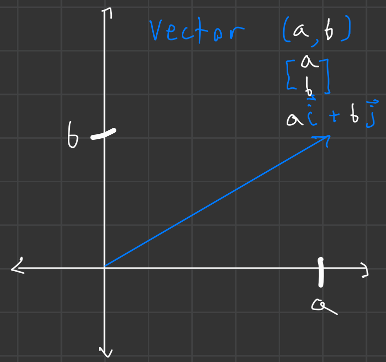
Math

1.1 Vectors & Linear Combinations

Vectors
- momentum
- acceleration
- point on a plane (a, b)
- $\langle a, b, c \rangle$
- $a\vec{i} + b\vec{j} + c\vec{k}$
- tail at origin
- functions : $\cos(x), \sin(x), e(x), \ldots$
- $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ ← $n \times 1$

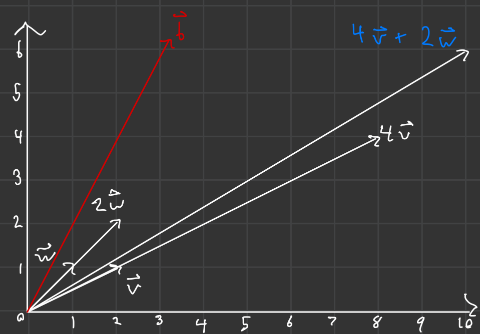


Operations on vectors
- Scalar multiplication
  - $c \cdot \vec{v}$, $c \in \mathbb{R}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \to c \cdot \vec{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$
- Vec. Addition
  - $\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
- Linear Combination
  - $c \cdot \vec{v} + d \cdot \vec{w}$

    for $c, d \in \mathbb{R}$

    is a linear comb. of $\vec{v}$ & $\vec{w}$

    with coefficients c & d



Find $c, d \in \mathbb{R}$ s.t.

$c\vec{v} + d\vec{w} = \vec{b}$

In practice
- Study $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$

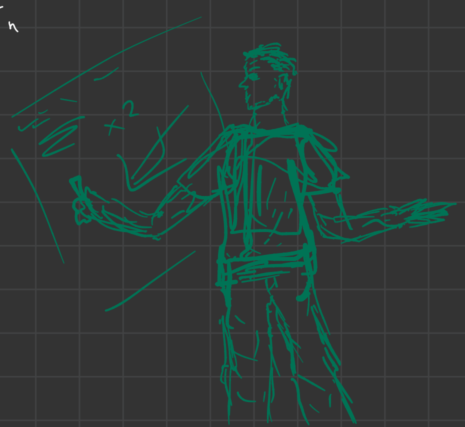
  for $c_1, c_2, \ldots, c_n \in \mathbb{R}$

$$A = \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}}_{n\text{ - columns}} \Big\} m\text{ - rows}$$

Ex. $m = 3$ $n = 2$

$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 7 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 6 \\ 0 & 3 \\ 7 & 3 \end{bmatrix}$



Ex $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\mathbb{R}^2$ = set of all 2×1 vec.

1. Find $c, d \in \mathbb{R}$ s.t. $c\vec{v} + d\vec{w} = \vec{b}$, specific $\vec{b}$
2. Describe set of all $c\vec{v} + d\vec{w}$ $c, d \in \mathbb{R}$

$\boxed{1}$

Systems of eq.

$b = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$

$c\begin{bmatrix} 2 \\ 1 \end{bmatrix} + d\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$

$\begin{bmatrix} 2c \\ c \end{bmatrix} + \begin{bmatrix} d \\ d \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$

$\begin{bmatrix} 2c + d \\ c + d \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$

$\Rightarrow$ $2c + d = 7$

$c + d = 5$

$-\frac{1}{2}(2c + 7 = 7)$

$+ \quad c + d = 5$

$0c + \frac{1}{2}d = \frac{3}{2}$

$2c + d = 7$

$\boxed{\begin{aligned} d &= 3 \\ c &= 2 \end{aligned}}$

or

Matrix Equation

$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 7 \\ 5 \end{bmatrix}}_{\vec{b}}$

$\boxed{2}$ $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ vectors in $\mathbb{R}^3$

$c\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + d\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2c \\ 3c \\ c \end{bmatrix} + \begin{bmatrix} d \\ d \\ 0 \end{bmatrix} = \begin{bmatrix} 2c + d \\ 3c + d \\ c \end{bmatrix}$ $\leftarrow$ Every linear combo. of $\vec{v}, \vec{w}$ is in this form for some $c, d \in \mathbb{R}$

1.2 - Dot product, etc

Ex] let $v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $w = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$

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

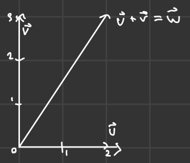
Generally

if $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, $w = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$

$n \times 1$ $\quad$ $n \times 1$

$$\boxed{\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n}$$

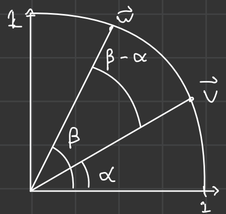
Ex] $\vec{u} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$ $\qquad \vec{w} \cdot \vec{w} \overset{\text{def}}{=} ||\vec{w}||$



Trig

Any unit vector in $\mathbb{R}^2$ is of the form $\vec{u} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$

$v = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$, $w = \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix}$



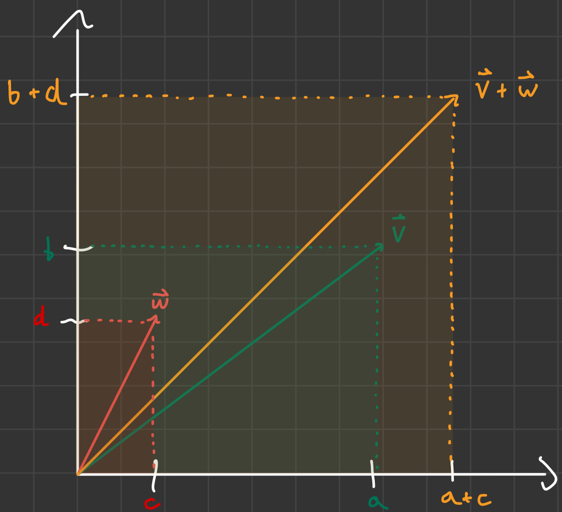
3 important theorems

1 $\vec{v}\cdot\vec{w} = 0 \iff \|\vec{v}+\vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$ Pythagorean Thm
2 $|\vec{v}\cdot\vec{w}| \leq \|\vec{v}\|\cdot\|\vec{w}\|$ Cauchy-Schwarz Ineq.
3 $\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ Triangle Ineq.

Proof of 1

$$\|v+w\|^2 = (v+w)\cdot(v+w)$$
$$= v\cdot v + v\cdot w + w\cdot v + w\cdot w$$
$$= v\cdot v + 2(v\cdot w) + w\cdot w$$
$$= \|v\|^2 + 2(v\cdot w) + \|w\|^2 \quad \square$$

$2(\vec{v}\cdot\vec{w})$ has to $= 0$



$\|\vec{v}\| = \sqrt{a^2+b^2}$, $\|\vec{w}\| = \sqrt{c^2+d^2}$

$\|\vec{v}+\vec{w}\| = \sqrt{(a+c)^2+(b+d)^2}$

$\vec{v}\cdot\vec{w} = ac + bd = 0$ iff $\vec{v}\perp\vec{w}$

$$\|\vec{v}\|\cdot\|\vec{w}\| = \sqrt{a^2+b^2}\cdot\sqrt{c^2+d^2}$$
$$= \sqrt{a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2}$$

$|\vec{v}\cdot\vec{w}| \leq \|\vec{v}\|\cdot\|\vec{w}\|$ Cauchy-Schwarz Inequality

$|ac+bd| \leq \sqrt{a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2}$

with equality iff $\exists\, t \in \mathbb{R}$ s.t. $\vec{v} = t\vec{w}$ or one of the vectors is $\begin{bmatrix}0\\0\end{bmatrix}$

$\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ Triangle inequality

$\sqrt{(a+c)^2+(b+d)^2} \leq \sqrt{a^2+b^2} + \sqrt{c^2+d^2}$

1.3 - Matrices and Column Space

Matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

3×2 $\qquad$ 2×4

$$v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^4 \qquad w \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

4×1 $\qquad$ 1×4

column vector $\qquad$ Row vector

* vectors are a special type of matrix

Matrix times vector

let $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$ $m \times n$

and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ $n \times 1$

$A \cdot \vec{x}$ is defined:

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boxed{x_1 \vec{a}_1 + x_2 \vec{a}_2 + \cdots + x_n \vec{a}_n}$$

Ex

$$\underset{\substack{A \\ 3\times 2}}{\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}} \underset{\substack{x \\ 2\times 1}}{\begin{bmatrix} 1 \\ -1 \end{bmatrix}} = 1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} - 1 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} (1)(1) + (2)(-1) \\ (3)(1) + (4)(-1) \\ (5)(1) + (6)(-1) \end{bmatrix} = \underset{3\times 1}{\boxed{\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}}}$$

Def

let $A_{m \times n}$. The column space of A is the set of all linear combinations of the columns of A

$$C(A) \text{ or } Col(A) \subseteq \mathbb{R}^m$$

For any vector $\vec{x} \in \mathbb{R}^n$, $A\vec{x}$ is a linear combo. of the columns of A.

$$Col(A) = \{ A\vec{x} \mid \vec{x} \in \mathbb{R}^n \}$$

Linear Independence / Dependence

Informally:

The columns of a matrix A are called independent if each new column of A is actually new, i.e. it is not a linear combination of the previous columns

The independence/dep. of the cols of A determine the "size" of $Col(A)$.

If you have more than n vectors in $\mathbb{R}^n$, they must be dependent!

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

Independent

$\begin{bmatrix} 0 \\ 4 \\ 6 \end{bmatrix}$ is not a scalar multiple of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}$ is not a lin. comb. of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 5 \end{bmatrix}$

This matrix has 3 indep. cols in $\mathbb{R}^3$

$\Rightarrow Col(A) = \mathbb{R}^3$

Ex $A = \begin{array}{c} \begin{array}{ccc} \vec{u} & \vec{v} & \vec{u}+\vec{v} \end{array} \\ \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 5 \\ 6 & 0 & 6 \end{bmatrix} \end{array}$

Lin. comb. of $\vec{u}, \vec{v}, \vec{u}+\vec{v}$

$c_1\vec{u} + c_2\vec{v} + c_3(\vec{u}+\vec{v})$

$= (c_1+c_3)\vec{u} + (c_2+c_3)\vec{v}$

$\underset{3\times 1}{\vec{u}} \in \mathbb{R}^3$, $Col(A)$ is just a plane

Span

$\text{span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$ = set of all lin. comb. of $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p$

$$= \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p \mid c_1, c_2, \cdots, c_p \in \mathbb{R}\}$$

Ex if $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}_{m \times n}$

$$\text{Col}(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$$

How many lin. indep. cols. does A have?
Rank of A

which are the first r lin. indep. cols?
They form the basis for Col(A)

How do you describe the dep. cols. in terms of the indep. cols?
- CR factorization
- Sys of eq

2.1 - Elimination

Consider a system

$$\begin{aligned} 1x_1 - 2x_2 + 1x_3 &= 0 \\ 0 \quad 2x_2 - 8x_3 &= 8 \\ 5x_1 \quad 0 \quad -5x_3 &= 10 \end{aligned} \Rightarrow \begin{bmatrix} x_1 \\ 0x_1 \\ 5x_1 \end{bmatrix} + \begin{bmatrix} -2x_2 \\ 2x_2 \\ 0x_2 \end{bmatrix} + \begin{bmatrix} 1x_3 \\ -8x_3 \\ -5x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}$$

$$\Rightarrow x_1 \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -8 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}$$

Solutions to a system

$\Updownarrow$

solutions $\vec{x} \mid A\vec{x} = \vec{b}$

Ex $\begin{aligned} 2x + 3y &= 5 \\ 4x + 2y &= 6 \end{aligned} \Rightarrow \underbrace{\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}}_{\text{coefficient matrix}} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$$\begin{aligned} -2(2x + 3y &= 5) \\ + \quad 4x + 2y &= 6 \\ \hline -4y &= -4 \end{aligned}$$

$\Rightarrow \boxed{y = 1}$

$\Rightarrow \boxed{x = 1}$

Analyze

$\overset{A}{\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}} \vec{x} = \vec{b}$ because $\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ are independent,

$\exists\, \vec{x} \in \mathbb{R}^2 \mid \vec{b} \in \mathbb{R}^2$ or $Col(A) = \mathbb{R}^2$

Independent

$\overset{A}{\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}}$ $\quad Col(A) = span\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$

Dependent

Augmented Matrix

Consider
$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ 0x_1 + 2x_2 - 8x_3 &= 8 \\ 5x_1 + 0x_2 - 5x_3 &= 10 \end{aligned} \Rightarrow \left[\overbrace{\begin{array}{ccc} 1 & -2 & 3 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{array}}^{\text{coefficient matrix } A} \middle| \overbrace{\begin{array}{c} 0 \\ 8 \\ 10 \end{array}}^{\vec{b}} \right]$$

$[A \mid \vec{b}]$ solve $A\vec{x} = \vec{b}$ by performing <u>row operations</u>

- Swap rows
- Scale a row by a non-zero const.
- Replace one row with the sum of itself & a non-zero multiple of another row

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right] \xrightarrow{-5R_1 + R_3 \to R_3} \left[\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -20 & 10 \end{array}\right] \xrightarrow{-5R_2 + R_3 \to R_3}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 20 & -30 \end{array}\right]$$

→ Keep goin w/ row sub

→ Back substitution → $20x_3 = -30$ → $x_3 = -\frac{3}{2}$ → ...

$$A\vec{x} = \vec{b} \xrightarrow{\text{row ops}} \underbrace{U}_{\text{upper triangular}}\vec{x} = \vec{c} \xrightarrow{U \text{ is invertible}} \vec{x} = U^{-1}\vec{c}$$

pretend we left off here

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 30 & -30 \end{array}\right] \underset{\frac{1}{2}R_3}{\xrightarrow{\frac{1}{2}R_2}} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right] \xrightarrow{4R_3 + R_2 \to R_2} \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \longrightarrow$$

$$\xrightarrow{R_1 - R_3 \to R_1} \left[\begin{array}{ccc|c} 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \xrightarrow{2R_2 + R_1 \to R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \Rightarrow \boxed{\begin{aligned} x_1 &= 1 \\ x_2 &= 0 \\ x_3 &= -1 \end{aligned}}$$

Gaussian elim. / row reduc.

