_{\mathcal{E}}, [C_2]_{\mathcal{E}}, \ldots, [C_m]_{\mathcal{E}}\}$
  - By the definition of matrix multiplication,

$$M \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = M[\vec{e}_1]_{\mathcal{E}} = C_1$$

and in general $M[\vec{e}_i]_{\mathcal{E}} = C_i$

- By the definition of induced transformation,

  $[T(\vec{e}_i)]_\varepsilon = M[\vec{e}_i]_\varepsilon = C_i$, and so

  $T(\vec{e}_i) = [C_i]_\varepsilon$

- Every input to T can be written as a linear combination of $\vec{e}_i$'s (because ε is a basis), and so, because T is linear, every output of T can be written as a linear combination of $[C_i]_\varepsilon$'s.

- range(T) = col(M)

- When trying to answer a question about the range of a linear transformation, we could think about the column space of its matrix instead (or vice versa)

- Ex. Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x - z \\ 4x - 2z \end{bmatrix}.$$ Find range(T) and rank(T)

- Let M be a matrix for T.

- range(T) = col(M)

- rank(T) = dim(range(T)) = dim(col(M))

- By inspection, $M = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \end{bmatrix}$

- By inspection, $\left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$ is a basis for col(M)

  and col(M) is one dimensional.

- Therefore, range(T) = span $\left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$ and

  rank(T) = 1

- Let M be a matrix. The rank of M is equal to the number of pivots in rref(M)
- If T is a linear transformation and M is a corresponding matrix, range(T) = col(M), and answering questions about M answers questions about T
- Let T be a linear transformation and let M be a matrix for T. Then nullity(T) is equal to the number of free variable columns in rref(M)
- For a matrix A, we have rank(A) = rank(A^T)
  - dim(col(A)) = dim(col(A^T))

The Rank-Nullity Theorem

- For a matrix A,

  rank(A) + nullity(A) = # of columns in A

- consider the matrix $M = [1 \ 2 \ 2]$

  - null(M) is the plane w/ equation $x+2y+2z=0$ and therefore is two-dimensional

  - Since M has 3 columns, rank$(M) = 1$, and so $\dim(\text{col}(M)) = \dim(\text{row}(M)) = 1$, which means that M has a one-dimensional set of normal vectors (if we include $\vec{0}$)

- Let $P \subseteq \mathbb{R}^4$ be the plane in $\mathbb{R}^4$ given in vector form by

$$\vec{x} = t \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

  - since the matrix $A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ -1 & 1 & -1 & 1 \end{bmatrix}$ is rank 2, and so has nullity 2. Therefore, there exists two linearly independent normal directions for P.

- Let T be a linear transformation. Then

  rank (T) + nullity (T) = dim (domain of T)

Module 12

Intro

- Functions can be divided into 2 categories: those can be undone and those cannot
  - invertible function - a function that can be undone

Invertible Functions

- The simplest function is the identity function
- Let X be a set. The identity function w/ domain and codomain X, notated $id: X \to X$, is the function satisfying $id(x) = x$ for all $x \in X$
- The identity function does nothing to its input
- A function is invertible if there exists an inverse function that when composed w/ the original function produces the identity function and vice versa
- Let $f: X \to Y$ be a function. f is invertible if there exists a function $g: Y \to X$ so that

$f \circ g = id$ and $g \circ f = id$. In this case,

we call g an inverse of f and write

$f^{-1} = g$

- for a function to be invertible, it must be one-to-one
- Let $f: X \to Y$ be a function.

  f is one-to-one (or injective) if

  distinct inputs to f produce distinct outputs.

  That is $f(x) = f(y) \Rightarrow x = y$
- Whenever a function f is one-to-one, there exists

  a function g so that $g \circ f = id$
- To declare that f is invertible, we also need

  $f \circ g = id$. To ensure this, we need

  f to be onto
- Let $f: X \to Y$ be a function.

  f is onto (or surjective) if every point in the

  codomain gets mapped to. That is $range(f) = Y$
- Every invertible function is both

  one-to-one and onto, and vice versa

Invertibility and Linear Transformations

- $T: \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if it is one-to-one and onto
- if T is one-to-one, the solution to $T(\vec{x}) = \vec{b}$ is always unique.
- Since the set of all solutions to $T(\vec{x}) = \vec{b}$ can be expressed as $\text{null}(T) + \{\vec{p}\}$, T is one-to-one if and only if $\text{nullity}(T) = 0$
- If T is onto, $\text{range}(T) = \mathbb{R}^m$ and so $\text{rank}(T) = m$
- Suppose T is one-to-one and onto, by the rank-nullity theorem,

  $\text{rank}(T) + \text{nullity}(T) = m + 0 = m = n$

  $= \dim(\text{domain of } T)$

  and so T has the same domain and codomain
- Properties equivalent to invertibility of a linear transformation:
  - $T: \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if:

- nullity$(T) = 0$ and rank$(T) = m$
- $m = n$ and nullity$(T) = 0$
- $m = n$ and rank$(T) = m$

- Ex. Let $P: \mathbb{R}^2 \to \mathbb{R}^2$ be the projection onto the x-axis and let $R: \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counterclockwise by $15°$. Classify each of P and R as invertible or not.
  - $P(\vec{e}_2) = P(2\vec{e}_2) = 0$, therefore P is not one-to-one and so not invertible.
  - Let $Q: \mathbb{R}^2 \to \mathbb{R}^2$ be rotation clockwise by $15°$. R and Q will undo each other, i.e. $R \circ Q = id$ and $Q \circ R = id$. Therefore, Q is an inverse of R, and so R is invertible

- Let T be a linear transformation. Then T^{-1} is also a linear transformation.

<u>Invertibility and Matrices</u>

- an identity matrix is a square matrix w/ ones on the diagonal and zeros everywhere else

- The $n \times n$ identity matrix is denoted as $I_{n \times n}$,
  or just I when its size is implied

- the inverse of a matrix A is a matrix B such that
  $AB = I$ and $BA = I$
  - In this case, B is called the inverse of A and is notated A^{-1}

- Ex. Determine whether the matrices $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $B = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ are inverses of each other

  - $AB = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & 5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

  - $BA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

  - Therefore, A and B are inverses of each other.

- Ex. Determine whether the matrices $A = \begin{bmatrix} 2 & 5 & 0 \\ -3 & -7 & 0 \end{bmatrix}$ and

$B = \begin{bmatrix} -7 & -5 \\ 3 & 2 \\ 1 & 1 \end{bmatrix}$ are inverses of each other.

- $AB = \begin{bmatrix} 2 & 5 & 0 \\ -3 & -7 & 0 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$

- $BA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 & 0 \\ -3 & -7 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -2 & 0 \end{bmatrix} \neq I$

- Therefore, A and B are not inverses of each other

- Facts about invertible matrices:
  - An $n \times m$ matrix A is invertible if and only if
    - $\text{nullity}(A) = 0$ and $\text{rank}(A) = n$
  - An $n \times n$ matrix A is invertible if and only if
    - $\text{nullity}(A) = 0$
    - $\text{rank}(A) = n$

Matix Algebra

- To solve the matrix $A\vec{x} = \vec{b}$, we need to eliminate A from the left side. This could be accomplished by using an inverse

- Suppose A is invertible, then

$$A\vec{x} = \vec{b} \Rightarrow A^{-1}A\vec{x} = A^{-1}\vec{b} \Rightarrow \vec{x} = A^{-1}\vec{b}$$

- Ex. Use the fact that $\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ to solve the system $\begin{cases} 2x + 5y = 2 \\ -3x - 7y = 1 \end{cases}$

- $\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

- Multiplying both sides by $\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}$,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -19 \\ 8 \end{bmatrix}$$

- Since order of matrix multiplication matters,

  $A\vec{x} = \vec{b}$ does not imply $A^{-1}A\vec{x} = \vec{b}A^{-1}$

  - If $\vec{b}$ is a column vector, $\vec{b}A^{-1}$ is undefined

Finding a Matrix Inverse

- $\vec{x} = A^{-1}\vec{b}$ is the solution to $A\vec{x} = \vec{b}$

- by picking different $\vec{b}$'s and solving for x,

  we can find A^{-1}

- Ex. Let $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$. Find A^{-1}

  - $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

  - Using row reduction,

$$A\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ has solution } \vec{x} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ has solution } \vec{x} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$$

  - Therefore,

$$A^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}$$

$$A^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$$

and so $A^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$

Elementary Matrices

- A strategy for finding an inverse matrix is to break a matrix into simpler ones whose inverses we can just write down
- A matrix is called an elementary matrix if it is an identity matrix w/ a single elementary row operation applied
- Examples:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  - i.e. multiply the last row by -5, add 7 times the last row to the first, swap the first two rows
- Elementary row operations:
  - multiply a row by a nonzero constant

- add a multiple of one row to another
- swap two rows

- Every elementary matrix is invertible

- Ex. Find the inverse of $E = \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

  - Since E corresponds to "add 7 times the last row to the first", E^{-1} must correspond to "subtract 7 times the last row from the first"

  - Therefore, $E^{-1} = \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

<u>Elementary Matrices and Inverses</u>

- M is invertible if and only if $\text{rref}(M) = I$
  - i.e. M is invertible if there is a sequence of elementary row operations that turn M into I
  - Each one of these row operations can be represented by an elementary matrix

- A matrix M is invertible if and only if there are elementary matrices $E_1, \dots, E_k$ so that $E_k \dots E_2 E_1 M = I$
  - Let $Q = E_k \dots E_2 E_1$, $QM = I$
  - If we can argue that $MQ = I$, then $Q = M^{-1}$
- If A is a square matrix and $AB = I$ for some matrix B, then $BA = I$
- We can find elementary matrices that turn the matrix into the identity matrix and multiply those elementary matrices together to find the inverse
- Ex. Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Find A^{-1} using elementary matrices

  - We can reduce A w/ the following steps:

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 0 \\ 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- the elementary matrices corresponding to these steps are:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

- We now have $E_3 E_2 E_1 A = I$, and so

$$A^{-1} = E_3 E_2 E_1$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1/4 & 0 \\ 0 & 1/4 & 1 \end{bmatrix}$$

Decomposition into Elementary Matrices

- If A is an invertible matrix, then the double inverse of A (i.e. $(A^{-1})^{-1}$) is A itself
- Suppose M is an invertible matrix. Then, there exists a sequence of elementary matrices $E_1, \dots, E_k$ so that $E_k \cdots E_2 E_1 M = I$ and $M^{-1} = E_k \cdots E_2 E_1$.

  Therefore, $M = (M^{-1})^{-1} = (E_k \cdots E_2 E_1)^{-1}$.

  We see that $(E_1^{-1} E_2^{-1} \cdots E_k^{-1}) E_k \cdots E_2 E_1 = I$ and $E_k \cdots E_2 E_1 (E_1^{-1} E_2^{-1} \cdots E_k^{-1}) = I$.

  And so $M = (E_k \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$
    - Notice that the order of matrix multiplication reversed
    - Each E_i^{-1} is also an elementary matrix
- A matrix M is invertible if and only if it can be written as the product of elementary matrices

Module 13

Change of Basis

- Given a basis $\mathcal{A}$ for $\mathbb{R}^n$, every vector $\vec{x} \in \mathbb{R}^n$ uniquely corresponds to the list of numbers $[\vec{x}]_{\mathcal{A}}$ (i.e. its coordinates w.r.t. $\mathcal{A}$)
- The operation of writing a vector in a basis is invertible
- $\vec{x} \in \mathbb{R}^n \begin{cases} \text{represent in } \mathcal{A} \text{ basis: } [\vec{x}]_{\mathcal{A}} \\ \text{represent in } \mathcal{B} \text{ basis: } [\vec{x}]_{\mathcal{B}} \end{cases}$
- There must be a function that converts between $[\vec{x}]_{\mathcal{A}}$ and $[\vec{x}]_{\mathcal{B}}$
    - The function inputs the list of numbers $[\vec{x}]_{\mathcal{A}}$, use those numbers as coefficients of the $\mathcal{A}$ basis to get the true vector $\vec{x}$, and then find the coordinates of that vector w.r.t. the $\mathcal{B}$ basis
- Ex. Let $\mathcal{A} = \{\vec{a}_1, \vec{a}_2\}$ where $a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}}$, $a_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}_{\mathcal{E}}$

and let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{E}}$, $b_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}_{\mathcal{E}}$

be bases for $\mathbb{R}^2$. Given that $[\vec{x}]_{\mathcal{A}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, find $[\vec{x}]_B$

- By definition,

$$\vec{x} = 2\vec{a}_1 - 3\vec{a}_2 = 2(\vec{e}_1 + \vec{e}_2) - 3(\vec{e}_1 - \vec{e}_2)$$

$$= -\vec{e}_1 + 5\vec{e}_2$$

- We need to write $\vec{x}$ as a linear combination of

$\vec{b}_1 = 2\vec{e}_1 + \vec{e}_2$ and $\vec{b}_2 = 5\vec{e}_1 + 3\vec{e}_2$.

$$\vec{x} = -\vec{e}_1 + 5\vec{e}_2 = \alpha(2\vec{e}_1 + \vec{e}_2) + \beta(5\vec{e}_1 + 3\vec{e}_2)$$

$$= (2\alpha + 5\beta)\vec{e}_1 + (\alpha + 3\beta)\vec{e}_2$$

- Equating the coefficients of $\vec{e}_1$ and $\vec{e}_2$,

$$\begin{cases} 2\alpha + 5\beta = -1 \\ \alpha + 3\beta = 5 \end{cases}$$

which has a unique solution $(\alpha, \beta) = (-28, 11)$.

- We conclude $[\vec{x}]_B = \begin{bmatrix} -28 \\ 11 \end{bmatrix}$

- Let $\mathcal{A}$ and B be bases for $\mathbb{R}^n$. The matrix M is called a change of basis matrix (which converts from $\mathcal{A}$ to $\mathcal{B}$) if for all $\vec{x} \in \mathbb{R}^n$

$M[\vec{x}]_{\mathcal{A}} = [\vec{x}]_{\mathcal{B}}$

- $[\mathcal{B} \leftarrow \mathcal{A}]$ stands for the change of basis matrix converting from $\mathcal{A}$ to $\mathcal{B}$
- $M = [\mathcal{B} \leftarrow \mathcal{A}]$

- Ex. Let $\mathcal{A} = \{\vec{a}_1, \vec{a}_2\}$ where $\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}_{\mathcal{E}}$

and let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{E}}$, $\vec{b}_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}_{\mathcal{E}}$

be bases for $\mathbb{R}^2$. Find the change of basis matrix, $[\mathcal{B} \leftarrow \mathcal{A}]$

- $[\mathcal{B} \leftarrow \mathcal{A}]$ is a 2×2 matrix such that

  $[\mathcal{B} \leftarrow \mathcal{A}][\vec{a}_1]_{\mathcal{A}} = [\vec{a}_1]_{\mathcal{B}}$ and

  $[\mathcal{B} \leftarrow \mathcal{A}][\vec{a}_2]_{\mathcal{A}} = [\vec{a}_2]_{\mathcal{B}}$
- We need to compute $[\vec{a}_1]_{\mathcal{B}}$ and $[\vec{a}_2]_{\mathcal{B}}$
- Repeating the procedure from the previous example, we find

$$[\vec{a}_1]_B = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ and } [\vec{a}_2]_B = \begin{bmatrix} 8 \\ -3 \end{bmatrix}$$

- and so $[B \leftarrow A] = \begin{bmatrix} -2 & 8 \\ 1 & -3 \end{bmatrix}$

-

$$\begin{tikzcd}
& x \in \mathbb{R}^n \arrow[dl, leftrightarrow, "\text{represent in } A \text{ basis}"'] \arrow[dr, leftrightarrow, "\text{represent in } B \text{ basis}"] & \\
{[\vec{x}]_A} \arrow[rr, shift right, "{[B \leftarrow A]}"'] & & {[\vec{x}]_B} \arrow[ll, shift right, "{[A \leftarrow B]}"']
\end{tikzcd}$$

- Suppose we have another basis C. We can obtain $[C \leftarrow A]$ by multiplying $[B \leftarrow A]$ on the left by $[C \leftarrow B]$
    - $[C \leftarrow A] = [C \leftarrow B][B \leftarrow A]$
- The arrow is backwards because we multiply vectors on the left by matrices
    - $$[\vec{x}]_C = [C \leftarrow A][\vec{x}]_A$$
      $$= [C \leftarrow B][B \leftarrow A][\vec{x}]_A$$

Change of Basis Matrix in Detail

- Let $M = [B \leftarrow A]$
    - Since we can change vectors back from B to A,

$M^{-1} = [\mathcal{A} \leftarrow \mathcal{B}]$

- $M^{-1}M = [\mathcal{A} \leftarrow \mathcal{B}][\mathcal{B} \leftarrow \mathcal{A}] = [\mathcal{A} \leftarrow \mathcal{A}] = I$

  $MM^{-1} = [\mathcal{B} \leftarrow \mathcal{A}][\mathcal{A} \leftarrow \mathcal{B}] = [\mathcal{B} \leftarrow \mathcal{B}] = I$

- An $n \times n$ matrix is invertible if and only if it is a change of basis matrix

- Let $\mathcal{A} = \{\vec{a}_1, \dots, \vec{a}_n\}$ be a basis for $\mathbb{R}^n$.

  It is always the case that $[\vec{a}_i]_{\mathcal{A}} = \begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$

  has 1 in its ith position and zeros elsewhere.

- Let $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ be another basis for $\mathbb{R}^n$ and define the matrix $M = [[\vec{a}_1]_{\mathcal{B}}\ [\vec{a}_2]_{\mathcal{B}} \cdots [\vec{a}_n]_{\mathcal{B}}]$ to be the matrix w/ columns $[\vec{a}_1]_{\mathcal{B}}, \dots, [\vec{a}_n]_{\mathcal{B}}$.

- Since multiplying a matrix by $[\vec{a}_i]_{\mathcal{A}}$ will pick out the ith column, we have that

  $M[\vec{a}_i]_{\mathcal{A}} = [\vec{a}_i]_{\mathcal{B}}$

- In other words, $M = [\mathcal{B} \leftarrow \mathcal{A}]$

Transformations and Bases

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let $\mathcal{B}$ be a basis for $\mathbb{R}^n$.