System → Augmented Matrix $\xrightarrow{\text{row ops.}}$ rref of $[A \mid \vec{b}]$ →

Using Gaussian Elim. / row reduc to get rref

Def ref & rref

A matrix A is in row echelon form if:

1) All rows of zeros are at the bottom
2) Each leading entry of a row is a column to the right of the leading entry above it.
3) All entries below a leading entry are zero.

A matrix A is in reduced echelon form if it satisfies the above conditions and:

4) Each leading entry is a 1
5) Each leading entry is the only non-zero entry in its column

Ex

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 30 & -30 \end{array}\right] \leftarrow \text{ref} \quad \Rightarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \leftarrow \text{rref}$$

Ex

$$\overset{\text{ref}}{\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right]} \xrightarrow{\text{row ops.}} \overset{\text{rref}}{\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]}$$

The point of rref

Suppose I have a sys. of eq. $A\vec{x} = \vec{b}$ s.t.

$$[A \mid \vec{b}] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{matrix} x_1 - 5x_3 = 1 \\ x_2 + x_3 = 4 \end{matrix} \Rightarrow \text{solution set} \begin{cases} x_1 = 5x_3 + 1 \\ x_2 = -x_3 + 4 \\ x_3 \text{ is free} \end{cases} \Rightarrow \begin{bmatrix} 5t+1 \\ -t+4 \\ t \end{bmatrix} \text{ s.t. } t \in \mathbb{R}$$

$\Rightarrow$ For each choice of $x_3 \in \mathbb{R}$, $x_3 = t$ we get a diff. solution to the original system.

$$\Rightarrow \text{solution set} \left\{ \begin{bmatrix} 5t+1 \\ -t+4 \\ t \end{bmatrix} \middle| \, t \in \mathbb{R} \right\} = \left\{ t\begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} \middle| \, t \in \mathbb{R} \right\}$$

and

1) Exactly 1 solution $\Leftrightarrow$ linearly indep. $\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$
2) ∞ solutions $\Leftrightarrow$ linearly dep., free var $\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{matrix}$
3) No solutions $\Leftrightarrow$ $0 \; 0 \cdots 0 \mid k$ $\quad k \neq 0$

Matrix operations

Let A be $m \times n$. We write a_{ij} to denote the (i, j) entry of A, with row i, col j

$A = [a_{ij}]$

Ex $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ $\quad b_{23} = 6$

Addition & scalar mult.

Let $A = [a_{ij}]$, $B = [b_{ij}]$

1) $A = B \iff a_{ij} = b_{ij}$ for all i, j
2) $A + B = [a_{ij} + b_{ij}]$
3) $cA = [c\, a_{ij}]$

Matrix mult. Def 1

Given matrices $\underset{m \times n}{A}$ and $\underset{n \times p}{B}$ we can define "$A \cdot B$"

$$AB = A \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} \overset{\text{def}}{=} \underbrace{\begin{bmatrix} | & | & & | \\ \underset{m \times n}{A}\, \underset{n \times 1}{b_1} & Ab_2 & \cdots & Ab_p \\ \underset{m \times 1}{|} & | & & | \end{bmatrix}}_{m \times p}$$

Def 2 $\underset{m \times n}{A}$ and $\underset{n \times p}{B}$

$$A \cdot B = \begin{bmatrix} — & r_1 & — \\ — & r_2 & — \\ & \vdots & \\ — & r_m & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} r_1 \cdot b_1 & r_1 \cdot b_2 & \cdots & r_1 \cdot b_p \\ r_2 \cdot b_1 & r_2 \cdot b_2 & \cdots & r_2 \cdot b_p \\ \vdots & \vdots & \ddots & \vdots \\ r_m b_1 & \cdots & \cdots & r_m \cdot b_p \end{bmatrix}$$

Ex $A = \underset{3 \times 2}{\begin{bmatrix} 1 & 0 \\ 4 & 2 \\ -1 & -3 \end{bmatrix}}$, $B = \underset{2 \times 2}{\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}}$ $\qquad AB = \begin{bmatrix} 2 & 3 \\ 6 & 16 \\ 1 & -9 \end{bmatrix}$

i, j entry of $A \cdot B$ is

$(\text{row}_A\ i) \cdot (\text{col}_B\ j)$

Matrix mult. is not commutative ($AB \neq BA$)

Ex $A = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 5 & -5 \end{bmatrix}$

$\underset{3 \times 1}{A}\,\underset{1 \times 3}{B} = \begin{bmatrix} -5 & -5 & 5 \\ 0 & 0 & 0 \\ 5 & 5 & -5 \end{bmatrix}$

$\underset{1 \times 3}{B}\,\underset{3 \times 1}{A} = (-5) + (0) + (-5) = -10$

$\underset{3 \times 2}{A} \cdot \underset{2 \times 5}{B} = 3 \times 5$

$\underset{2 \times 5}{B} \cdot \underset{3 \times 2}{A} = \text{DNE}$

$\underset{2\times2}{A} \cdot \underset{2\times2}{B}$ and $\underset{2\times2}{B} \cdot \underset{2\times2}{A}$ } not equal

$AB = AC \not\Rightarrow B = C$ even if $A \neq 0$

$A \cdot B = 0 \not\Rightarrow A = 0$ or $B = 0$

1.4 – CR Factorization

Recall, the rank of matrix A is the # of indep. cols.

- zero indp. cols. $\Rightarrow A = 0$
- one indp. col. $\Rightarrow \text{rank}\, A = 1$

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

Each column of A is a scalar of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\Rightarrow \text{rank}\, A = 1$

and

Each row of A is a scalar of $[1 \ 2 \ -1 \ 0]$ $\Rightarrow$ "row rank" $= 1$

a matrix A can be "decomposed/factored" into

$$\underset{m\times n}{A} = \overset{C}{\begin{bmatrix} m \times r \end{bmatrix}} \overset{R}{\begin{bmatrix} r \times n \end{bmatrix}}$$ where $r = \text{rank}\, A$

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

$$\underset{3\times1}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} \underset{1\times4}{\begin{bmatrix} 1 & 2 & -1 & 0 \end{bmatrix}} = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & 0 \\ 3 & 6 & -3 & 0 \end{bmatrix} = A$$

In general, C contains the independent columns and R contains instructions for the dependent columns.

$$A \xrightarrow{\text{rref}} \begin{bmatrix} \boxed{1 \ \ 2 \ \ -1 \ \ 0} \\ 0 \ \ 0 \ \ 0 \ \ 0 \\ 0 \ \ 0 \ \ 0 \ \ 0 \end{bmatrix} \leftarrow R$$

Ex

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 4 & 2 \\ 3 & 7 & 6 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

let $C = \underset{3\times2}{\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix}}$, $R = \underset{2\times3}{\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \end{bmatrix}}$

Invertable Matrices

Def

Let A be an $n \times n$ matrix. A is called invertible if there exists another $n \times n$ matrix C s.t.

$$\underline{AC = I_n = CA} \qquad \underline{I_n \vec{b} = \vec{b}}$$

$$\underline{AC\vec{b} = I_n \vec{b} = \vec{b}}$$

If C exists, • it's called the inverse of A, and is denoted $C = A^{-1}$
• it's unique

Ex

$E_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ ← Row op.: $2R_2 + R_1 \to R_1$

$E_2 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ ← Inverse row op.: $-2 + R_1 \to R_1$

$$E_1 E_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$$E_2 E_1 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$E_2 = E_1^{-1}$ or $E_1 = E_2^{-1}$ $\to$ E_1 and E_2 are inverses

Ex

$A = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$ ← invertable?

need: $A \overbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}^{C} = I_2 \Rightarrow \begin{bmatrix} a+2c & b+2d \\ 4a+8c & 4b+8d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$\begin{aligned} a + 2c &= 1 \\ b + 2d &= 0 \\ 4a + 8c &= 0 \\ 4b + 8d &= 1 \end{aligned} \Rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 4 & 0 & 8 & 0 & 0 \\ 0 & 4 & 0 & 8 & 1 \end{array}\right] \xrightarrow{R_3 - 4R_1 \to R_1} \left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 4 & 0 & 8 & 1 \end{array}\right]$$ ← contradiction

$\Rightarrow$ The system is inconsistent

$\Rightarrow$ No such C

$\Rightarrow$ A is not invertable

Thm

$\underset{n \times n}{A}$ is invertible iff $\text{rank}(A) = n$

or

$\underset{n \times n}{A}$ is invertible iff $\det(A) \neq 0$

Thm.
Let A be an $n \times n$ invertible matrix. Then, for any $\vec{b} \in \mathbb{R}^n$, the system $A\vec{x} = \vec{b}$ has exactly one solution given by $\vec{x} = A^{-1}\vec{b}$

Pf. $\underset{n\times n}{A}\vec{x} = \vec{b}$

A is invertable $\Rightarrow \exists A^{-1}$ s.t. $AA^{-1} = I_n = A^{-1}A$

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

Suppose $\vec{y}$ is some other solution $\Rightarrow A\vec{y} = b \Rightarrow \vec{y} = A^{-1}\vec{b} = \vec{x}$

$\Rightarrow \vec{y} = \vec{x}$ $\square$

If $[A \mid \vec{b}]$ is a system with invertible coefficient matrix A, then it can be solved $\forall \vec{b}$

If A is invertible, then $Col(A) = \mathbb{R}^n$.

In particular, any $\vec{b}' \in \mathbb{R}^n$ is in $Col(A) = \mathbb{R}^n$.

Rank vs. Invertibility

Suppose A is invertible but has less than n indp cols.

$A\begin{bmatrix} -c \\ -d \\ 1 \end{bmatrix} = 0 \Rightarrow A^{-1}A\begin{bmatrix} -c \\ -d \\ 1 \end{bmatrix} = A^{-1}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \leftarrow 1 \neq 0$ $\qquad \begin{bmatrix} -c \\ -d \\ 1 \end{bmatrix} \in Null(A)$

ABSENT FOR ONE DAY

This day covered invertible matrices and their properties row op. on I_n $\rightsquigarrow$ same row op on A

$\to E$ $\qquad\qquad\qquad\qquad$ $\to E\cdot A$

In general, the operation $kR_i + R_j \to R_j$ with $k \in \mathbb{R}$ is given by "I_n with k in the (j, i) entry"

Row reduction to $\mathrm{rref}(A)$ can be viewed as a sequence of matrix products.

$$\underbrace{E_s \cdots E_2 E_1}_{s \text{ steps of row reduc}} A = \mathrm{rref}(A)$$

Thm. $\underset{m\times n}{A}$ is invertible iff $\mathrm{rref}(A) = I_n$

Pf

Suppose $\mathrm{rref}(A) = I_n$

Then $\mathrm{rref}(A) = E_s \cdots E_2 E_1 A \Rightarrow I_n = \overbrace{(E_s \cdots E_2 E_1)}^{A^{-1}} A$

$\Rightarrow$ $A^{-1} = E_s \cdots E_2 E_1$ $\quad$ * technically we sould check $A(E_s \cdots E_2 E_1) = I_n$

Assume A is invertible $\Rightarrow \mathrm{rref}(A) = E_s \cdots E_2 E_1 A$ ← a product of invertible matrices

"sub proof" If $\underset{n\times n}{A}$ and $\underset{n\times n}{B}$ are invertible, then $A\cdot B$ is also inv.

$(AB)\cdot(B^{-1}A^{-1}) = I_n \Rightarrow (AB)(B^{-1}A^{-1}) = A\underbrace{BB^{-1}}_{I_n}A^{-1} = A\cdot I_n \cdot A^{-1} = I_n$

and $\quad B^{-1}A^{-1}\cdot AB = B^{-1}\cdot I_n \cdot B = I_n$

$\Rightarrow \quad (AB)^{-1} = B^{-1}A^{-1}$ $\square$

$\Rightarrow$ $\mathrm{rref}(A)$ is invertible $\Rightarrow$ $\mathrm{rank}(\mathrm{rref}(A)) = n$

$\Rightarrow$ all rows are indep. $\Rightarrow$ no rows of zero $\Rightarrow$ pivot in every col.

$\Rightarrow$ $\mathrm{rref}(A) = I_n$ $\square$

LU Factorization

Let A $m\times n$ matrix

Goal: reduce A to ref using only "replace" ops (no swap or scale)

If this can be done, then A can be decomposed into

$$\underset{m\times n}{A} = \underset{m\times m}{L}\,\underset{m\times n}{U} \quad \text{where}$$

L = lower triangular, invertible, 1's on diagonal
U = upper triangular ref.