  The matrix for T with respect to $\mathcal{B}$, notated $[T]_{\mathcal{B}}$, is the $n \times n$ matrix satisfying $[T\vec{x}]_{\mathcal{B}} = [T]_{\mathcal{B}} [\vec{x}]_{\mathcal{B}}$.

  - We say the matrix $[T]_{\mathcal{B}}$ is the representation of T in the $\mathcal{B}$ basis

- Ex. Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}_{\mathcal{E}}$, $\vec{b}_2 = \begin{bmatrix} 5 \\ -7 \end{bmatrix}_{\mathcal{E}}$

  be a basis for $\mathbb{R}^2$ and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that stretches in the $\vec{e}_1$ direction by the factor of 2. Find $[T]_{\mathcal{E}}$ and $[T]_{\mathcal{B}}$

  - Since $T\vec{e}_1 = 2\vec{e}_1$ and $T\vec{e}_2 = \vec{e}_2$, we know

$$[T]_{\mathcal{E}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad [T]_{\mathcal{E}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and so

$$[T]_{\mathcal{E}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

- We can find $[T]_{\mathcal{B}}$ in two ways:

    1. directly from definition

    2. using change of basis matrices

1. To find $[T]_{\mathcal{B}}$, we need to figure out what T does to $\vec{b}_1$ and $\vec{b}_2$. Since T is described in terms of $\vec{e}_1$ and $\vec{e}_2$, we will express $\vec{e}_1$ and $\vec{e}_2$ in the $\mathcal{B}$ basis first

$$[\vec{e}_1]_{\mathcal{B}} = [\mathcal{B} \leftarrow \mathcal{E}] \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}$$

$$[\vec{e}_2]_{\mathcal{B}} = [\mathcal{B} \leftarrow \mathcal{E}] \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$$

- Now we know

$$[T]_{\mathcal{B}}[\vec{e}_1]_{\mathcal{B}} = [T\vec{e}_1]_{\mathcal{B}} = [2\vec{e}_1]_{\mathcal{B}} = \begin{bmatrix} -14 \\ 6 \end{bmatrix}$$

$$[T]_{\mathcal{B}}[\vec{e}_2]_{\mathcal{B}} = [T\vec{e}_2]_{\mathcal{B}} = [\vec{e}_2]_{\mathcal{B}} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$$

- Using what we know, we can solve for the entries of $[T]_{\mathcal{B}}$

$$[T]_{\mathcal{B}} = \begin{bmatrix} -13 & -35 \\ 6 & 16 \end{bmatrix}$$

2. We already know $[T]_{\mathcal{E}}$, and so

$$[T]_{\mathcal{B}} = [\mathcal{B} \leftarrow \mathcal{E}][T]_{\mathcal{E}}[\mathcal{E} \leftarrow \mathcal{B}]$$

- Further, we know $[\mathcal{E} \leftarrow \mathcal{B}] = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$

and $[\mathcal{B} \leftarrow \mathcal{E}] = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$

- Putting everything together,

$$[T]_{\mathcal{B}} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} -13 & -35 \\ 6 & 16 \end{bmatrix}$$

Similar Matrices

- Some bases are better than others to represent a linear transformation

- Ex. Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}_{\mathcal{E}}$, $\vec{b}_2 = \begin{bmatrix} 5 \\ -7 \end{bmatrix}_{\mathcal{E}}$ be a basis for $\mathbb{R}^2$ and let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that stretches in the $b_1 = 2\vec{e}_1 - 3\vec{e}_2$ direction by a factor of 2 and reflects vectors in the $\vec{b}_2 = 5\vec{e}_1 - 7\vec{e}_2$ direction. Find $[S]_{\mathcal{E}}$ and $[S]_{\mathcal{B}}$.
  - We know that $S\vec{b}_1 = 2\vec{b}_1$, $S\vec{b}_2 = -\vec{b}_2$.

    Therefore, $[S]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

  - Use change of basis matrices to find $[S]_{\mathcal{E}}$:

    $[S]_{\mathcal{E}} = [\mathcal{E} \leftarrow \mathcal{B}][S]_{\mathcal{B}}[\mathcal{B} \leftarrow \mathcal{E}]$, and that

    $[\mathcal{E} \leftarrow \mathcal{B}] = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$, $[\mathcal{B} \leftarrow \mathcal{E}] = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$

  - Therefore,

$$[S]_{\mathcal{E}} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -43 & -30 \\ 63 & 44 \end{bmatrix}$$

- In the example above, $[S]_{\mathcal{B}}$ and $[S]_{\mathcal{E}}$ are related to each other
    - $[S]_{\mathcal{B}} = [\mathcal{B} \leftarrow \mathcal{E}][S]_{\mathcal{E}}[\mathcal{E} \leftarrow \mathcal{B}]$
    - We call these matrices similar
- The matrices A and B are called similar matrices, denoted $A \sim B$, if A and B represent the same linear transformation but in possibly different bases.
- $A \sim B$ if there is an invertible matrix X so that $A = XBX^{-1}$
    - The X is always a change-of-basis matrix

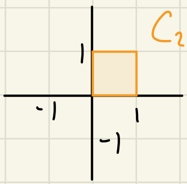
Module 14

Volumes

- We use the term volume to also mean area where appropriate
- the unit n-cube is the n-dimensional cubes with sides given by the standard basis vectors and lower-left corner located at the origin. That is

$$C_n = \left\{ \vec{x} \in \mathbb{R}^n : \vec{x} = \sum_{i=1}^{n} \alpha_i \vec{e}_i \text{ for some } \alpha_1, \dots, \alpha_n \in [0,1] \right\} = [0,1]^n$$

- C_2 is the unit square w/ lower-left corner at the origin



- C_n always has volume 1
- By analyzing the image of C_n under linear transformation,

We can see how much a given transformation changes volume.

- Ex. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x-y \\ x+\frac{1}{2}y \end{bmatrix}$.

Find the volume of $T(C_2)$

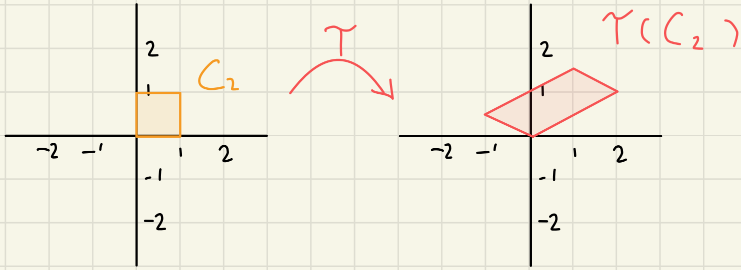
- C_2 has sides $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

- Applying the linear transformation T to $\vec{e}_1, \vec{e}_2$,

$$T(\vec{e}_1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad T(\vec{e}_2) = \begin{bmatrix} -1 \\ 1/2 \end{bmatrix}$$

- plotting $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1/2 \end{bmatrix}$, we see

$T(C_2)$ is a parallelogram w/ base $\sqrt{5}$ and height $\frac{2\sqrt{5}}{5}$



- Therefore, the volume of $T(C_2)$ is 2

- Let $vol(X)$ stand for the volume of the set X.

Given a linear transformation $S: \mathbb{R}^n \to \mathbb{R}^n$,

we can define a number

$$\text{Vol Change}(S) = \frac{\text{Vol}(S(C_n))}{\text{Vol}(C_n)} = \frac{\text{Vol}(S(C_n))}{1}$$

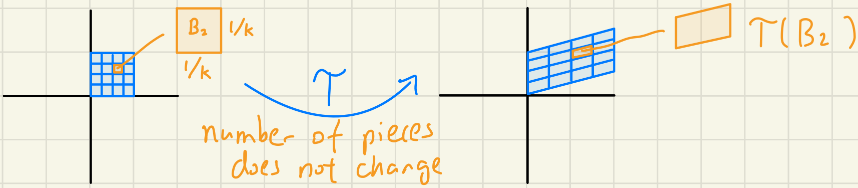
$$= \text{Vol}(S(C_n))$$

- Because S is a linear transformation, Vol Change(S) describes how S changes the volume of any figure

- Theorem. Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let $X \subseteq \mathbb{R}^n$ be a subset w/ volume α.
Then the volume of $T(X)$ is $\alpha \cdot$ Vol Change(T)

- Theorem. Suppose $T: \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, $X \subseteq \mathbb{R}^n$ is a subset, and the volume of $T(X)$ is α. Then for any $\vec{p} \in \mathbb{R}^n$, the volume of $T(X + \{\vec{p}\})$ is α.
  - translations do not change volume

- Theorem. Fix k and let B_n and C_n scaled to have side lengths $\frac{1}{k}$ and let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then Vol Change $(T) = \dfrac{\mathrm{Vol}(T(B_n))}{\mathrm{Vol}(B_n)}$



   - There are k^n copies of B_n in C_n and k^n copies $T(B_n)$ in $T(C_n)$. Thus,

$$\text{Vol Change}(T) = \frac{\mathrm{Vol}(T(C_n))}{\mathrm{Vol}(C_n)} = \frac{k^n \mathrm{Vol}(T(B_n))}{k^n \mathrm{Vol}(B_n)}$$

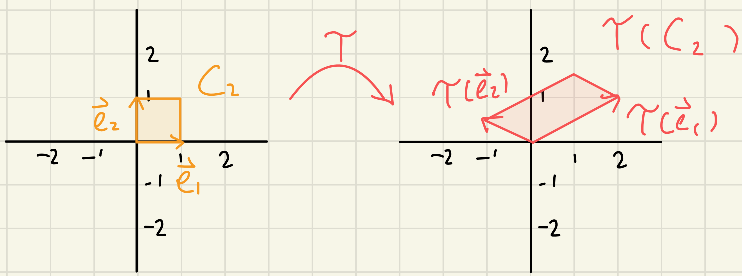
$$= \frac{\mathrm{Vol}(T(B_n))}{\mathrm{Vol}(B_n)}$$

The Determinant

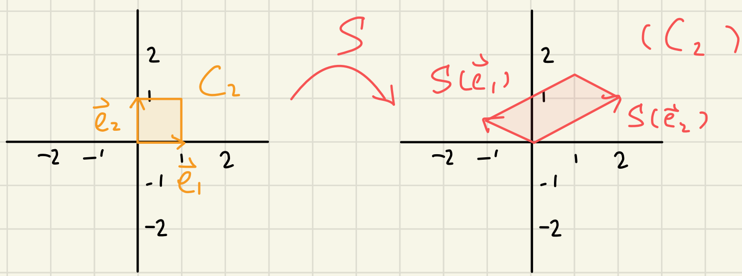
- The determinant of a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$, denoted $\det(T)$ or $|T|$, is the oriented volume of the image of the unit n-cube
   - The determinant of a square matrix is the determinant of its induced transformation

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. We say T is orientation preserving if the ordered basis $\{T(\vec{e}_1), \dots, T(\vec{e}_n)\}$ is positively oriented and we say T is orientation reversing if the ordered basis $\{T(\vec{e}_1), \dots, T(\vec{e}_n)\}$ is negatively oriented.
    - If $\{T(\vec{e}_1), \dots, T(\vec{e}_n)\}$ is not a basis for $\mathbb{R}^n$, then T is neither orientation preserving nor orientation reversing

Orientation Preserving



Orientation Reversing



- For an arbitrary linear transformation $Q : \mathbb{R}^n \to \mathbb{R}^n$

and a set $X \subseteq \mathbb{R}^n$, we define the

oriented volume of $Q(X)$ to be

$+\operatorname{Vol} Q(X)$ if Q is orientation preserving and

$-\operatorname{Vol} Q(X)$ if Q is orientation reversing

- Ex. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x+y \\ -x+\frac{1}{2}y \end{bmatrix}. \text{ Find } \det(T).$$

  - From the previous example, we computed

    $\operatorname{Vol} T(C_2) = 2$

  - Since T is orientation preserving, $\det(T) = 2$

- Ex. Let $S: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x+y \\ x+y \end{bmatrix}. \text{ Find } \det(S)$$

  - By drawing a picture, we see that

    $S(C_2)$ is a square and $\operatorname{Vol} S(C_2) = 2$

  - Since $S(\vec{e}_1) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $S(\vec{e}_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

    form a negatively oriented basis,

S is orientation reversing.

- Therefore, $\det(S) = -\mathrm{Vol}\, S(C_2) = -2$

- Ex. Let $P : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the line w/ equation $x + 2y = 4$. Find $\det(P)$.
  - Because P projects everything to a line, we know $P(C_2)$ must be a line segment and therefore has volume 0
  - Thus $\det(P) = 0$

Determinants and Composition

- If a linear transformation T changes volume by a factor of α and S changes volume by a factor of β, then $S \circ T$ changes volume by a factor of $\beta\alpha$
- Determinants are multiplicative
- Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^n$ be linear transformations. Then

  $\det(S \circ T) = \det(S)\det(T)$

Determinants of Matrices

- The determinant of a matrix is defined as the determinant of its induced transformation
- Theorem. Let A and B be $n \times n$ matrices. Then $\det(AB) = \det(A)\det(B)$
- Theorem (Volume Theorem I). For a square matrix M, $\det(M)$ is the oriented volume of the parallelepiped given by the column vectors
    - Parallelepiped - the n-dimensional analog of a parallelogram
- Three types of elementary matrices
- Multiply a row by a non-zero constant
    - Let E_m be such an elementary matrix
    - Scaling one row of I is equivalent to scaling one column of I
    - Columns of E_m specify a parallelepiped that is scaled by α in one direction
    - i.e. if

$$E_m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \text{ then}$$

$$\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} \to \{\vec{e}_1, \vec{e}_2, \alpha\vec{e}_3\}$$

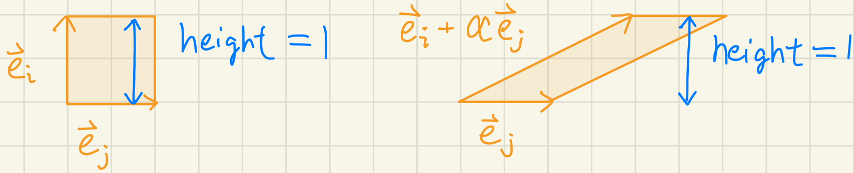
- Thus $\det(E_m) = \alpha$

- Swap two rows
    - Let E_s be such an elementary matrix
    - Swapping two rows of I is equivalent to swapping two columns of I
    - This reverses the orientation of the basis given by the columns.
    - i.e. if $E_s = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

      then $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} \to \{\vec{e}_2, \vec{e}_1, \vec{e}_3\}$
    - Thus $\det(E_s) = -1$

- Add a multiple of one row to another
    - Let E_a be such an elementary matrix

- The columns of E_α are the same as the columns of I except that one column where $\vec{e}_i$ is replaced by $\vec{e}_i + \alpha\vec{e}_j$
- Has the effect of shearing C_n in the $\vec{e}_j$ direction



- Since C_n is sheared in a direction parallel to one of its other sides, its volume is not changed. Thus $\det(E_\alpha) = 1$

- By decomposing a matrix into the product of elementary matrices, we can use the multiplicative property of the determinant to compute the determinant of an invertible matrix
- Ex. Use elementary matrices to find the determinant of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
  - We can row-reduce A

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- The elementary matrices that correspond to the above step:

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

- And so $E_3 E_2 E_1 A = I$

$$A = E_1^{-1} E_2^{-1} E_3^{-1} I = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

- Using the fact that the determinant is multiplicative, we get

$$\det(A) = \det\left(\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \right)$$

$$= \det\left(\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \right) \det\left(\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \right) \det\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \right)$$

$$= (1)(-2)(1) = -2$$

Determinants and Invertibility

- Let M be an $n \times n$ matrix that is not invertible.

Then, we must have $\text{nullity}(M) > 0$ and $\dim(\text{col}(M)) = \text{rank}(M) < n$.

- At least one line of vectors, $\text{null}(M)$ gets collapsed to $\vec{0}$
- The column space of M loses dimensions.
- The volume of the parallelepiped given by the columns of M must be 0, and so $\det(M) = 0$

- Theorem. A is invertible if and only if $\det(A) \neq 0$

- $\det(A^{-1}) = \dfrac{1}{\det(A)}$

<u>Determinants and Transposes</u>

- Theorem (Volume Theorem II). The determinant of a square matrix A is equal to the oriented volume of the parallelepiped given by the rows of A
  - $\det(A) = \det(A^T)$

Module 15

Intro

- If T stretches the vector $\vec{v}$, then T, in that direction, can be described by $\vec{v} \to \alpha\vec{v}$
    - Eigen directions - the "stretch" directions for a linear transformation
    - Eigenvectors - the vectors that are stretched
- Let X be a linear transformation or a matrix. An eigenvector for X is a non-zero vector that doesn't change directions when X is applied. That is, $\vec{v} \neq \vec{0}$ is an eigenvector for X if $X\vec{v} = \lambda\vec{v}$ for some scalar λ. We call λ the eigenvalue of X corresponding to the eigenvector $\vec{v}$
- Eigen - German for characteristic, representative, intrinsic
- Ex. Let $P: \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the line ℓ given by $y=x$. Find the eigenvectors and eigenvalues for P.

- We are looking for vectors $\vec{v} \neq \vec{0}$ s.t.

  $P\vec{v} = \lambda \vec{v}$ for some λ.

- Since $P(\ell) = \ell$, we know for any $\vec{v} \in \ell$,

  $P(\vec{v}) = 1\vec{v} = \vec{v}$

- Therefore, any non-zero multiple of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  is an eigenvector for P w/

  corresponding eigenvalue $= 1$

- By considering the null space of P, we see

  $$P\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- And so $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and all its non-zero multiples

  are eigenvectors of P w/

  corresponding eigenvalue 0.