Ex $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{bmatrix} \xrightarrow[\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}]{E_1:\ -2R_1+R_2\to R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 4 & 5 \end{bmatrix} \xrightarrow[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}]{E_2:\ -3R_1+R_3\to R_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 4 & 2 \end{bmatrix} \xrightarrow[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}]{E_3:\ -2R_2+R_3\to R_3} \underbrace{\overset{U}{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}}}_{\text{ref}}$

$E_3 E_2 E_1 A = U$

multiply by $(E_3 E_2 E_1)^{-1} = \overbrace{E_1^{-1} E_2^{-1} E_3^{-1}}^{L}$

A has an LU factorization if

$A = \begin{bmatrix} \begin{bmatrix} k\times k \end{bmatrix} & \\ & \end{bmatrix}_{n\times n}$ must be invertible

$A^{-1} = U^{-1} L^{-1}$

Ex $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{bmatrix}$ $\quad A \to E_1 A \to E_2 E_1 A \to E_3 E_2 E_1 A$

$\Rightarrow (E_3 E_2 E_1) A = U \Rightarrow A = (E_3 E_2 E_1)^{-1} U$

$\Rightarrow A = \underbrace{E_1^{-1} E_2^{-1} E_3^{-1}}_{L} U$

$E_1 = \underset{-2R_1+R_2\to R_2}{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \Rightarrow E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$E_2 = \underset{-3R_1+R_3\to R_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}} \Rightarrow E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$

$E_3 = \underset{-2R_2+R_3\to R_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}} \Rightarrow E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

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

$LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \underset{\text{ref}(A)}{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}} = A$

For sparse (lots of zero entries) matrices, LU is much faster than row reduction, and much much faster than A^{-1}

3.1 – Vector spaces & sub spaces

n dimensional Euclidean space : $\mathbb{R}^n$

$$\mathbb{R}^n = \left\{ \underset{n \times 1}{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}} \;\middle|\; a_i \in \mathbb{R} \right\}$$

Def.

A vector space V is a non-empty set equipped with 2 operations

(1) Addition : If $u, v \in V$, then $u+v$ is defined and $u+v \in V$

(2) Scalar mult. : If $u \in V$ and $c \in \mathbb{R}$, cu is defined and $cu \in V$

and a long list of axioms are satisfied (pg 89)

eg. $u+v = v+u$

$c(u+v) = cu + cv$

$\vdots$

There must be a well defined notion of a zero vector, $\vec{0}$ → $c \cdot \vec{0} = \vec{0}$ $u + \vec{0} = u$

Ex

$\mathbb{R}^n$, $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

(1) P_n = set of all polynomials in one variable of degree $\leq n$

$\vec{0} = 0$ polynomial

(2) $M_{m \times n}$ or $\mathbb{R}^{m \times n}$ = set of all $m \times n$ matrices

eg. $M_{3\times 2} = \mathbb{R}^{3 \times 2} = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \;\middle|\; a_{ij} \in \mathbb{R} \right\}$ $\vec{v} = \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 1 & 1 \end{bmatrix} \in M_{3\times 2}$

(3) Let $I \subseteq \mathbb{R}$ is an interval

$\mathcal{F}(I)$ = set of all functions $f : \underset{\text{domain}}{I} \longrightarrow \underset{\text{codomain}}{\mathbb{R}}$

$f, g \in \mathcal{F}(I) \rightarrow (f+g)(x) = f(x) + g(x)$

$f \in \mathcal{F}(I)$ $(c \cdot f)(x) = c f(x)$

Define $\left.\begin{array}{l} \vec{0} : I \longrightarrow \mathbb{R} \\ \vec{0}(x) = 0 \\ x \in I \end{array}\right\}$ zero function

non example

$\mathbb{H}$ = upper half plane of $\mathbb{R}^2$

$c \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{H} \xrightarrow{c=-1} \begin{bmatrix} -1 \\ -1 \end{bmatrix} \notin \mathbb{H}$

So $\mathbb{H}$ is not a vector space b/c it's not "closed under scalar mult."

To be a vector space, all linear combos must stay in V

If $\vec{u}, \vec{v} \in V$, then $c\vec{u} + d\vec{v} \in V$ for all $c, d \in \mathbb{R}$

Def

Let V be a vector space and $H \subseteq V$
Then H is called a subspace of V if H is also a vector space.

Ex $\underset{m \times n}{A}$, $\text{col}(A)$ is a subspace $\mathbb{R}^m$

Ex $H \subseteq \mathbb{R}^2$ subset but not subspace

fails #3

Subspace test

V is a vec. space and $H \subseteq V$. Then H is a subspace of V iff

(1) $\vec{0} \in H$

(2) if $\vec{u}, \vec{v} \in H$, then $\vec{u} + \vec{v} \in H$

"closed under vector addition"

(3) if $\vec{u} \in H$ and $c \in \mathbb{R}$, $c\vec{u} \in H$

"closed under scalar multiplication"

Possible subspaces of $\mathbb{R}^2$

$\{\vec{0}\}$ zero subspace point

$\text{span}\{\vec{v}\}$, $\vec{v} \neq 0$ lines through $(0,0)$

$\mathbb{R}^2$ whole space plane

Possible subspaces of $\mathbb{R}^3$

$\{\vec{0}\}$ zero subspace point

$\text{span}\{\vec{v}\}$, $\vec{v} \neq 0$ lines through $(0,0)$

$\text{span}\{\vec{u}, \vec{v}\}$ $\vec{u}, \vec{v}$ are indp plane

$\mathbb{R}^3$ 3D space

Ex $U = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \middle| \ a, b, d \in \mathbb{R} \right\}$

$U \subseteq M_{2 \times 2}$ subspace?

subspace test:

(1) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in U$ ✓

(2) closed under addition?

Let $\begin{bmatrix} a_1 & b_1 \\ 0 & d_1 \end{bmatrix} \in U$ and $\begin{bmatrix} a_2 & b_2 \\ 0 & d_2 \end{bmatrix} \in U$

$\Rightarrow \begin{bmatrix} a_1 & b_1 \\ 0 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ 0 & d_1 + d_2 \end{bmatrix} \in U$ ✓

(3) ✓

U is a subspace of $M_{2 \times 2}$

Ex $T = \{ A \in \mathbb{R}^{3\times 2} \mid \text{rank } A = 2 \text{ or rank } A = 0 \}$ $\quad T \subseteq M_{3\times 2}$ subspace

(1) $\vec{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \in T$ b/c $\text{rank } \vec{0} = 0$ ✓

(2) $\underset{\text{rank}=2}{A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}, \underset{\text{rank}=2}{B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}}$ $\quad \underset{\text{rank}=1}{A + B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}} \Rightarrow$ $A, B \in T$, but $A + B \notin T$ ✗

T is not a subspace of $M_{3\times 2}$

Four fundamental subspaces associate to $A_{m\times n}$

(1) $\text{Col}(A)$ = set of all lin. combos of cols of $A \leftarrow \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$

$= \text{span}\{ \vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n \}$ is a subspace of $\mathbb{R}^m$

(2) Row space of A = set of all lin. combos of rows of $A \leftarrow \begin{bmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_m \end{bmatrix}$

$= \text{span}\{ \vec{r}_1, \vec{r}_2, \cdots \vec{r}_n \}$

(3) Nullspace of A is a subspace of $\mathbb{R}^n$

$\text{Null}(A) = \{ \vec{x} \in \mathbb{R}^n \mid \underset{m\times n}{A} \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{0}} \}$

(4) "Left Nullspace" coming soon

Thm

Let $A_{m\times n}$. $\text{Null}(A)$ is a subspace of $\mathbb{R}^n$

Pf

(1) $\vec{0} \in \text{Null}(A)$ since $A\vec{0} = \vec{0}$ ✓

(2) Let $\vec{u}, \vec{v} \in \text{Null}(A)$

since $\vec{u} \in \text{Null}(A) \Rightarrow A(\vec{u}) = \vec{0}$

since $v \in \text{Null}(A) \Rightarrow A(\vec{v}) = \vec{0}$

$A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} = \vec{0} + \vec{0} = \vec{0}$ ✓

(3) Let $\vec{u} \in \text{Null}(A)$ and $c \in \mathbb{R}$

Since $\vec{u} \in \text{Null}(A) = A\vec{u} = \vec{0}$

wts. $A(c\vec{u}) = \vec{0} \iff c\vec{u} \in \text{Null}(A)$

$A(c\vec{u}) = c \cdot A\vec{u} = c \cdot \vec{0} = \vec{0}$ ✓ $\Rightarrow c\vec{u} \in \text{Null}$ ✓

$\Rightarrow \text{Null}(A)$ is a subspace of $\mathbb{R}^n$

Computing Null(A)

A system of lin. eq. is "homogeneous" if it can be written in the form $A\vec{x} = \vec{0}$ for some A

Given $\underset{m\times n}{A}$, $Null(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$ = solutions to homogeneous system $A\vec{x} = \vec{0}$

Trivial solution: $\vec{x} = \vec{0}$ $\quad A\vec{0} = \vec{0}$

Ex) $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \end{bmatrix}$ solve $A\vec{x} = \vec{0}$ for non-trivial $\vec{x} \in \mathbb{R}^n$

$[A \mid \vec{0}] \xrightarrow{rref} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ $\quad$ sol set $= \left\{ \begin{array}{l} x_1 = -2x_2 + 3x_3 \\ \left.\begin{array}{l} x_2 \\ x_3 \end{array}\right\} \text{free} \end{array} \right.$

Any solution to $A\vec{x} = \vec{0}$ is of the form

$\vec{x} = s\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$ for some $s, t \in \mathbb{R}$

$\Rightarrow Null(A) = \left\{ s\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \,\middle|\, s, t \in \mathbb{R} \right\} = \text{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \right\}$

Observation

$A\vec{x} = \vec{0}$ has a non-trivial solution iff $[A \mid \vec{0}]$ has a free var.

* If # columns of A > # of rows of A, then $[A \mid \vec{0}]$ has a free var.
* If # of cols. of A > # of leading 1's, then $[A \mid \vec{0}]$ has a free var.

Thm $\underset{m\times n}{A}$, $Null(A) \neq \{\vec{0}\}$ iff $rank(A) < n$

PF

Assume $Null(A) \neq \{\vec{0}\} \Rightarrow \exists$ a non-trivial solution to $A\vec{x} = 0$

$\Rightarrow \exists\, \vec{x} \neq \vec{0}$ s.t. $A\vec{x} = \vec{0}$

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$

$A\vec{x} = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{0} \Rightarrow x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{0}$ $\quad$ At least one $x_i \neq 0$

$x_i\vec{a}_i = \underbrace{-x_2\vec{a}_2 - x_3\vec{a}_3 - \cdots - x_n\vec{a}_n}_{\text{without } x_i\vec{a}_i}$

$\vec{a}_i = -\frac{x_1}{x_i}\vec{a}_2 - \cdots - \frac{x_n}{x_i}\vec{a}_n \leftarrow$ without $\frac{x_i}{x_i}\vec{a}_i$ $\Rightarrow rank(A) \leq n-1 < n$ $\quad [\Leftarrow]$

Def of linear Independence!

A set of vectors $\{v_1, v_2, \cdots, v_p\}$ in a vector space V are called linearly independent iff the only solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p = \vec{0}$ is the trivial solution $c_1 = c_2 = \cdots = c_p = 0$.

If $\exists$ a non-trivial solution, the set of vectors is dependent.

Ex $v_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $v_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}$, $v_3 = \begin{bmatrix}2\\1\\0\end{bmatrix} \in \mathbb{R}^3$

non-trivial solutions to $\begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = \vec{0}$?

$[A \mid \vec{0}] = \left[\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \underset{\text{Solution set}}{\Rightarrow} \begin{cases} x_1 = 2x_3 \\ x_2 = -x_3 \\ x_3 \text{ free} \end{cases} \Rightarrow \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}2x_3\\-x_3\\x_3\end{bmatrix} = x_3\begin{bmatrix}2\\-1\\1\end{bmatrix}$

$\Rightarrow \text{Null}(A) = \left\{ t\begin{bmatrix}2\\-1\\1\end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\} = \text{span}\left\{\begin{bmatrix}2\\-1\\1\end{bmatrix}\right\}$

In particular, any $t \neq 0$ gives a non trivial solution to $A\vec{x} = \vec{0}$.

If $\{v_1, v_2, \dots, v_p\}$ in $\mathbb{R}^n$ with $p > n$, then it must be indep.

Ex $A = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$, $B = \begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}$, $C = \begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}$ as vectors in $V = M_{2\times2} = \mathbb{R}^{2\times2}$

Suppose $c_1 A + c_2 B + c_3 C = \vec{0} \Rightarrow \begin{bmatrix} c_1 + c_2 + c_3 & c_2 + 2c_3 \\ 0 & c_1 + c_3 \end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$

$\Rightarrow \begin{array}{l} c_1 + c_2 + c_3 = 0 \\ c_2 + 2c_3 = 0 \\ c_1 + c_3 = 0 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[I_3 \,\middle|\, \begin{matrix}0\\0\\0\end{matrix} \right] \Rightarrow c_1 = c_2 = c_3 = 0 \Rightarrow \{A, B, C\}$ forms a linearly indep. set in $M_{2\times2}$

Bases and Dimension

Recall, $e_1, e_2, \dots, e_n \in \mathbb{R}^n$ denotes columns of I_n.

$e_i = \begin{bmatrix}0\\0\\\vdots\\1\\0\\\vdots\\0\end{bmatrix} \leftarrow i^{\text{th}}$ position The "standard basis" on $\mathbb{R}^n$

eg. $\mathbb{R}^3$, $e_1 = \underset{\vec{i}}{\begin{bmatrix}1\\0\\0\end{bmatrix}}$, $e_2 = \underset{\vec{j}}{\begin{bmatrix}0\\1\\0\end{bmatrix}}$, $e_3 = \underset{\vec{k}}{\begin{bmatrix}0\\0\\1\end{bmatrix}}$

$\{e_1, e_2, \dots, e_n\}$ in $\mathbb{R}^n$ has two important properties

(1) $\text{span}\{e_1, e_2, \dots, e_n\} = \mathbb{R}^n$

(2) $\{e_1, e_2, \dots, e_n\}$ form a linearly independent set.