Finding Eigenvectors

- Let M be a square matrix. The vector $\vec{v} \neq 0$ is an

  eigenvector of M if and only if there exists a

Scalar λ so that $M\vec{v} = \lambda\vec{v}$

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

- The operation $\vec{v} \to M\vec{v} - \lambda\vec{v}$ can be achieved by multiplying $\vec{v}$ by the single matrix $E_\lambda = M - \lambda I$
    - $E_\lambda \vec{v} = (M - \lambda I)\vec{v} = M\vec{v} - \lambda\vec{v} = \vec{0}$
    - $\vec{v}$ is a non-zero vector satisfying $\vec{v} \in null(E_\lambda)$

<u>Characteristic Polynomial</u>

- Because eigenvectors must be non-zero, we can only find an eigenvector if $null(E_\lambda) \neq \{\vec{0}\}$
    - We would like to know when $null(E_\lambda)$ is non-trivial
- Since E_λ is an $n \times n$ matrix, E_λ has a non-trivial null space if and only if E_λ is not invertible, which is true if and only if $\det(E_\lambda) = 0$
- By viewing $\det(E_\lambda)$ as a function of λ, we can figure out when $\det(E_\lambda) = 0$
- Characteristic polynomial - the quantity $\det(E_\lambda)$ viewed as a function of λ

- For a matrix A, the characteristic polynomial of A is $\text{char}(A) = \det(A - \lambda I)$

- Ex. Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

    - By the definition of the characteristic polynomial of A, we have

      $\text{char}(A) = \det(A - \lambda I)$

      $= \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix}\right)$

      $= (1-\lambda)(4-\lambda) - 6 = \lambda^2 - 5\lambda - 2$

- Properties of $\text{char}(A)$ for an $n \times n$ matrix A:

    - $\text{char}(A)$ is a polynomial

    - $\text{char}(A)$ has degree n

    - The coefficient of the λ^n term in $\text{char}(A)$ is ± 1

        - $+1$ if n is even

        - -1 if n is odd

    - $\text{char}(A)$ evaluated at $\lambda = 0$ is $\det(A)$

- The roots of char(A) are the eigenvalues of A

Using the Characteristic Polynomial to Find Eigenvalues

- Ex. Find the eigenvectors/values of $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$

  - We first compute char(A)

    $$\text{char}(A) = \det \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix}$$

    $$= (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4 = (4-\lambda)(-1-\lambda)$$

  - We solve for char(A) = 0 to find eigenvalues

    $\lambda_1 = -1, \lambda_2 = 4$

  - We know non-zero vectors in $\text{null}(A - \lambda_1 I)$ are eigenvectors with eigenvalues -1. Computing,

    $$\text{null}(A - \lambda_1 I) = \text{null}\left(\begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$$

  - And so the eigenvectors of A corresponding to the eigenvalue $\lambda_1 = -1$ are the non-zero multiples of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

- Similarly, for $\lambda_2 = 4$, we compute

$$\text{null}(A - \lambda_2 I) = \text{null}\left(\begin{bmatrix} -3 & 2 \\ 3 & -2 \end{bmatrix}\right) = \text{span}\left\{\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$$

- And so the eigenvectors of A corresponding to the eigenvalue $\lambda_2 = 4$ are the non-zero multiples of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

- Every eigenvalue for a matrix is a root of the characteristic polynomial

- Ex. Find the eigenvectors/values of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

- First, we find the roots of char(A) by setting it to 0

$$\text{char}(A) = \det\begin{bmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix}$$

$$= (1-\lambda)(4-\lambda) - 6 = \lambda^2 - 5\lambda - 2 = 0$$

- By the quadratic formula,

$$\lambda_1 = \frac{5-\sqrt{33}}{2}, \quad \lambda_2 = \frac{5+\sqrt{33}}{2}$$

are the roots of char(A)

- We need to find null $(A - \lambda_1 I)$ and null $(A - \lambda_2 I)$
- We row-reduce $A - \lambda_1 I$

$$\begin{bmatrix} 1 - \frac{5-\sqrt{33}}{2} & 2 \\ 3 & 4 - \frac{5-\sqrt{33}}{2} \end{bmatrix} \to \begin{bmatrix} \frac{-3+\sqrt{33}}{2} & 2 \\ 3 & \frac{3+\sqrt{33}}{2} \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & \frac{4}{-3+\sqrt{33}} \\ 1 & \frac{3+\sqrt{33}}{6} \end{bmatrix} \to \begin{bmatrix} 1 & \frac{4(3+\sqrt{33})}{(-3+\sqrt{33})(3+\sqrt{33})} \\ 1 & \frac{3+\sqrt{33}}{6} \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & \frac{3+\sqrt{33}}{6} \\ 1 & \frac{3+\sqrt{33}}{6} \end{bmatrix} \to \begin{bmatrix} 1 & \frac{3+\sqrt{33}}{6} \\ 0 & 0 \end{bmatrix}$$

- Thus, we conclude that the eigenvectors w/ eigenvalue $\frac{5-\sqrt{33}}{2}$ are the non-zero multiples of

$$\begin{bmatrix} \frac{3+\sqrt{33}}{6} \\ -1 \end{bmatrix}$$

- Similarly, the eigenvectors w/ eigenvalue $\frac{5\sqrt{33}}{2}$

are the non-zero multiples of $\begin{bmatrix} \frac{3-\sqrt{33}}{6} \\ -1 \end{bmatrix}$

Transformations and Eigenvalues

- Let $R: \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counter-clockwise by $90°$.

  There are no non-zero vectors that don't change direction when R is applied.

- $M_R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is a matrix for R, and

  $\text{char}(M_R) = \lambda^2 + 1$

- The polynomial λ^2+1 has no real roots, which means that M_R and R has no real eigenvalues

- λ^2+1 does have complex roots of $\pm i$

- If we allow complex numbers as scalars and view R as a transformation from $\mathbb{C}^2 \to \mathbb{C}^2$, it would have eigenvalues and eigenvectors

- Theorem. If A is a square matrix, then A always has an eigenvalue provided complex eigenvalues are permitted.

Module 16

Intro

- Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ to be the linear transformation w/ matrix $M = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.
  - $\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector for T w/ eigenvalue -1
  - $\vec{v}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is an eigenvector for T w/ eigenvalue 4
  - Let $\vec{a} = \vec{v}_1 + \vec{v}_2$
  - Now $T(\vec{a}) = T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$
    $= -\vec{v}_1 + 4\vec{v}_2$
- Let $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$. $\mathcal{V}$ is a basis for $\mathbb{R}^2$
  - Since $[\vec{a}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $T\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{V}} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}_{\mathcal{V}}$

- When represented in the $\mathcal{V}$ basis, computing T would be easy

- $T(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha T(\vec{v}_1) + \beta T(\vec{v}_2)$

  $= -\alpha\vec{v}_1 + 4\beta\vec{v}_2$

- And so $T\begin{bmatrix}\alpha \\ \beta\end{bmatrix}_{\mathcal{V}} = \begin{bmatrix}-\alpha \\ 4\beta\end{bmatrix}_{\mathcal{V}}$

- T, when acting on vectors written in the $\mathcal{V}$ basis, just multiplies each coordinate by an eigenvalue

  - $[T]_{\mathcal{V}} = \begin{bmatrix}-1 & 0 \\ 0 & 4\end{bmatrix}$

- $[T]_{\mathcal{E}} = \begin{bmatrix}1 & 2 \\ 3 & 2\end{bmatrix}$ and $[T]_{\mathcal{V}} = \begin{bmatrix}-1 & 0 \\ 0 & 4\end{bmatrix}$ are equally valid

<u>Diagonalization</u>

- The process of diagonalizing a matrix A is finding a diagonal matrix that is similar to A

- Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and suppose that $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis so that

$$[T]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$ is a diagonal matrix.

- $\vec{b}_1, \ldots, \vec{b}_n$ are eigenvectors for T
- $[T]_{\mathcal{B}}[\vec{b}_1]_{\mathcal{B}} = \alpha_1 [\vec{b}_1]_{\mathcal{B}} = [\alpha_1 \vec{b}_1]_{\mathcal{B}}$, and in general, $[T]_{\mathcal{B}}[\vec{b}_i]_{\mathcal{B}} = \alpha_i [\vec{b}_i]_{\mathcal{B}} = [\alpha_i \vec{b}_i]_{\mathcal{B}}$
- For $i = 1, \ldots, n$, we have $T\vec{b}_i = \alpha_i \vec{b}_i$
- Since $\mathcal{B}$ is a basis, $\vec{b}_i \neq 0$ for any i, and so each $\vec{b}_i$ is an eigenvector for T w/ corresponding eigenvalue α_i

- Theorem. A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ can be represented by a diagonal matrix if and only if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors for T.
  - If $\mathcal{B}$ is such a basis, then $[T]_{\mathcal{B}}$ is a diagonal matrix
- A matrix is diagonalizable if it is similar to a

diagonal matrix

- Suppose A is an $n \times n$ matrix. A induces some transformation $T_A : \mathbb{R}^n \to \mathbb{R}^n$. By definition, this means $A = [T_A]_{\mathcal{E}}$. The matrix B is similar to A if there is some basis $\mathcal{V}$ so that $B = [T_A]_{\mathcal{V}}$.
    - $A = [\mathcal{E} \leftarrow \mathcal{V}][T_A][\mathcal{V} \leftarrow \mathcal{E}]$
    $= [\mathcal{E} \leftarrow \mathcal{V}] B [\mathcal{V} \leftarrow \mathcal{E}]$
    - A and B are similar if there is some invertible change-of-basis matrix P so $A = PBP^{-1}$
    - B will be a diagonal matrix if and only if P is the change-of-basis matrix for a basis of eigenvectors
        - B will be a diagonal matrix w/ eigenvalues along the diagonal

- Ex. Let $\begin{bmatrix} 1 & 2 & 5 \\ -11 & 14 & 5 \\ -3 & 2 & 9 \end{bmatrix}$ be a matrix and

$\vec{v}_1 = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ are

eigenvectors for A. Diagonalize A.