Def

Let H be a non-zero subspace of V. a basis for H is a set of vectors $\{b_1, b_2, \dots, b_3\}$ s.t.

(1) $\text{span}\{b_1, b_2, \dots b_p\} = H$

(2) $\{b_1, b_2, \dots, b_p\}$ is indp.

Pf.

Suppose $\{b_1, \dots, b_p\}$ is a basis for some subspace $H \subseteq V$, and let $x \in H$.

Suppose we have $c_1\vec{b}_1 + c_2\vec{b}_2 + \dots + c_p\vec{b}_p = x = d_1\vec{b}_1 + d_2\vec{b}_2 + \dots + d_p\vec{b}_p \Rightarrow (c_1 - d_1)\vec{b}_1 + (c_2 - d_2)\vec{b}_2 + \dots + (c_p - d_p)\vec{b}_p = \vec{0}$

$\Rightarrow c_i = d_i$ for all i $\square$

Ex $\mathbb{P}_2 = \{a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$

$\{1, x, x^2\}$ is called the std basis for $\mathbb{P}_n$

More generally std. basis for $\mathbb{P}_n$ is $\{1, x, \dots, x^n\}$

Bases for $\mathbb{R}^n$

Let $A = [a_1 \cdots a_n]$ be $n \times n$, invertible

(1) For any $b \in \mathbb{R}^n$, the system $A\vec{x} = \vec{b}$ is consistent with solution $\vec{x} = A^{-1}\vec{b}$

$\Rightarrow \vec{b} \in \text{span}\{a_1, a_2, \cdots, a_n\} \Rightarrow \text{span}\{a_1, a_2, \cdots, a_n\} \in \mathbb{R}^n$

(2) The only solution to $A\vec{x} = \vec{0}$ is $\vec{x} = A^{-1}\vec{0} = \vec{0}$

$\Rightarrow \{a_1, a_2, \cdots, a_n\}$ are indp

Std bases for $M_{2\times 2}$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \text{span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

Thm

$\{v_1, v_2, \cdots, v_n\}$ is a basis for $\mathbb{R}^n$ iff $A = [v_1\ v_2 \cdots v_n]$ is invertible

Cor

If A is invertible it gives a basis and cols of A^{-1} also give a basis.

Cor

$\exists$ infinitely many bases for $\mathbb{R}^n$

eg $\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $a \neq 0$

Key fact

If $H \subseteq V$ is a subspace and $\{b_1, b_2, \cdots, b_p\}$ is a basis for H, then any other basis will also have exactly p vectors

Def

Let $H \neq \{0\}$ be a subspace of V. The $\underbrace{\text{dimension of } H}_{\dim H}$ is the number of vectors in any basis for H. $\dim\{\vec{0}\} = 0$

Ex]

(1) $\dim \mathbb{R}^n = n$

(2) $\dim \mathbb{P}_2 = 3$

$\mathbb{P}_2 = \{a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$ $\quad \{1, x, x^2\}$

$\dim \mathbb{P}_n = n+1$ $\quad \{1, x, x^2, \dots, x^n\}$

(3) $\dim M_{2\times 2} = 2 \cdot 2$

(4) If $\vec{u} \neq 0 \in V$, $H = \text{span}\{\vec{u}\}$ is a subspace of V.

$\text{span}\{\vec{u}\} = H$ ✓ $\qquad \{\vec{u}\}$ is indp ✓

$\dim H = 1$

(4.5) $\vec{u}, \vec{v} \in V$, $\quad H = \text{span}\{\vec{u}, \vec{v}\}$

$\vec{u}, \vec{v} \neq \vec{0}$

$$\dim H = \begin{cases} 1 & \text{if } \{\vec{u}, \vec{v}\} \text{ is dep} \\ 2 & \text{if } \{\vec{u}, \vec{v}\} \text{ is indep} \end{cases}$$

Bases for the Four Fundamental spaces

Ex] Basis for Null(A)

Let $\underset{m\times n}{A} = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$ $\qquad [A \mid \vec{0}] \xrightarrow{\text{rref}} \left[\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

$$\Rightarrow \text{solution set} = \begin{cases} x_1 = 2x_2 + x_4 - 3x_5 \\ x_2 \ \text{free} \\ x_3 = -2x_4 + 3x_5 \\ x_4 \ \text{free} \\ x_5 \ \text{free} \end{cases} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 3x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow \mathcal{B}_1 = \underbrace{\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}}_{\text{forms a basis for Null}(A)} \Rightarrow \text{Null}(A) \Rightarrow \dim(\text{Null}(A)) = 3 = \#\text{ free vars} = \text{nullity of } A$$

Basis for Col(A)

Identifying the "pivot cols" of A

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathcal{B}_2 = \left\{ \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} \right\} \Rightarrow \dim(\text{col}(A)) = \text{rank}(A) = 2$$

Basis for Row(A)

Non-zero rows of rref(A) form a basis for Row(A)

$$\Rightarrow \mathcal{B}_3 = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -3 \end{bmatrix} \right\}$$

Connection to CR

if $\underset{m\times n}{A} = CR$ is a CR factorization, then

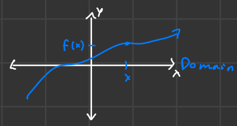
(1) $\{\vec{c}_1, \cdots, \vec{c}_r\}$ is a basis for $\mathrm{Col}(A)$

(2) $\{\vec{R}_1, \cdots, \vec{R}_r\}$ is a basis for $\mathrm{Row}(A)$

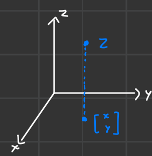
(3) The dependencies described in R form a basis for $\mathrm{Null}(A)$

Linear transformations

Calc 1/2 : $f : \underset{\text{domain}}{\mathbb{R}} \to \underset{\text{codomain}}{\mathbb{R}}$



Calc 3 : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$



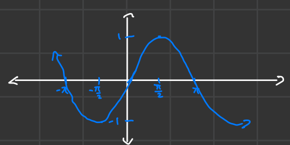
For us : $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $\quad T(\vec{x}) = \vec{y}$ $\quad \underset{n\times 1}{[\;]} \xrightarrow{T} \underset{m\times 1}{[\;]}$

For each $\underset{\text{input}}{\vec{x}} \in \mathbb{R}^n$, $\underset{\text{output}}{T(\vec{x})} \in \mathbb{R}^m$ is called the <u>image of $\vec{x}$ via T</u>.

The set of all possible "images" via T is called the <u>Image of T</u> or <u>Range of T</u>

$$\mathrm{Range}(T) = \mathrm{Image}(T) = \mathrm{Im}(T)$$
$$= \{T(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\} = \{\vec{y} \in \mathbb{R}^m \mid y = T(\vec{x}) \text{ for some } x \in \mathbb{R}^n\}$$

$\boxed{\text{Ex}}$ $F(x) = \sin x : \mathbb{R} \to \mathbb{R}$



$f(\pi/2) = 1$

$1 \in \mathbb{R}$

$\boxed{\text{Ex}}$ $T : \mathbb{R}^2 \to \mathbb{R}^2$ $\quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longmapsto \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ $\quad T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$

$\nexists \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ s.t. $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\mathrm{Im}(T) = \{T(\vec{x}) \mid x \in \mathbb{R}^2\} = \left\{\begin{bmatrix} x_1 \\ 0 \end{bmatrix} \middle| x_1 \in \mathbb{R}\right\} = \mathrm{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} \subseteq \mathbb{R}^2$

$\dim(\mathrm{Im}(T)) = 1$

Matrix transformation

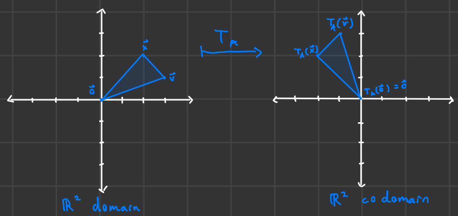
Given an $m \times n$ matrix A, if $\vec{x} \in \mathbb{R}^n$, then $\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} \in \mathbb{R}^m$

So, the "rule" $\vec{x} \longmapsto A\vec{x}$ defines a transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$

$T_A(\vec{x}) = A\vec{x}$ $\quad T_A$ is the transformation associated to A.

Ex $\underset{2\times 2}{A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}$ defines a transformation $T_A : \overset{\text{cols}}{\mathbb{R}^2} \to \overset{\text{rows}}{\mathbb{R}^2}$

$T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$ is the image of $\begin{bmatrix} x \\ y \end{bmatrix}$ via T_A



Def

Let V, W be vector spaces. A function $T : V \to W$ is called a linear transformation if

(1) $T(\underbrace{\vec{u} + \vec{v}}_{\in V}) = \underbrace{T(\vec{u}) + T(\vec{v})}_{\in W}$ for all $\vec{u}, \vec{v} \in V$

(2) $T(c \cdot \vec{u}) = c \cdot T(\vec{u})$ for all $\vec{u} \in V$ and $c \in \mathbb{R}$

Prop: A matrix transformation is a linear transformation

Pf

A is $m \times n$, $T_A : \mathbb{R}^n \to \mathbb{R}^m$ $\quad T_A(x) = Ax$

Let $\vec{u}, \vec{v} \in \mathbb{R}^n$

$T_A(\vec{u} + \vec{v}) = A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} = T_A(\vec{u}) + T_A(\vec{v})$ $\quad$ (2) ✓

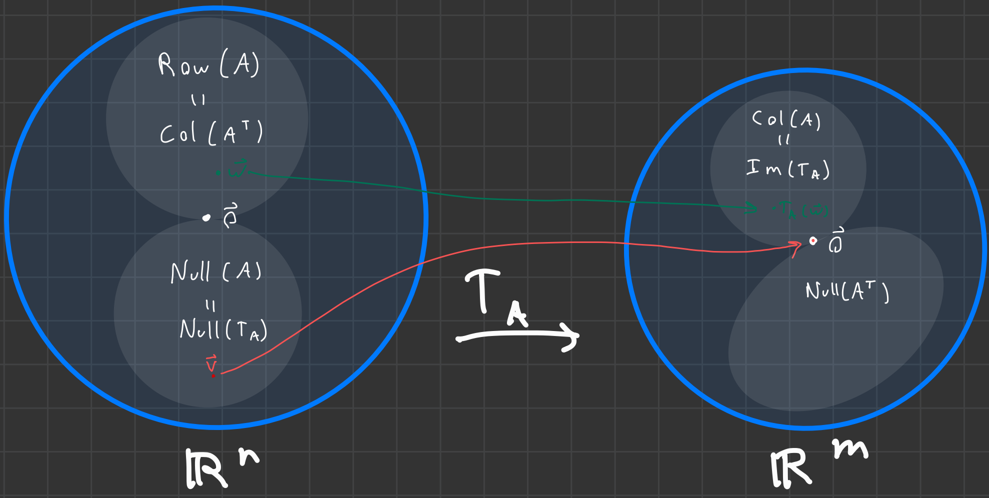
Ex $T : \mathbb{P}_3 \to \mathbb{P}_2$ $\quad T(P) = \frac{d}{dx}(P)$ is a linear transformation

Subspaces associated to $T_A : \mathbb{R}^n \to \mathbb{R}^m$

$$\begin{aligned} \text{Null}(T_A) &= \{\vec{x} \in \mathbb{R}^n \mid T_A(\vec{x}) = \vec{0}\} \\ &= \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\} \\ &= \text{Null}(A) \end{aligned}$$

$\text{Image}(T_A) = \{T_A(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\}$

$\text{Range}(T_A) = \{A\vec{x} \in \mathbb{R}^m \mid \vec{x} \in \mathbb{R}^n\} = \text{Col}(A) \subseteq \mathbb{R}^m$



Def transpose

Let A be $m \times n$, $A = [a_{ij}]$

The transpose of A is the matrix $A^T = \underbrace{[a_{ji}]}_{n \times m}$

Prop $\underset{m \times n}{A}$

(1) $(A^T)^T = A$

(2) $(A+B)^T = A^T + B^T$

(3) $(AB)^T = B^T A^T$

Ex

$$A = \underbrace{\begin{bmatrix} 1 & 0 & -1 & 3 & 2 \\ 2 & 4 & 2 & 3 & -4 \\ 3 & 2 & -1 & 3 & 2 \\ 0 & 1 & 1 & 1 & -2 \end{bmatrix}}_{4 \times 5} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{rank}(A) = 3 \implies \text{Null}(A) = \left\{ s \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$$

$\mathcal{B}_1 = \{\vec{u}, \vec{v}\}$ basis for Null(A) $\rightarrow$ $\dim(\text{Null}(A)) = 2 \leftarrow$ # of free vars $\rightarrow \text{Null}(A) \subseteq \mathbb{R}^5$

Col(A) $\mathcal{B}_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \\ 1 \end{bmatrix} \right\} \rightarrow \dim(\text{Col}(A)) = 3 = \text{rank}(A), \quad \text{Col}(A) \subseteq \mathbb{R}^4$

Row(A) $\mathcal{B}_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \rightarrow \dim(\text{Row}(A)) = 3 = \text{rank}(A), \quad \text{Row}(A) \subseteq \mathbb{R}^5$

Thm "rank-nullity thm" or dimension thm

Let $\underset{m\times n}{A}$.

(1) $\underbrace{\dim \text{Row}(A)}_{\text{rank}} + \underbrace{\dim \text{Null}(A)}_{\text{nullity}} = n$

(2) $\dim \text{Col}(A) + \dim \text{Null}(A^T) = m$

ie.
$\dim \text{Col}(A) = r$
$\dim \text{Row}(A) = r$
$\dim \text{Null}(A) = n-r$
$\dim \text{Null}(A^T) = m-r$

Thm "1-1 thm"

Let $\underset{m\times n}{A}$, all or none of these are true

(1) Columns of A are linearly indp.

(2) $\text{rank}(A) = n \le m$

(3) $\text{Null}(A) = \{\vec{0}\}$

(4) Pivot in every col

Thm "onto-thm",

Let $\underset{m\times n}{A}$, T.F.A.E.

(1) Rows of A are lin. indp.

(2) $\text{rank}(A) = m \le n$

(3) $\text{Null}(A^T) = \{\vec{0}\}$

(4) Pivot in every row

Chapter 4

Recall, $\vec{u}, \vec{v} \in \mathbb{R}^n$. $\vec{u}\cdot\vec{v} = \vec{u}_1\vec{v}_1 + \cdots + \vec{u}_n\vec{v}_n = [u_1 \cdots u_n]\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u^T v$

Def u, v are orthogonal if $u \cdot v = u^T v = 0 \in \mathbb{R}$

Def Let $W \subseteq \mathbb{R}^n$ be a subspace.
The orthogonal compliment of W in $\mathbb{R}^n$ is
$W^\perp = \{\vec{x} \in \mathbb{R}^n \mid \vec{x}\cdot\vec{y} = 0 \ \forall \ \vec{y} \in W\}$

Ex Let $W = \text{span}\{e_1, e_2\}$ in $\mathbb{R}^3$.

note: $W_1, W_2 \subseteq \mathbb{R}^n$

$W_1^\perp = W_2 \iff W_2^\perp = W_1$

$W^\perp = \text{span}\{e_3\}$

Pf Let $\vec{v}$ be on the z-axis

$\Rightarrow \vec{v} = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}$ for $c \in \mathbb{R}$

And let $\vec{w} \in W$

$\Rightarrow \vec{w} = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \Rightarrow \vec{v} \cdot \vec{w} = v^T w = [0\ 0\ c] \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} = 0 \Rightarrow \vec{v} \in W^\perp \Rightarrow z\text{-axis} \subseteq W^\perp$

Now let $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \in W^\perp \Rightarrow u \cdot y = 0 \quad \forall \ y \in W$

$\Rightarrow u \cdot e_1 = u \cdot e_2 = 0 \Rightarrow u_1 = u_2 = 0 \Rightarrow u = \begin{bmatrix} 0 \\ 0 \\ u_3 \end{bmatrix}$ for some $u_3 \in \mathbb{R} \Rightarrow u \in z\text{-axis}$

$\Rightarrow W^\perp \subseteq z\text{-axis} \Rightarrow W^\perp = \text{span}\{e_3\}$ $\square$

Thm Let $\underset{m \times n}{A}$

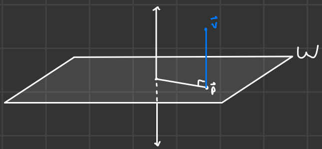
(1) $(\text{Row}(A))^\perp = \text{Null}(A)$

(2) $(\text{Col}(A))^\perp = \text{Null}(A^T)$

Pf Let $\underset{m \times n}{A} = \begin{bmatrix} - r_1 - \\ \vdots \\ - r_m - \end{bmatrix}$, $\vec{x} \in \text{Null}(A) \Rightarrow A\vec{x} = \vec{0}$

4.2 ish - Projection

Given a subspace $W \subseteq \mathbb{R}^n$ and a vector $\vec{v} \in \mathbb{R}^n$, find the vector in W which is closest to $\vec{v}$.



$\vec{p}$ is the "projection of $\vec{v}$ onto W." This can be acomplised with a transform (matrix mult.)

Closest means: $\|\vec{v} - \vec{p}\| \leq \|\vec{v} - \vec{y}\|$ for $\vec{y} \in W$

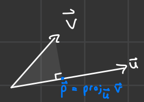
i.e. $\vec{p}$ is the vector which minimized the quantity $\|\vec{v} - \vec{y}\| \ \big| \ \vec{y} \in W$

How to compute $\vec{p} = \text{proj}_W \vec{v}$

Step 1

Project onto a line. Let $\vec{v}, \vec{u} \in \mathbb{R}^n$. The projection of $\vec{v}$ onto $\vec{u}$ is

$$\vec{p} = \text{proj}_{\vec{u}} \vec{v} = \underbrace{\left(\frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}}\right)}_{\in \mathbb{R}} \vec{u}$$



Note:

$$\text{proj}_{c\vec{u}} \vec{v} = \left(\frac{\vec{v}\cdot(c\vec{u})}{(c\vec{u})\cdot(c\vec{u})}\right) c\vec{u} = \frac{c(\vec{u}\cdot\vec{v})}{c^2(\vec{u}\cdot\vec{u})} c\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}} \vec{u} = \text{proj}_{\vec{u}} \vec{v}$$

$$= \text{proj}_L \vec{v} \quad \text{where} \quad L = \text{span}\{\vec{u}\}$$

This can also be accomplished by a matrix multiplication: Let $\vec{u} \in \mathbb{R}^n$, $\vec{u} \neq 0$.

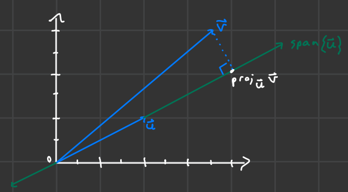
Define $P = P_{\vec{u}} = \underbrace{\frac{1}{\vec{u}\cdot\vec{u}}}_{\mathbb{R}} \underbrace{\underset{n\times 1}{\vec{u}}\,\underset{1\times n}{\vec{u}^T}}_{n\times n} = \frac{\vec{u}\vec{u}^T}{\vec{u}^T\vec{u}}$

P is called the projection matrix associated to $\vec{u}$

$\forall\, \vec{v} \in \mathbb{R}^n$, $P\vec{v} = \text{proj}_{\vec{u}} \vec{v}$

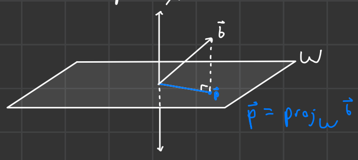
- $P = P^T \Rightarrow P$ is symmetric
- $P^2 = P$
- $I_n - P$ projects onto $L^\perp$
- $\text{Null}(P) = L^\perp$

Ex Let $\vec{v} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \in \mathbb{R}^2$



Step 2

Project onto bigger subspaces. Let $W \subseteq \mathbb{R}^n$ be a subspace and $\vec{b} \in \mathbb{R}^n$

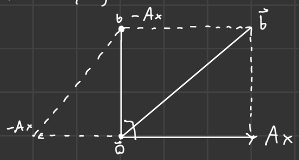


Goal: Find the vector $p \in W$ which is closest to $\vec{b}$.

Suppose $\{a_1, a_2, \dots, a_p\}$ is a basis for $W \subseteq \mathbb{R}^n$. Then every vector in W can be written uniquely as $A\vec{x}$ for same $\vec{x} \in \mathbb{R}^p$ where $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_p \\ | & | & & | \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$,

i.e. trying to find the vector $Ax \in W$ closest to $\vec{b}$.

2D



← $\vec{b} - Ax$ is orthogonal to Ax

$$\Rightarrow \vec{b} - A\vec{x} \in (\mathrm{Col}(A))^{\perp} = \mathrm{Null}(A^T) \Rightarrow A^T\underbrace{(b - Ax)}_{\in \mathrm{Null}(A^T)} = \vec{0} \Rightarrow A^T\vec{b} = A^TAx$$

$$A^TAx = A^Tb$$

"Normal Equations"

Lemma

If the cols. of $\underset{m\times n}{A}$ are lin. indp., then $\underbrace{\underset{n\times m}{A^T}\,\underset{m\times n}{A}}_{n\times n}$ is invertible and symetric

Pf

$n = \mathrm{rank}(A) + \dim \mathrm{Null}(A) \Rightarrow n = n + \underset{0}{\dim \mathrm{Null}(A)} \Rightarrow \mathrm{Null}(A) = \{\vec{0}\}$

Cols of A indp. $\Rightarrow \mathrm{Null}(A) = \{\vec{0}\}$. Suppose $\vec{x} \in \mathrm{Null}(A^TA) \Rightarrow A^TAx = \vec{0}$

$\Rightarrow x^TA^TAx = \vec{0} \Rightarrow (Ax)^TAx = 0 \Rightarrow (Ax)\cdot(Ax) = 0 \Rightarrow \|Ax\| = 0 \Rightarrow A\vec{x} = \vec{0}$

$\Rightarrow \vec{x} \in \mathrm{Null}(A) = \{\vec{0}\} \Rightarrow \vec{x} = \vec{0} \Rightarrow \mathrm{Null}(A^TA) = \{\vec{0}\}$

$\mathrm{Null}(A^TA) = 0 \Rightarrow \underbrace{\dim \mathrm{Null}(A^TA)}_{0} + \mathrm{rank}\,(\underset{n\times n}{A^TA}) = n \Rightarrow A^TA$ is invertible $\square$

Thm A^TA is symetric.

Pf $(A^TA)^T = A^T(A^T)^T = \underline{A^TA}$ $\square$

Thm If cols. of A are indp. and $\vec{b} \in \mathbb{R}^n$, then the unique solution to $A^TAx = A^Tb$ is given by $\underline{\hat{x} = (A^TA)^{-1}A^T\vec{b}}$

$\Rightarrow$ the vector in $\mathrm{Col}(A)$ closest to $\vec{b}$ is $\underline{\mathrm{proj}_W \vec{b} = A\hat{x} = A(A^TA)^{-1}A^T\vec{b}}$

Def Let $W \in \mathbb{R}^n$ be a subspace with basis $\{a_1, a_2, \dots, a_p\}$ and $\vec{b} \in \mathbb{R}^n$. Set $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_p \\ | & | & & | \end{bmatrix}$. The projection matrix associated to W is

$P = A\cdot(A^TA)^{-1}A^T \Rightarrow P\vec{b} = \mathrm{proj}_W \vec{b}$

Ex $W = \text{span}\{e_1, e_2\} \subseteq \mathbb{R}^3$ ← xy-axis $\quad \vec{b} = \begin{bmatrix}2\\2\\1\end{bmatrix}$

P = proj matrix of W

$A = \begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix} \quad A^T A = \begin{bmatrix}1&0\\0&1\end{bmatrix} \quad P = A(A^TA)^{-1}A^T = \begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$

$Pb = \underbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}}_{P}\underbrace{\begin{bmatrix}2\\2\\1\end{bmatrix}}_{b} = \underbrace{\begin{bmatrix}2\\2\\0\end{bmatrix}}_{\text{proj}_W b}$ ✓

4.3 - Least Squares Approximation

Let A be $m \times n$ and $b \in \mathbb{R}^m$ s.t. $[A|b]$ is inconsistent.

i.e. there does not exist any vector $x \in \mathbb{R}^n$ s.t. $Ax = b$

The next best thing is to find approximate solutions

i.e. find $\hat{x} \in \mathbb{R}^n$ s.t. $A\hat{x} \approx b$

i.e. $A\hat{x} \neq b$ but $\vec{e} = b - A\vec{x}$ is as small as possible.

How? Minimize $\|b - A\vec{x}\|$ as $\vec{x}$ ranges over $\mathbb{R}^n$

Def Let $[A|b]$ be an inconsistent system. a least squares solution to this system is a vector $\hat{x} \in \mathbb{R}^n$ s.t. $0 < \|b - A\hat{x}\| \leq \|b - A\vec{x}\| \ \forall \vec{x} \in \mathbb{R}^n$

i.e. $A\hat{x}$ is the vector in $Col(A)$ which is closest to $\vec{b}$

$\Rightarrow A\vec{x} = \text{proj}_W \vec{b}$ ← let's call this $\vec{p}$, where $W = Col(A)$

$\Rightarrow \vec{b} - \vec{p} \in Col(A)^\perp = Null(A^T) \Rightarrow A^T(\vec{b} - \vec{p}) = 0 \Rightarrow A^T\vec{b} = A^T\vec{p}$

$\Rightarrow A^TA\hat{x} = A^Tb$ ← normal equations

Thm

Let $[A|b]$ be an inconsistent system.

$$\underbrace{\left\{\begin{array}{c}\text{Least squares}\\ \text{solutions to}\\ [A|b]\end{array}\right\}}_{\substack{\text{best possible}\\ \text{approximate solutions}}} = \left\{\begin{array}{c}\text{Actual solutions}\\ \text{to the normal}\\ \text{equations}\\ A^TA\hat{x} = A^Tb\end{array}\right\}$$

Application Best fit line!

Ex] Consider the points $(0,6)$ $(1,0)$ and $(2,0)$

We want $f(t) = A + Bt$ st. $f(0) = 6 = A + B\cdot 0$
$f(1) = 0 = A + B$
$f(2) = 0 = A + 2B$

i.e. we want to solve the system

$$\begin{aligned} A &= 6 \\ A + B &= 0 \\ A + 2B &= 0 \end{aligned}$$

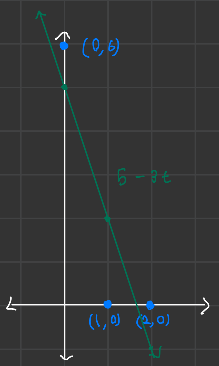
we can't b/c $\left[\begin{array}{cc|c} 1 & 0 & 6 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{array}\right]$ is inconsistent

Let $C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, $b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ $\Rightarrow$ Normal Eqs: $[C^TC \mid C^Tb]$ ← always consistent!