- First, we find the eigenvalues that correspond to the eigenvectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$.

$$A\vec{v}_1 = \begin{bmatrix} 20 \\ 20 \\ 4 \end{bmatrix} = 4\vec{v}_1, \quad A\vec{v}_2 = \begin{bmatrix} 8 \\ 8 \\ 8 \end{bmatrix} = 8\vec{v}_2,$$

$$A\vec{v}_3 = \begin{bmatrix} 12 \\ 36 \\ 12 \end{bmatrix} = 12\vec{v}_3$$

and so the eigenvalues corresponding to $\vec{v}_1$ is 4, to $\vec{v}_2$ is 8, and to $\vec{v}_3$ is 12.

- The change-of-basis matrix that converts from the $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ basis to the standard basis is

$$P = \begin{bmatrix} 5 & 1 & 1 \\ 5 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}, \text{ and } P^{-1} = \begin{bmatrix} 1/4 & 0 & -1/4 \\ 1/4 & -1/2 & 5/4 \\ -1/2 & 1/2 & 0 \end{bmatrix}$$

- Define D to be the 3×3 matrix w/ the eigenvalues of A along the diagonal (in the order: 4, 8, 12).

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{bmatrix}$$

  - Matrix A written in the basis of eigenvectors

- $A = PDP^{-1}$, and D is the diagonalized form of A.

Non-Diagonalizable Matrices

- Not every matrix is diagonalizable

- Ex. Is the matrix $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ diagonalizable?

  - Computing, $\text{char}(R) = \lambda^2 + 1$ has no real roots, therefore, R has no real eigenvalues.

- Consequently, R has no real eigenvectors, and so R is not diagonalizable.

- Ex. Is the matrix $D = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$ diagonalizable?

  - For every vector $\vec{v} \in \mathbb{R}^2$, $D\vec{v} = 5\vec{v}$, and so every non-zero vector in $\mathbb{R}^2$ is an eigenvector for D
  - Thus, $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ is a basis of eigenvectors for $\mathbb{R}^2$, and so D is diagonalizable.

- Ex. Is the matrix $J = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ diagonalizable?

  - Computing. $\text{char}(J) = (5-\lambda)^2$, which has a double root at 5. Therefore, 5 is the only eigenvalue of J.
  - The eigenvectors of J all lie in

$$\text{null}(J - 5I) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}.$$

  - Since this is a one-dimensional space, there is no basis for $\mathbb{R}^2$ consisting of eigenvectors for J.

- Therefore, J is not diagonalizable.

- Ex. Is the matrix $K = \begin{bmatrix} 5 & 1 \\ 0 & 2 \end{bmatrix}$ diagonalizable?

  - Computing, $\text{char}(K) = (5-\lambda)(2-\lambda)$, which has roots 5 and 2. Therefore, 5 and 2 are the eigenvalues of K.

  - The eigenvectors of K lie in one of

    $$\text{null}(K - 5I) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} \quad \text{or}$$

    $$\text{null}(K - 2I) = \text{span}\left\{\begin{bmatrix} -1 \\ 3 \end{bmatrix}\right\}$$

  - Picking one eigenvector from each nullspace,

    $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix}\right\}$ is a basis for $\mathbb{R}^2$

    consisting of eigenvectors of K

  - Thus, K is diagonalizable.

- We can check if an $n \times n$ matrix is diagonalizable by determining whether there is a basis of eigenvectors in $\mathbb{R}^n$

Geometric and Algebraic Multiplicities

- eigenspaces - subspaces where vectors are stretched by only one eigenvalue
- Let A be an $n \times n$ matrix w/ eigenvalues $\lambda_1, \dots, \lambda_n$. The eigenspace of A corresponding to the eigenvalue λ_i is the null space of $A - \lambda_i I$
  - i.e. the space spanned by all eigenvectors that have the eigenvalue λ_i
- The geometric multiplicity of an eigenvalue λ_i is the dimension of the corresponding eigenspace.
- The algebraic multiplicity of λ_i is the number of times λ_i occurs as a root of the characteristic polynomial of A
  - i.e. the number of times $x - \lambda_i$ occurs as a factor

- The multiplicity of a root of a polynomial is the power of that root in the factored polynomial

- Ex. Let $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of R
  - Computing, $\text{char}(R) = \lambda^2 + 1$ which has no real roots.
  - Therefore R has no real eigenvalues.

- Ex. Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of D
  - Computing, $\text{char}(D) = (5-\lambda)^2$, so 5 is an eigenvalue of D w/ algebraic multiplicity 2.
  - The eigenspace of D corresponding to 5 is $\mathbb{R}^2$. Thus, the geometric multiplicity of 5 is 2.

- Ex. Let $J = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of J

- Computing, $\text{char}(J) = (5-\lambda)^2$, so 5 is an eigenvalue of J w/ algebraic multiplicity 2.
- The eigenspace of J corresponding to 5 is $\text{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$. Thus, the geometric multicity of 5 is 1.

- Ex. Let $K = \begin{bmatrix}5 & 1\\0 & 2\end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of K
  - Computing, $\text{char}(K) = (5-\lambda)(2-\lambda)$, so 5 and 2 are eigenvalues of K, both w/ multiplicity 1.
  - The eigenspace of K corresponding to 5 is $\text{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$, and the eigenspace corresponding to 2 is $\text{span}\left\{\begin{bmatrix}1\\3\end{bmatrix}\right\}$. Thus, both 2 and 5 have a geometric multiplicity of 1.

- Theorem. (Fundamental Theorem of Algebra)

  Let p be a polynomial w/ degree n. Then, if complex roots are allowed, the sum of the multiplicities of the roots of p is n.

- Theorem. Let λ be an eigenvalue of the matrix A.

  Then geometric mult(λ) $\leq$ algebraic mult(λ).

- Theorem. An $n \times n$ matrix A is diagonalizable if and only if the sum of its geometric multiplicities is equal to n. Further, provided complex eigenvalues are permitted, A is diagonalizable if and only if all its geometric multiplicities equal its algebraic multiplicities.

Appendix 1

Equation

- an equation encodes a relationship between quantities
- $\underbrace{\text{slices of cake}}_{C} = \underbrace{\text{slices I ate}}_{M} + \underbrace{\text{slices my brother ate}}_{B}$
    - C, M, B are variables or unknowns
- system of equations: expresses additional relationships
    - i.e. there are 6 pieces of cakes and my brother ate twice as many pieces as me

$$\begin{cases} C = M + B \\ B = 2M \\ C = 6 \end{cases}$$

"the relationship $C = M + B$ holds and the relationship $B = 2M$ holds and the relationship $C = 6$ holds"

- systems of equations can appear through vector equations

- suppose $\vec{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

  $\vec{w}$ is a linear combination of $\vec{u}$ and $\vec{v}$

  if and only if the vector equation $\vec{w} = a\vec{u} + b\vec{v}$

  has a solution for some a and b

  - written in coordinates,

    $$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = a\begin{bmatrix} 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} a+2b \\ 2a+3b \end{bmatrix}$$

  - equating coordinates, a system of equations appears:

    $$\begin{cases} a+2b = 1 \\ 2a+3b = 1 \end{cases}$$

- every vector equation, by way of coordinates, corresponds to a system of equations

Systems of Linear Equations

- there is no guarantee that a general equation, like $x^4 - e^x + 7 = 0$, has a solution, and it might be impossible to decide if it has a solution

- for linear equations and systems of linear equations we can always tell whether there is a solution and what the solutions are
- a linear equation in the variables $x_1, \dots, x_n$ is one that can be expressed as $a_1x_1 + a_2x_2 + \dots + a_nx_n = c$ for constants $a_1, \dots, a_n$ and c
- a system of linear equations is a system of equations consisting of one or more linear equations
- every vector equation corresponds to an equivalent system of linear equations and vice versa
    - equivalent: expressing the same relationships between variables

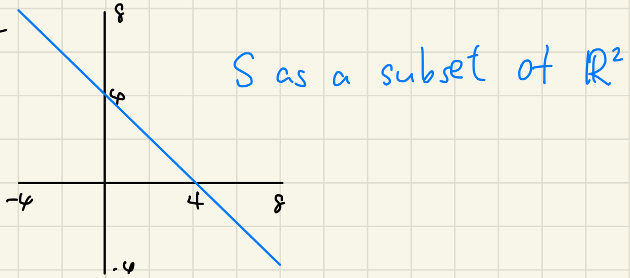
Solution Sets

- a solution to an equation is a particular choice of values for the variables that satisfy (i.e. make true)
- $x+y=4$ has a solution of $x=y=2$, however, we also have $x=4, y=0$; $x=6, y=-2$, etc.