$= \left[\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 5 & 0 \end{array}\right] \xrightarrow{rref} \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -3 \end{array}\right] \Rightarrow$ The unique L.S.S. to $[C \mid b]$ is $\hat{x} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$

$f(t) = 5 - 3t$ is the best fit line

$\Rightarrow C^TC\hat{x} = C^Tb \Rightarrow \hat{x} = (C^TC)^{-1}Cb$



What did we really minimize?
$\|b - Cx\|$ as x varies over $\mathbb{R}$?

Quiz 2

- 20 min
- 4 Fundamental subspaces of A
  - ↳ Bases / dimension
  - ↳ ran-nullity thm → FTLA part 1/2
  - ↳ orthog. compliments $\begin{cases} (\text{Row } A)^\perp = \text{Null}(A) \\ (\text{Col } A)^\perp = \text{Null}(A^T) \end{cases}$

4.4 - Orthonormal bases and GSP

Goal: Given a basis $\mathcal{B}$, replace $\mathcal{B}$ with a better basis where all vectors are orthogonal

Def a set of vectors $\{u_1, \cdots, u_p\}$ in $\mathbb{R}^n$ is called
(1) an orthogonal set if $u_i \cdot u_j = 0$ for all $i \neq j$
(2) an orthonormal set if it is orthogonal and $\|u_i\| = 1$

You can always* normalize an orthogonal set to get an orthonormal one

$$\underbrace{\{u_1, u_2\}}_{\substack{\text{orthogonal} \\ u_1, u_2 \neq 0}} \rightsquigarrow \underbrace{\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}\right\}}_{\text{orthonormal}}$$

Reason 1

If $\{u_1, u_2, \ldots, u_p\}$ is an orthogonal set of non-zero vectors, then $\{u_1, u_2, \cdots, u_p\}$ is a linearly indp set.

Pf

Suppose $c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_p\vec{u}_p = \vec{0}$ (N.T.S. $c_1 = c_2 = \cdots = c_p = 0$)

Notice: $0 = \vec{0} \cdot \vec{u}_1 = (c_1\vec{u}_1 + \cdots + c_p\vec{u}_p) \cdot \vec{u}_1 = c_1(\underbrace{\vec{u}_1 \cdot \vec{u}_1}_{=?}) + c_2(\underbrace{\vec{u}_2 \cdot \vec{u}_1}_{=0}) + \cdots + c_p(\underbrace{\vec{u}_p \cdot \vec{u}_1}_{=0})$

$\Rightarrow 0 = c_1(\vec{u}_1 \cdot \vec{u}_1) \Rightarrow 0 = c_1\|\vec{u}_1\|^2$

Since $\vec{u}_1 \neq 0 \Rightarrow \|u_1\|^2 \neq 0 \Rightarrow c_1 = 0$. Similarly $c_2 = c_3 = \cdots = c_p = 0$ $\square$

Def Let $W \subseteq \mathbb{R}^n$ be a subspace. An orthogonal (respectively orthonormal) basis for W is a basis which is also an orthogonal set (respectively orthonormal)

Ex Std basis for $\mathbb{R}^n$ is an orthonormal basis

$$e_i = \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{n \times 1} \leftarrow i \qquad \|e_i\| = 1 \qquad e_i \cdot e_j = 0 \quad i \neq j$$

Reason 2

Let $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ be an orthogonal basis for $W \leq \mathbb{R}^n$.

For any $\vec{y} \in \mathbb{R}^n$, $\text{proj}_W \vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_p\vec{u}_p$ where $c_j = \frac{\vec{y}\cdot\vec{u}_j}{\vec{u}_j\cdot\vec{u}_j}$ $1 \leq j \leq p$.

If $\vec{y} \in W$, then $\text{proj}_W y = y$, and we get an exact formula for the coefficients of $\vec{y}$ with respect to $\mathcal{B}$

If $\mathcal{B}$ is an orthonormal basis, then all denominators are 1

Reason 3

If $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ is an orthonormal set in $\mathbb{R}^n$, then

$$\underset{n\times p}{Q = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_p \\ | & | & & | \end{bmatrix}}$$ is a really nice matrix

note: a matrix with orthonormal cols. is called a orthogonal matrix

(1) $Q^TQ = \begin{bmatrix} - & u_1 & - \\ - & u_2 & - \\ & \vdots & \\ - & u_p & - \end{bmatrix}\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_p \\ | & | & & | \end{bmatrix} = I_p$

$\|u_1\| = 1 \Rightarrow \|u_1\|^2 = 1 \Rightarrow u_1 u_1 = 1$

(2) In general, $QQ^T \neq I_n$. But if Q is square, then $Q^TQ = I_p \Rightarrow QQ^T = I_p$

$\Rightarrow Q^{-1} = Q^T$ awesome!

(3) Geometric properties

- $\|Q\vec{x}\| = \|\vec{x}\|$ "Q preserves length"
- $(Q\vec{x})\cdot(Q\vec{y}) = \vec{x}\cdot\vec{y}$ "Q preserves angles"

Reason 4 = Reason 3 + Reason 2

Let W be a subspace of $\mathbb{R}^n$ and suppose $\mathcal{B} = \{u_1, u_2, \dots, \cancel{u_p}\}$ is an orthonormal basis for W.

set $\underset{n\times p}{Q = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_p \\ | & | & & | \end{bmatrix}}$ $\Rightarrow$ Then the matrix which projects onto W is

$$P = Q(Q^TQ)^{-1}Q^T = Q(I_p)^{-1}Q^T = QQ^T$$

ie. For any $b \in \mathbb{R}^n$, $\text{proj}_W \vec{b} = QQ^T\vec{b}$

Reason 5

Orthonormal bases always exist!

Gram - Schmidt Process (GSP)

Replaces any basis with an orthonormal one.

$$\text{Basis} \xrightarrow{\text{"orthogonalize"}} \begin{array}{c}\text{Orthog}\\ \text{basis}\end{array} \xrightarrow{\text{"normalize"}} \begin{array}{c}\text{orthonorm}\\ \text{basis}\end{array}$$

$$\text{Basis} \xrightarrow{\text{G.S.P.}} \text{orthonorm basis}$$

Ex] $x_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, x_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}, x_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} \in \mathbb{R}^4$

$\mathcal{B} = \{x_1, x_2, x_3\}$ is a basis for $W = \text{span}\{x_1, x_2, x_3\}$

$x_1 \cdot x_2 = 3 \qquad x_1 \cdot x_1 = 4$

$x_1 \cdot x_3 = 2 \qquad x_2 \cdot x_2 = 3$

$x_2 \cdot x_3 = 2 \qquad x_3 \cdot x_3 = 2$

Step 1

Pick $v_1 = x_1$ set $W_1 = \text{span}\{v_1\} = \text{span}\{x_1\}$

Step 2

Let v_2 be: $v_2 = x_2 - \text{proj}_{W_1} x_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \left(\frac{x_2 \cdot x_1}{x_1 \cdot x_1}\right) x_1 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}$

Step 2.5

If needed scale v_2 to simplify further computations

set $v_2' = 4v_2 = \begin{bmatrix}-3\\1\\1\\1\end{bmatrix}$, set $W_2 = \text{span}\{v_1, v_2'\}$

* So far, we've constructed an orthog. basis for W_2 i.e. $v_1 \cdot v_2' = 0$

Step 3

Let v_3 be: $v_3 = x_3 - \text{proj}_{W_2} x_3 \overset{\text{since } \{v_1, v_2'\} \text{ is an orthog basis}}{=} x_3 - (\text{proj}_{v_1} x_3 + \text{proj}_{v_2'} x_3)$

$= \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \left(\left(\frac{x_3 \cdot v_1}{v_1 \cdot v_1}\right) v_1 + \left(\frac{x_3 \cdot v_2'}{v_2' \cdot v_2'}\right) v_2'\right) = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$

Step 3.5

Scale v_3 to make life easier. $v_3' = 3v_3 = \begin{bmatrix}0\\-2\\1\\1\end{bmatrix}$

Once we've replaced each of the original basis vectors, we're done: $\mathcal{B}' = \{v_1, v_2', v_3'\} = \left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}-3\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\-2\\1\\1\end{bmatrix}\right\}$ is an orthog basis for W.

Step 4

Normalize each vector:

$\mathcal{B}'' = \left\{\frac{v_1}{\|v_1\|}, \frac{v_2'}{\|v_2'\|}, \frac{v_3'}{\|v_3'\|}\right\}$ is an orthonormal basis for W

QR Factorization

If A is an $m\times n$ matrix with indp cols, then there exists an $A = QR$ where Q has orthonormal cols and R is $n\times n$, invertible, upper triangle (with positive entries on the diagonal). Note: R is not the same as in CR.

How? $\underset{m\times n}{A}$

(1) Find basis for $Col(A)$ (all columns form a basis)

(2) Apply GSP to obtain an orthonormal basis for $Col(A)$ $\{u_1, u_2, \dots u_n\}$

(3) For $Q = \underset{m\times n}{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix}}$ $\quad R = Q^TA$

why does this work?

$QR = \underline{QQ^TA} = A$ ← proj matrix for $Col(A)$

$QQ^T \neq I_m$ in general but

$QQ^T\vec{b} = proj_{Col(A)}\vec{b} = \vec{b}$ ← $\vec{b} \in col(A)$

Application

Let A be $m\times n$ w/ indp cols, $A = QR$. $\forall b \in \mathbb{R}^m$ the unique least squares solution to $[A|b]$ is given by $\underline{\hat{x} = R^{-1}Q^T\vec{b}}$

Pf

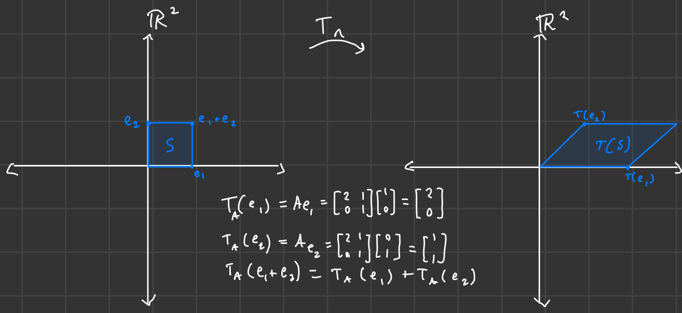
Let $p = proj_W \vec{b}$, $W = col(A)$

$A = QR \Rightarrow A\hat{x} = A(R^{-1}Q^T\vec{b}) = QRR^{-1}Q^Tb = QQ^T\vec{b} = proj_W \vec{b} = p$

Chapter 5 - determinants

Geometric idea

Ex $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ where $x \longmapsto Ax$ $\quad A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$ $\quad T_A(x) = Ax$



Area$(S) = 1$

Area$(T_A(S)) = 2\cdot 1 = 2$

$T_A(e_1) = Ae_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$T_A(e_2) = Ae_2 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$T_A(e_1+e_2) = T_A(e_1) + T_A(e_2)$

Thm

$T: \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation and $A = [T]$.
For any region S in $\mathbb{R}^2$, $\text{Area}(T(S)) = |\det(A)| \cdot \text{Area}(S)$

Minor and Cofactors

Let A be an $n \times n$ matrix with i, j entry a_{ij} $A = [a_{ij}]$

Minor

The i, j minor of A is the submatrix formed by deleting the i^{th} row and j^{th} col. of A

Ex $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$. $(1,1)$ minor of A is $\begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} = M_{11}$

$(2,3)$ minor of A is $\begin{bmatrix} 1 & 5 \\ 0 & -2 \end{bmatrix} = M_{23}$

Cofactor

The (i,j) cofactor of A is $C_{ij} = (-1)^{ij} \det(M_{ij})$

Ex Some A = previous A

$(2,1)$ Cofactor of A: $C_{21} = (-1)^{2+1} \det(M_{21})$
$= -\det\left(\begin{bmatrix} 5 & 0 \\ -2 & 0 \end{bmatrix}\right) = \boxed{0}$

$(3,3)$ cofactor of A: $C_{33} = (-1)^{3+3} \det\left(\begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}\right) = \boxed{-6}$

Def

Let $A = [a_{ij}]_{m \times n}$

will always be equal

Cofactor expansion along row i: $a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in}$

Cofactor expansion down col. j: $a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + a_{nj} C_{nj}$

Ex $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$ C.E. across row 2

$a_{21} C_{21} + a_{22} C_{22} + a_{23} C_{23}$

$= 2(-1) \underbrace{\begin{vmatrix} 5 & 0 \\ -2 & 0 \end{vmatrix}}_{0} + 4(1) \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}}_{0} - 1(-1) \underbrace{\begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix}}_{-2} = -2$

C.E. down Col 3

$= 0 \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} - 1(-1) \underbrace{\begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix}}_{-2} + 0 \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} = -2$

Fact

For an $n \times n$ matrix A cofactor expansion across any row or down any column always gives the same result

Def this value is $\det(A)$

Ex $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ -1 & 2 & 1 \end{bmatrix}$ $\det(A) = 1\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} + (-1)\begin{vmatrix} 2 & 3 \\ 3 & 2 \end{vmatrix} = 1(-1) - 2(-4) - (-5)$

$= 12$

Ex $B = \begin{bmatrix} 2 & 3 & 7 \\ 0 & -1 & 8 \\ 0 & 0 & 4 \end{bmatrix}$ $\det(B) = 2\begin{vmatrix} -1 & 8 \\ 0 & 4 \end{vmatrix} = 2(-4) = -8$

Thm If A is "triangular" upper, lower, or diagonal
then $\det(A)$ = product of the diagonal entries

Det and Row operations

Let A be $n \times n$ and $k \in \mathbb{R}$

Row op.		Resulting det
swap $R_i \leftrightarrow R_j$	$\longrightarrow$	$-\det(A)$
scale $k \cdot R_i$	$\longrightarrow$	$k \det(A)$
replace $kR_i + R_j \to R_j$	$\longrightarrow$	$\det(A)$ $\longleftarrow$ nice!

Ex $A = \begin{bmatrix} 2 & 10 & 4 \\ 3 & 13 & 3 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \underbrace{\begin{bmatrix} 1 & 5 & 2 \\ 3 & 13 & 3 \\ 0 & 1 & 2 \end{bmatrix}}_{A_1} \xrightarrow[R_2]{-3R_1+R_2} \underbrace{\begin{bmatrix} 1 & 5 & 2 \\ 0 & -2 & -3 \\ 0 & 1 & 2 \end{bmatrix}}_{A_2} \xrightarrow{R_2 \leftrightarrow R_3} \underbrace{\begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -3 \end{bmatrix}}_{A_3} \xrightarrow{2R_2+R_3 \to R_3} \underbrace{\begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{A_4}$

$\det(A_4) = (1)(1)(1) = 1 \Rightarrow \det(A_3) = \det(A_4) = 1 \Rightarrow \det(A_2) = -\det(A_3) = -1$

$\Rightarrow \det(A_2) = \det(A_1) = -1 \Rightarrow \det(A_1) = \frac{1}{2}\det(A) \Rightarrow \det(A) = -2$

More properties of det

Let A, B be $n \times n$ $k \in \mathbb{R}$

(1) $\det(A^T) = \det(A)$

(2) $\det(k \cdot A) = k^n \cdot \det(A)$

(3) $\det(A \cdot B) = \det(A) \cdot \det(B)$

(4) If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$

Pf assuming (3)

Since A is invertible, $A \cdot A^{-1} = I_n$. $\det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$

$\det(I_n) = \begin{vmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{vmatrix} = 1 \Rightarrow \det(A) \cdot \det(A^{-1}) = 1$ $\square$ Notice: $\det(A) \neq 0$

Thm $n\times n$ A is invertible iff $\det(A) \neq 0$

PF A invertible $\Rightarrow \det(A) \neq 0$ ✓

Assume $\det(A) \neq 0$

$\underbrace{E_s \cdots E_2 E_1}_{s \text{ total steps}} A = \text{rref}(A) \Rightarrow \det(E_s \cdots E_2 E_1 A) = \det(\text{rref}(A))$

$\Rightarrow \underbrace{\det(E_s)}_{\neq 0} \cdots \underbrace{\det(E_2)}_{\neq 0} \underbrace{\det(E_1)}_{\neq 0} \det(A) = \det(\text{rref}) \Rightarrow \det(\text{rref}(A)) \neq 0$

$\Rightarrow \text{rref}(A) = I_n \Rightarrow A$ is invertible!

note: $E_s \cdots E_2 E_1 = A^{-1}$

Cramer's Rule

Solve $Ax = b$ using determinants.

$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}, \quad b \in \mathbb{R}^n$$

notation
$$\underset{1 \le i \le n}{A_i(\vec{b})} = \begin{bmatrix} | & | & & | & & | \\ a_1 & a_2 & \cdots & \vec{b} & \cdots & a_n \\ | & | & & | & & | \end{bmatrix}$$

$\vec{b}$: ith col of A

Thm

If A is $n\times n$ invertible and $b \in \mathbb{R}^n$, then the unique solution to $Ax = b$ is given by $x_i = \dfrac{\det(A_i(\vec{b}))}{\det A}$

Ex $A = \begin{bmatrix} 3 & 7 \\ 1 & 2 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$\det \neq 0$
$\Rightarrow A$ invertible

solve $Ax = b$

$$x_1 = \frac{\det(A_1(b))}{\det A} = \frac{\begin{vmatrix} 3 & 7 \\ 4 & 2 \end{vmatrix}}{\begin{vmatrix} 3 & 7 \\ 1 & 2 \end{vmatrix}} = 22$$

$$x_2 = \frac{\begin{vmatrix} 3 & 3 \\ 1 & 4 \end{vmatrix}}{\begin{vmatrix} 3 & 7 \\ 1 & 2 \end{vmatrix}} = -9 \quad \Rightarrow \quad x = \begin{bmatrix} 22 \\ -9 \end{bmatrix}$$

Let $A = [a_{ij}]^{n\times n}$, $C = [c_{ij}]$ where c_{ij} is the $(i,j)^{th}$ cofactor of A.

C is the matrix of cofactors of A

Def then $C^T = [c_{ji}]$ is called

- adjoint of A $\quad \text{adj}(A)$
- Classical adjoint of A
- adjugate

$\text{adj}(A) = C^T$

$$\underset{3\times 3}{\begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix}} \underset{3\times 3}{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}} = \begin{bmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{bmatrix}$$

Thm Let A be $n \times n$, and let $\mathrm{adj}(A)$ be the adjoint of A. Then

$$\mathrm{adj}(A) \cdot A = \det(A) \cdot I_n = A \cdot \mathrm{adj}(A)$$

If A is invertible, $\quad A \cdot \left(\frac{1}{\det(A)} \mathrm{adj}(A) \right) = I_n$

$\det A \neq 0$

Cor If A is $n \times n$ invertible, then $A^{-1} = \frac{1}{\det(A)} \mathrm{adj}(A)$

Ex $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$C_{11} = (-1)^{1+1} \det([d]) = d$ $\quad$ $C_{21} = (-1)^{2+1} \det([b]) = -b$

$C_{12} = (-1)^{1+2} \det([c]) = -c$ $\quad$ $C_{22} = a$

$\mathrm{adj}(A) = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \Rightarrow$ if A is invertible $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Chapter 6

Recall, a matrix A is called diagonal if $a_{ij} = 0$ for all $i \neq j$

Ex 0 matrix ✓ $\qquad I_n$ ✓

Ex $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax \\ by \end{bmatrix}$

Most transformations are not so simple, but some are this simple "with respect to the right basis."

Ex $T = T_A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$. $u = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$T(u) = Au = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$

"T is not diagonal"

$T(v) = Av = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2v$

$T(w) = Aw = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = w$

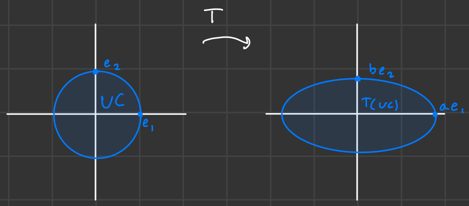
Def Let V be a vector space an $\mathcal{B} = \{ b_1, \dots, b_2 \}$ a basis for V. For any $\vec{v} \in V$, there are unique coefficients $c_1, c_2, \dots, c_p$ s.t. $\vec{v} = c_1 \vec{b}_1 + \dots + c_p \vec{b}_p$

$c_1, \dots, c_p$ are called the coordinates of $\vec{v}$ w.r.t. $\mathcal{B}$.

$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$ is the coordinate vector of $\vec{v}$ w.r.t. $\mathcal{B}$

By the way

$T:\mathbb{R}^2 \to \mathbb{R}^2$ given by $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} ax \\ by \end{bmatrix}$



$$\text{area(ellipse)} = \left|\det[T]\right| \cdot \text{area(UC)} = ab\pi$$

Let $u = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. If we view $\mathbb{R}^2$ w.r.t. $\mathcal{B} = \{v, w\}$, look what happens; coordinate vector of u w.r.t. $\mathcal{B}$:

$$\left[\begin{array}{cc|c} v & w & u \end{array}\right] \Rightarrow \qquad u = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow [u]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Coordinate vector of $T_2(u) = \begin{bmatrix} -5 \\ -1 \end{bmatrix} = -4\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow [T_2(u)]_{\mathcal{B}} = \begin{bmatrix} -4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2(-2) \\ 1(3) \end{bmatrix}$

Def Let A be $n \times n$.

Suppose $\exists$ a non-zero vector $\vec{x} \in \mathbb{R}^n$ s.t. $A\vec{x} = \lambda \vec{x}$ for some λ.

The scalar $\lambda \in \mathbb{R}$ is called an eigenvalue of A

The vector $\vec{x}$ is called an eigenvector of A corresponding to λ

Ex $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $w = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$Av = 2v$

v is an eigen vector corresponding $\lambda = 2$

$\lambda = 2$ is an eigen value of A

$Aw = 1w$

w is an eigen vector of A corresponding to $\lambda = 1$

$\lambda = 1$ is an eigen value of A

Two Questions

(1) How do we find eigen values of A?

(2) Once we have eigen values, how do we find eigen vectors corresponding to those eigenvalues?

(2) suppose A is $n \times n$, $\lambda \in \mathbb{R}$.

λ is an eigenvalue of A $\iff$ $A\vec{x} = \lambda \vec{x}$ for some $x \neq 0$ in $\mathbb{R}^n$

Notice: $A\vec{x} = \lambda\vec{x}$ $\iff$ $A\vec{x} - \lambda\vec{x} = \vec{0}$

$\iff A\vec{x} - (\lambda I_n)\vec{x} = \vec{0}$

$\iff (A - \lambda I_n)\vec{x} = \vec{0}$

$\iff \vec{x} \in Null(A - \lambda I_n)$

We call $Null(A - \lambda I_n)$ the eigen space of A corresponding to λ.

E_λ ↘ $Null(A - \lambda I_n)$ = set of all eigen vectors corresponding to λ and $\vec{0}$

(1) λ is an eigenvalue of A $\iff$ $A\vec{x} = \lambda\vec{x}$ for some $\vec{x} \neq 0$

$\iff (A - \lambda I_n)\vec{x} = \vec{0}$ $\quad \vec{x} \neq 0$

$\iff Null(A - \lambda I_n) \neq \{\vec{0}\}$

$\iff A - \lambda I_n$ is not invertible

$\iff \underbrace{\det(A - \lambda I_n)}_{\text{characteristic polynomial}} = 0$

How do we find such λ?

Treat λ like a variable and solve $\det(A - \lambda I_n) = 0$

Ex $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$ $\qquad A - \lambda I_2 = \begin{bmatrix} 3-\lambda & -2 \\ 1 & -\lambda \end{bmatrix}$

$\det(A - \lambda I_2) = (3-\lambda)(-\lambda) + 2$

$= -3\lambda + \lambda^2 + 2$

$= (\lambda - 2)(\lambda - 2)$

Ex $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$ $\qquad$ Find eigen values of A

Treat λ like a variable

$A - \lambda I_2 = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$

$\det(A - \lambda I_2) = (2-\lambda)(-6-\lambda) - 9 = -12 + 4\lambda + \lambda^2 - 9$

$= \lambda^2 + 4\lambda - 21$ ⟵ Find roots!

$= (\lambda + 7)(\lambda - 3) \Rightarrow \quad \lambda_1 = -7, \ \lambda_2 = 3$

* $\lambda \in \mathbb{R}$ is an eigen value of A iff λ is a root of the characteristic polynomial.
* If $\lambda = 0$ is an eigenvalue of A, $\exists\, \vec{x} \neq \vec{0} \in \mathbb{R}^n$ s.t. $A\vec{x} = \lambda\vec{x} = \vec{0}$

$\Rightarrow \exists\, \vec{x} \neq \vec{0}$ in $Null(A)$ $\Rightarrow$ A is not invertible

Claim: A is $n \times n$ $Null(A) \neq \{\vec{0}\}$ then A is not invertible

Since $Null(A) \neq 0 \Rightarrow \dim Null(A) > 0 \Rightarrow rank(A) < n$
$\Rightarrow A$ is not invertible rank nullity theorem all the way!

Alt $\lambda = 0$ is an eigenvalue of A
$\Leftrightarrow \det(A - \underset{0}{\lambda} I_n) = 0$
$\Leftrightarrow \det(A) = 0$
$\Leftrightarrow A$ is not invertible

Thm An $n \times n$ matrix A is invertible iff $\lambda = 0$ is not an eigen value of A.

Def Let $\underset{n\times n}{A}$

(1) Suppose $\det(A - \lambda I_n) = (\lambda - r)^m p(\lambda)$ where $r \in \mathbb{R}$, $m \geq 1$, $p(\lambda)$ poly in λ

Then $\lambda = r$ is an eigenvalue of A with algebraic multiplicity m.