- the solution set, also called the complete solution, to an equation / system of equations is the set of all possible solutions
  - the solution set to $x+y=4$ is

    $$S = \{(x,y) : y = 4-x\}$$

    - S contains infinitely many solutions

- 

- via solution sets, equations and systems of equations can represent geometric objects

Consistent & Inconsistent Systems

- $x^2 = 0$ w/ solution set $S_x \subseteq \mathbb{R}$

  $y^2 = 4$ w/ solution set $S_y \subseteq \mathbb{R}$

  $z^2 = -1$ w/ solution set $S_z \subseteq \mathbb{R}$

  - $S_x = \{0\}$ consists of 1 number
  - $S_y = \{-2, 2\}$ consists of 2 numbers

- $S_2 = \{\}$ consists of no numbers
- the first two equations are consistent
- the third equation is inconsistent

- an equation or system of equations is consistent if it has at least 1 solution
  - i.e. the solution set is not empty
  - otherwise, the equation or system of equations is inconsistent

Equivalent Systems

- two equations or systems of equations are equivalent if they have the same solution sets
  - i.e. $x = 2y$ and $\frac{1}{2}x = y$

Row Reduction

- consider the vector equation $t\vec{u} + s\vec{v} + r\vec{w} + \vec{p}$ where

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} -2 \\ -5 \\ 1 \end{bmatrix}, \quad \vec{p} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}$$

- by expanding in terms of coordinates:

$$\begin{cases} t + 2s - 2r = -15 & \text{row}_1 \\ 2t + s - 5r = -21 & \text{row}_2 \\ t - 4s + r = 18 & \text{row}_3 \end{cases}$$

- most general way: substitution
  - i.e. solve for t in row_1, substitute the result in row_2, etc.
- for linear systems, we can use row reduction
  - also known as elimination
- we can keep track of those numbers in an augmented matrix
  - augmented matrix: matrix w/ extra information

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 1 & -4 & 1 & 18 \end{array}\right]$$

  - contains two types of information:
    - coefficients of variables t, s, r
    - constants on the right side
  - vertical line separates the two types

of numbers

- $$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 1 & -4 & 1 & 18 \end{array}\right] \xrightarrow{\text{row}_3 \to \text{row}_3 - \text{row}_1} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 0 & -6 & 3 & 33 \end{array}\right]$$

$$\xrightarrow{\text{row}_2 \to \text{row}_2 - 2\text{row}_1} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 0 & -3 & -1 & 9 \\ 0 & -6 & 3 & 33 \end{array}\right]$$

$$\xrightarrow{\text{row}_3 \to \text{row}_3 - 2\text{row}_2} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 0 & -3 & -1 & 9 \\ 0 & 0 & 5 & 15 \end{array}\right]$$

$$\xrightarrow{\text{to equations}} \begin{cases} t + 2s - 2r = -15 \\ -3s - r = 9 \\ 5r = 1 \end{cases}$$

The Rules of Row Reduction

- the 3 elementary row operations:
    - swapping two rows ($\text{row}_i \leftrightarrow \text{row}_j$)
    - multiplying a row by a nonzero scalar ($\text{row}_i = k \cdot \text{row}_i$)

- adding a multiple of one row to another
  ($row_i \to row_i + k \cdot row_j$)

- strategy for solving a system:
  1. rewrite the system as an augmented matrix
  2. use elementary row operations to zero-out the lower triangle of the matrix
  3. convert the matrix back to a system of equations
  4. read off the solution (substituting when necessary)

Appendix 2

Reduced Row Echelon Form

- abbreviated rref
- a matrix X is in reduced row echelon form if:
    - the first non-zero entry in every row is a 1
        - these entries are called pivots or leading ones
    - above and below each leading ones are zeros
    - the leading ones form an echelon (staircase) pattern.
        - i.e. if row i has a leading one, every leading one appearing in row $j > i$ also appear to the right of the leading one in row i
    - all rows of zeros occur at the bottom of X
- columns of a reduced row echelon form matrix that

contain pivots are called pivot columns

- every matrix, M, has a unique row echelon form, written $\text{rref}(M)$, which can be obtained from M by applying elementary row operations

<u>Row-Reduction Algorithm</u>

- let M be a matrix

1. If M takes the form $M = [\vec{0} \mid M']$ (i.e. its first column is all zeros), apply the algorithm to M
2. If not, perform a row-swap so the upper left entry of M is non-zero
3. let α be the upper-left entry of M. Perform the row operation $\text{row}_1 \to \frac{1}{\alpha}\text{row}_1$. The upper-left entry of M is now 1 and is called a pivot.
4. Use row operations of the form $\text{row}_i \to \text{row}_i + \beta\,\text{row}_i$ to zero every entry below the pivot.
5. Now, M has a form $M = \left[\begin{array}{c|c} 1 & ?? \\ \hline \vec{0} & M' \end{array}\right]$, where

M' is a submatrix of M.

Apply the algorithm to M'

- the resulting matrix is in pre-reduced row echelon form. To put the matrix in reduced row echelon form, additionally apply step 6.

6. Use the row operations of the form

$row_i \to row_i + \beta row_j$ to zero above each pivot

- Ex. apply the row-reduction algorithm to the matrix

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

- since M starts w/ a column of zeros, we separate it w/ the rest

$$M = \left[\begin{array}{c|cccc} 0 & 0 & 0 & -2 & -2 \\ 0 & 1 & 2 & 3 & 2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right]$$

- next, we perform a row swap to bring a non-zero entry to the upper left

$$\xrightarrow{row_1 \leftrightarrow row_2} \left[\begin{array}{c|cccc} 0 & 1 & 2 & 3 & 3 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right]$$

- the upper left entry is already a 1, so we can use it to zero all entries below

$$\xrightarrow{row_3 \to row_3 - 2row_1} \left[\begin{array}{c|cccc} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

- now we work on the submatrix

$$\left[\begin{array}{cc|ccc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

- the submatrix has a first column of zeros

$$\left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

- now we turn the upper left entry into a 1 and use that pivot to zero all entries below

$$\xrightarrow[\text{row}_3 \to \text{row}_3 + \text{row}_2]{\text{row}_2 \to -\frac{1}{2}\text{row}_2} \left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- The matrix is now in pre-reduced row echelon form. To put it in row echelon form, we zero above each point

$$\xrightarrow{\text{row}_1 \to \text{row}_1 - 3\text{row}_2} \begin{bmatrix} 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Free Variables & Complete Solutions

- $$\text{rref}\left(\left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 1 & -4 & 1 & 18 \end{array}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

- which correspond to the system $\begin{cases} x & = -1, \\ y & = -4, \\ z & = 3 \end{cases}$

from which the solution is immediate

- $\left[\begin{array}{cc|c} 1 & 3 & 2 \\ 2 & 6 & 4 \end{array}\right] \xrightarrow{\text{row reduces to}} \left[\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 0 & 0 \end{array}\right]$,

which isn't exactly a solution

- since the original system had only 1 equation's worth of information, we cannot solve for both x and y based on the original system
- we introduce the arbitrary equation $y = t$:

$$\begin{cases} x + 3y = 2 \\ 2x + 6y = 4 \\ y = t \end{cases} \xrightarrow{\text{row reduces to}} \begin{cases} x + 3y = 2 \\ y = t \end{cases}$$

  - we omit the equation $0 = 0$ since it adds no information

- we now solve for x and y in terms of t

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2-3t \\ t \end{bmatrix} = t\begin{bmatrix} -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

- any choice of t produces a valid solution to the original system
- we call t a parameter and y a free variable
- $t\begin{bmatrix} -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ is the vector form of the line $x+3y=2$

- pick all variables corresponding to non-pivot colomns to be free varicbles
  - free varicble colomns: non-pivot non-augmented colomns of a row-reduced matrix

- Ex. use row reduction to find the complete solution to

$$\begin{cases} x+y+z=1 \\ y-z=2 \end{cases}$$

  - corresponding augmented matrix for the system:

$$A = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

- A is already in the pre-reduced row echelon form, so we only need to zero above each pivot

$$\xrightarrow{row_1 \to row_1 - row_2} \left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 2 \end{array}\right] = rref(A)$$

- the third column of $rref(A)$ is a free variable column, so we introduce the arbitrary equation $z=t$ and solve the system in terms of t

$$\begin{cases} x + 2z = -1 \\ y - z = 2 \\ z = t \end{cases}$$

- written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

- written as a set, the solution set is

$$\left\{ \vec{x} \in \mathbb{R}^3 : \vec{x} = t \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$$

- $\{ 0x + 0y + z = 1 \}$
  - the solution set for the above system is the xy-plane in $\mathbb{R}^3$ shifted up by 1 unit
  - corresponds to $\left[\begin{array}{ccc|c} 0 & 0 & 1 & 1 \end{array}\right]$, which is already in reduced row echelon form
  - column 3: pivot column; columns 1 and 2: free variable columns
  - we introduce 2 arbitrary equations, $x = t$ and $y = s$

$$\begin{cases} 0x + 0y + z = 1 \\ x = t \\ y = s \end{cases}$$

- $$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ s \\ 1 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- Ex. Find the solution set for system

$$\begin{cases} \quad\quad\quad -2w = -2 \\ y + 2z + 3w = 2 \\ 2y + 4z + 5w = 3 \end{cases}$$

in the variables x, y, z, and w

- augmented matrix:

$$M = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 2 & -2 \\ 0 & 1 & 2 & 3 & 2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right]$$

$$\text{rref}(M) = \left[\begin{array}{cccc|c} 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- columns 1 and 3 are free variable columns, so we introduce $x = t$ and $z = s$

$$\begin{cases} y + 2z = -1 \\ w = 1 \\ x = t \\ z = s \end{cases}$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} t \\ -1-2s \\ s \\ 1 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

- thus, the solution set to the system is

$$\left\{ x \in \mathbb{R}^4 : \vec{x} = t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \text{ for some } t, s \in \mathbb{R} \right\}$$

Free Variables & Inconsistent Systems

- when evaluating the number of solutions to a system, pay attention to whether or not the system is consistent
- a system of linear equations has either 0 solution, 1 solution, or infinitely many solutions

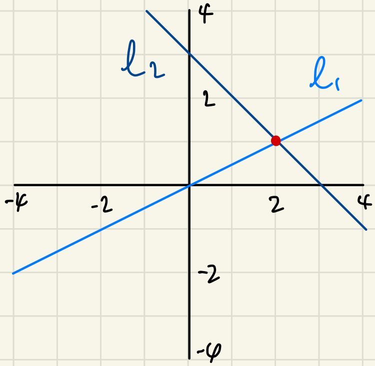
The Geometry of Systems of Equations

- for system $\begin{cases} x - 2y = 0 \\ x + y = 3 \end{cases}$, the only values of

x and y that satisfies both equations is

$(x, y) = (2, 1)$

- each row, viewed in isolation, specifies a line in $\mathbb{R}^2$



- the two lines intersect at $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

- solutions lie in $\ell_1 \cap \ell_2$

- geometrically, a solution to a system of equations is the intersection of objects specified by the individual equations
- the lines for an inconsistent system are parallel and never intersect

Planes & Hyperplanes

- in general, a single equation in n variables requires $n-1$ free variables to describe its complete solution
  - the only exception is the trivial equation $0x_1 + \ldots + 0x_n = 0$, which requires n free variables
  - i.e. the solution to $x + 2y - z = 3$ is a plane, by inspection
- the dimension of the solution set is always one less than the dimension of the ambient space ($\mathbb{R}^2$, $\mathbb{R}^3$, etc.)
  - such sets are called hyperplanes because they are flat and plane like
- unlike a plane, the dimension of a hyperplane need not be 2
- in higher dimensions, solution sets are formed by intersecting hyperplanes.

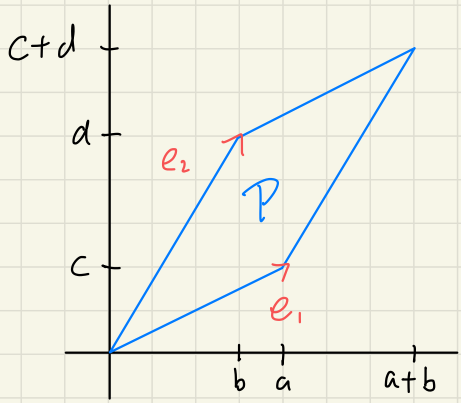
Appendix 3

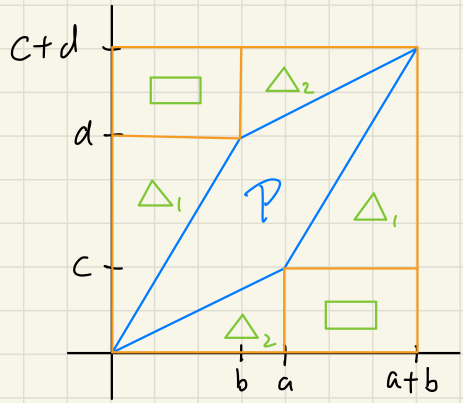
Computing 2 x 2 Determinants

- Theorem. Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then, $\det(M) = ad - bc$

- Let $\vec{c}_1 = \begin{bmatrix} a \\ c \end{bmatrix}$, $\vec{c}_2 = \begin{bmatrix} b \\ d \end{bmatrix}$. We need to compute

  the area of the parallelogram P w/ sides $\vec{c}_1, \vec{c}_2$





- Vol(P) = area of big rectangle

  − area of little rectangles − area of triangles

- Vol(P) =

$$\underbrace{(a+b)(c+d)}_{\text{big }\square} - \underbrace{2bc}_{\text{small }\square\text{'s}} - \underbrace{bd}_{\Delta_1\text{'s}} - \underbrace{ac}_{\Delta_2\text{'s}} = ad - bc$$

- We did not consider anything negatively oriented

  - turns out that $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$

    for all cases

- Ex. Compute the determinant of $M = \begin{bmatrix} 1 & 6 \\ 2 & 7 \end{bmatrix}$ by using the 2×2 formula and by decomposing M into the product of elementary matrices.

  - Using the 2×2 formula,

    $\det(M) = (1)(7) - (6)(2) = -5$

  - Row-reducing,

$$\underbrace{\begin{bmatrix} 1 & -6 \\ 0 & 1 \end{bmatrix}}_{E_1} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1/5 \end{bmatrix}}_{E_2} \underbrace{\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}}_{E_3} M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and so

$$M = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix} = E_1^{-1} E_2^{-1} E_3^{-1}$$

$\det(E_1^{-1}) = \det(E_3^{-1}) = 1 \qquad \det(E_2^{-1}) = -5$

$\det(M) = (1)(-5)(1) = -5$

Computing 3x3 Determinants

- Theorem. Let $M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Then,

$$\det(M) = aei + bfg + cdh - gec - hfa - idb$$

- Rule of Sarrus

    - Let $M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

    - Step 1. Augment M w/ copies of its first two columns

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

    - Step 2. Multiply together then add the entries along the three diagonals of the new matrix. These are called the diagonal products

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

Sum of diagonal products : $aei + bfg + cdh$

- Step 3. Multiply together and then subtract the entries along the three anti-diagonals. These are called anti-diagonal products

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

Difference of anti-diagonal products:

$-gec - hfa - idb$

- Step 4. Add the diagonal products and subtract the anti-diagonal products to get the determinant

$\det(M) = aei + bfg + cdh - gec - hfa - idb$

- Ex. Compute $\det\left(\begin{bmatrix} 1 & 4 & 0 \\ -2 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix}\right)$

- $\left[\begin{array}{ccc|cc} 1 & 4 & 0 & 1 & 4 \\ -2 & 3 & 1 & -2 & 3 \\ 0 & 2 & 1 & 0 & 2 \end{array}\right]$

- Sum of diagonal products:

  $(1)(3)(1) + (4)(1)(0) + (0)(2)(-2) = 3$

- Difference of anti-diagonal products:

  $-(0)(3)(0) - (2)(1)(1) - (1)(-2)(4) = 6$

- Thus, $\det(M) = 3 + 6 = 9$

- The Rule of Sarrus does not apply to larger matrices

Determinant Formulas and Orientation

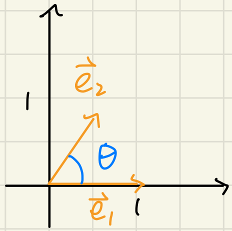
- Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ be an ordered basis for $\mathbb{R}^2$, and let $M = [\vec{b}_1 | \vec{b}_2]$. Since $\vec{b}_1$ and $\vec{b}_2$ are linearly independent, $\det(M) \neq 0$. By applying the definition of the determinant,

- $\det(M) > 0$ means that $\mathcal{B}$ is a right-handed basis
- $\det(M) < 0$ means that $\mathcal{B}$ is a left-handed basis

- Ex. Determine whether the ordered basis $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix} \right\}$ is left-handed or right-handed.
  - Let $A = \begin{bmatrix} 1 & -3 \\ 2 & 2 \end{bmatrix}$ be the matrix whose columns are the elements of the given ordered basis.
  - Using the formula for 2×2 determinants gives us $\det(A) = (1)(2) - (-3)(2) = 8 > 0$
  - And so we conclude $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix} \right\}$ is a right-handed basis.
- Recall the ordered basis $\mathcal{Q} = \{\vec{e}_1, \vec{u}_\theta\}$ where

$\vec{u}_\theta = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ is the unit vector which

forms an angle of θ with the positive x-axis



- Visually, Q should be right-handed when $\theta \in (0, \pi)$,

  left-handed when $\theta \in (\pi, 2\pi)$,

  not a basis when $\theta = 0$ or $\theta = \pi$

- Computing the determinant of the matrix $Q = [\vec{e}_1 | \vec{u}_\theta]$,

  $$\det(Q) = \det([\vec{e}_1 | \vec{u}_\theta]) = \det\left(\begin{bmatrix} 1 & \cos\theta \\ 0 & \sin\theta \end{bmatrix}\right) = \sin\theta.$$

  - $\det(Q) = \sin\theta > 0$ when $\theta \in (0, \pi)$,

    $\det(Q) = \sin\theta < 0$ when $\theta \in (\pi, 2\pi)$.

    $\det(Q) = \sin\theta = 0$ when $\theta \in \{0, \pi\}$