Ex) $\det(A - \lambda I_n) = (\lambda + 7)^2 (\lambda + 3)^3$

Then the eigenvalues of A are $\underbrace{\lambda_1 = -7, \lambda_2 = -7,}_{-7 \text{ has alg. mult. } 2} \underbrace{\lambda_3 = 3, \lambda_4 = 3, \lambda_5 = 3}_{3 \text{ has alg. mult. } 3}$

(2) Let λ be an eigenvalue of A

The geometric multiplicity of λ is $\dim Null(A - \lambda I_n)$.

i.e. geometric mult of $\lambda = \dim E_\lambda$

* an eigenvalue can have multiple eigen vectors, not vice versa

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

step 1 $|A - \lambda I_3| = \begin{vmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{vmatrix} = (4-\lambda)\begin{vmatrix} 1-\lambda & 6 \\ -1 & 8-\lambda \end{vmatrix} - (2)\begin{vmatrix} -1 & 6 \\ -1 & 8-\lambda \end{vmatrix} + (2)\begin{vmatrix} -1 & 6 \\ -1 & 8-\lambda \end{vmatrix}$

$= \lambda^3 - 13\lambda^2 + 40\lambda - 36 = (\lambda - 9)(\lambda - 2)^2$

$\Rightarrow \lambda_1 = 9$ is an eigenvalue of A w/ alg. mult. 1

$\Rightarrow \lambda_2 = 2$ is an eigenvalue of A w/ alg. mult. 2

step 2 Compute Null $(A - 9I_3)$

$\left[\begin{array}{ccc|c} -5 & -1 & 6 & 0 \\ 2 & -8 & 6 & 0 \\ 2 & -1 & -1 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow$ A basis for $\mathcal{E}_{\lambda_1}$ is $\mathcal{B}_{\lambda_1} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

every eigenvector of A corresponding to λ_1 is a non-zero multiple of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$\dim \mathcal{E}_{\lambda_1} = \dim \text{Null}(A - 9I_3) = 1$

$\Rightarrow \lambda_1 = 9$ has geometric mult. 1

Compute Null$(A - 2I_3)$

$[A - 2I_3 \mid 0] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & -1/2 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow$ geomult of $\lambda_2 = 2$ is 2

$\mathcal{B}_{\lambda_2} = \left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathcal{E}_{\lambda_2}$

i.e every eigenvec. of A corresponding to $\lambda_2 = 2$ is of the form $\vec{x} = s\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$, $\vec{x} \neq 0$

Two cool theorems

Could have repeats

Thm Let $A = [a_{ij}]$ be $n \times n$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

Then (1) $\det(A) = \prod_{k=1}^{n} \lambda_k$

(2) $\text{trace}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$

Ex $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\quad \det(A - \lambda I_2) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc$

$= \lambda^2 - (a+d)\lambda + ad - bc$

$p(\lambda)$ characteristic polynomial of 2×2 A $= \underline{\lambda^2 - \text{trace}(A)\lambda + \det(A)}$

Suppose $\lambda_1 = r$ and $\lambda_2 = s$ are eigenvals of A

$\Rightarrow p(\lambda) = (\lambda - r)(\lambda - s) = \lambda^2 - (r+s)\lambda + rs$

$\Rightarrow \lambda^2 - (r+s)\lambda + rs = \underline{\lambda^2 - \text{trace}(A)\lambda + \det(A)}$

Ex $A = \begin{bmatrix} 2 & 7 & 5 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{bmatrix}$ $\quad \det(A) = 2 \cdot 3 \cdot 4 \quad$ trace$(A) = 2 + 3 + 4$

$|A - \lambda I_3| = \begin{vmatrix} 2-\lambda & 7 & 5 \\ 0 & 3-\lambda & 6 \\ 0 & 0 & 4-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda)(4-\lambda) \Rightarrow$ eigvals: 2, 3, 4

Thm If $v_1, v_2, \ldots, v_p$ are eigvecs corresponding to distinct eigvals $\lambda_1, \lambda_2, \ldots, \lambda_p$, then $\{v_1, \cdots, v_p\}$ is indp.

$\lambda_i \neq \lambda_j$ for $i \neq j$

$p=1$: ✓

$p=2$: $Av_1 = \lambda_1 v_1$ $\quad \lambda_1 \neq \lambda_2$ $\quad$ suppose $c_1\vec{v}_1 + c_2\vec{v}_2 = 0$

$A v_2 = \lambda_2 v_2$

$\Rightarrow A(c_1 v_1 + c_2 v_2) = \vec{0} \quad \Rightarrow c_1 A v_1 + c_2 A v_2 = \vec{0}$

$\Rightarrow c_1 \lambda_1 \vec{v}_1 + c_2 \lambda_2 \vec{v}_2 = \vec{0} \quad \Rightarrow c_2 \lambda_2 \vec{v}_2 - c_2 \lambda_1 \vec{v}_2 = 0$

$\Rightarrow c_2(\lambda_2 - \lambda_1)\vec{v}_2 = \vec{0} \quad \Rightarrow c_2 = 0$. similarly, $c_1 = 0$

$\Rightarrow \{v_1, v_2, \ldots, v_p\}$ is lin indp.

Def Two matrices $\underset{n\times n}{A}$ and $\underset{n\times n}{B}$ are called similar if there exists an invertible matrix P s.t. $B = PAP^{-1} \Leftrightarrow P^{-1}BP = A$

Notation: A is similar to B $\rightsquigarrow$ $A \sim B$

$A \sim B \Leftrightarrow B \sim A$

Ex $A \sim A$ ie. $A = I_n A I_n^{-1}$ reflexive

Ex $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ $\qquad B = PAP^{-1} \Leftrightarrow BP = PA$

$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{P}\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}}_{A} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = BP = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

why do we care?

Thm If $\underset{n\times n}{A} \sim \underset{n\times n}{B}$,

then A and B have the same eigenvalues! same char poly!

Pf Suppose $\underset{n\times n}{A} \sim \underset{n\times n}{B}$

$\Rightarrow \exists \underset{n\times n}{P}$ invertible s.t. $A = P^{-1}BP$

$\det(A - \lambda I_n) = \det(P^{-1}BP - \lambda I_n) = \det(P^{-1}BP - \lambda P^{-1}P) = \det(P^{-1}(B - \lambda I_n)P)$

$= \det(P^{-1}) \cdot \det(B - \lambda I_n) \cdot \det(P) = \det(B - \lambda I_n)$ $\square$

When is A similar to a diagonal matrix?

Ex $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ claim: $A \sim D$

Let $P = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. $\quad D = P^{-1}AP \iff PD = AP$

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

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

$$\Rightarrow P^{-1}\underbrace{\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{P} = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}_{D}$$

eig. vecs. of A: $\lambda = 2$, $\lambda = 1$

eigvals of A

Def $\underset{n\times n}{A}$ is diagonalizable if A is similar to a diagonal matrix.

i.e. if $\exists$ diagonal D and invertible P s.t. $P^{-1}AP = D$ "P diagonalizes A"

Note: Suppose $\underset{n\times n}{A} \sim D_{\text{diagonal}} \Rightarrow D = P^{-1}AP \iff PD = AP$

$$P = \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}}_{\text{eigenvectors of } A}, \quad D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \Rightarrow PD = \begin{bmatrix} | & | & & | \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \\ | & | & & | \end{bmatrix}$$

$$AP = \begin{bmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_n \\ | & | & & | \end{bmatrix} \Rightarrow Av_1 = \lambda_1 v_1,\ Av_2 = \lambda_2 v_2,\ \ldots,\ Av_n = \lambda_n v_n$$

If A is diagonalizable, then $P^{-1}AP = D$ $\quad P = \begin{bmatrix} \text{eigen-} \\ \text{vectors} \end{bmatrix}$, $D = \begin{bmatrix} \text{eigen-} \\ \text{values} \end{bmatrix}$

Thm $\underset{n\times n}{A}$ is diagonalizable iff there exists n linearly independent eigenvecs.

i.e. iff $\exists$ a basis for $\mathbb{R}^n$ of eigenvectors of A.

Warning:

$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ $\quad A \not\sim B$

For $\lambda = 2$: Compute $\mathcal{E}_{\lambda=2} = \text{Null}(A - 2I_2)$

$[A - 2I_2] = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \dim \mathcal{E}_\lambda = 1 \Rightarrow$ can only find 1 indp. eigen vector

$\Rightarrow A$ is not diagonalizable. $\quad$ alg mult of $\lambda = 2$ $= 2 \neq 1 =$ geometric mult of $\lambda = 2$

Suppose $\overset{\text{indp cols}}{A}x = b$ is inconsistent Goal: Solve L.S.S. using QR

Suppose $\underset{m\times n}{A} = \underset{m\times n}{Q}\underset{n\times n}{R}$ $R = Q^TA$ $QQ^TA = A$

orthonormal columns invertible

$\exists!\ \hat{x} = R^{-1}Q^T\vec{b}$

Finding $P^{-1}AP = D$

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

Step 1: Find eigenvalues

$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = -1$

AM=1 AM=2 AM=1

Step 2: Find eigenvectors

$\lambda_1 \Rightarrow \left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\right\}, \quad \lambda_2 \Rightarrow \left\{\begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}3\\0\\1\\0\end{bmatrix}\right\}, \quad \lambda_3 \Rightarrow \left\{\begin{bmatrix}3/2\\0\\-1/3\\1\end{bmatrix}\right\}$

Step 3: Form P and D

$P = \begin{bmatrix} 1 & 2 & 3 & -3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ $P^{-1} = \begin{bmatrix} 1 & -2 & -3 & 3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Ex $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix} \Rightarrow \lambda_1 = 1, \ \underset{AM=2}{\lambda_2 = 5}$

For $\lambda_2 = 5$, $[A - 5I_3] = \begin{bmatrix} -4 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (1 free var) $\Rightarrow \dim \mathcal{E}_{\lambda_2} = 1$ GM = 1 < AM = 2 !

A is not diagonalizable

Summary A is $n \times n$

(1) For any λ of A, $0 < GM \text{ of } \lambda \leq AM \text{ of } \lambda$

(2) If A has n distinct eigenvalues, then A is diagonalizable

(3) A is diagonalizable iff GM of λ = AM of λ $\forall$ eigenvalues λ of A

(4) Eigenvectors corresponding to distinct eigenvalues need not be orthogonal

ex $\begin{bmatrix}1\\0\\0\\0\end{bmatrix} \cdot \begin{bmatrix}2\\1\\0\\0\end{bmatrix} \neq 0$

$A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ $\lambda_1 = a+b, \ \lambda_2 = a-b \Rightarrow |A - \lambda I_2| = \begin{vmatrix} a-\lambda & b \\ b & a-\lambda \end{vmatrix} = (a-\lambda)^2 - b^2 = \cdots$

$= (\lambda - (a+b))(\lambda - (a-b))$. $[A - (a+b)I_2] = \begin{bmatrix} -b & b \\ b & -b \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$. $\mathcal{E}_{\lambda_1} = \text{spn}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$

$[A - (a-b)I_2] = \begin{bmatrix} b & b \\ b & b \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. $\mathcal{E}_{\lambda_2} = \text{spn}\left\{\begin{bmatrix}1\\-1\end{bmatrix}\right\}$

For any $\vec{x} \in \mathbb{R}$, $T(\vec{x}) = [T]\vec{x}$

More generally: If $\mathcal{B} = \{b_1, b_2, \dots, b_n\}$ is a basis for $\mathbb{R}^n$, $T: \mathbb{R}^n \to \mathbb{R}^n$ is a linear trans, then the $\mathcal{B}$ matrix of T is

$$[T]_{\mathcal{B}} = \begin{bmatrix} | & | & & | \\ [T(b_1)]_{\mathcal{B}} & [T(b_2)]_{\mathcal{B}} & \cdots & [T(b_n)]_{\mathcal{B}} \\ | & | & & | \end{bmatrix}$$

Recall if $x \in \mathbb{R}^n$, $\mathcal{B} = \{b_1, \dots, b_n\}$

$\exists! \ \vec{x} = c_1 b_1 + c_2 b_2 + \cdots + c_n b_n \Rightarrow [x]_{\mathcal{B}} = [$

$[T]_{\mathcal{B}}$ satisfies: $[T(\vec{x})]_{\mathcal{B}} = [T]_{\mathcal{B}} [x]_{\mathcal{B}}$

Ex) $T: \mathbb{R}^2 \to \mathbb{R}^2$ given by $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 3x - 2y \\ x \end{bmatrix}$

with respect to the standard basis:

Std matrix of T: $T(e_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$, $T(e_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$

$[T] = \begin{bmatrix} | & | \\ T(e_1) & T(e_2) \\ | & | \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$ "A"

The respect to a basis of eigen vecs

$\mathcal{B} = \{\vec{v}, \vec{w}\} = \left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} \Rightarrow T(v) = T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2v$

$T(w) = T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = w$

$T(b_1) = 2b_1 = 2b_1 + 0b_2 \Rightarrow [T(b_1)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$T(b_2) = b_2 = 0b_1 + 1b_2 \Rightarrow [T(b_2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

So, the $\underbrace{\mathcal{B}\text{-matrix}}_{\text{basis of eigen vecs}}$ of T is: $[T]_{\mathcal{B}} = \begin{bmatrix} | & | \\ [T(b_1)]_{\mathcal{B}} & [T(b_2)]_{\mathcal{B}} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ diagonal!

* with the right choice of basis, T can be represented by a diagonal matrix.

For any $x \in \mathbb{R}^2$, $[T(x)]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} [x]_{\mathcal{B}}$

ex $x = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \in \mathbb{R}^2$ $\quad T(x) = \begin{bmatrix} 3(3) - 2(2) \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ or $\underbrace{\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}}_{[T]} \underbrace{\begin{bmatrix} 3 \\ 2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 5 \\ 3 \end{bmatrix}}_{T(x)}$ ✓

$[x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$[T(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 5 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{[T(x)]_{\mathcal{B}}} = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}_{[T]_{\mathcal{B}}} \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{[x]_{\mathcal{B}}}$

$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} P^{-1} \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{P}$

Symmetric Matrices

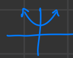
Def $\underset{m\times n}{A} = A^T \qquad [a_{ij}] = [a_{ji}]$

Ex Let A be any matrix $m\times n$

$(A^TA)^T = A^T(A^T)^T = A^TA$ ($n\times n$)

$(AA^T)^T = AA^T$ ($m\times m$)

Warning: Eigenvalues can be complex numbers

$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\det(A - \lambda I_2) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 \leftarrow$ no real roots 

$\Rightarrow \lambda = \pm i$ are evals of A. $\mathbb{C} = \{a+bi \mid a,b \in \mathbb{R}\}$ note: $i^2 = -1$

① $(a+bi) + (c+di) = (a+c) + (b+d)i$

② ex/ $(bi)(di) = bd\,i^2 = -bd$

$i\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} i \\ i \end{bmatrix}$, $i\begin{bmatrix} i \\ i \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

Main Facts Let $\underset{n\times n}{A} = A^T$

(1) All eigenvals of $A \in \mathbb{R}$ i.e. no complex (n total)

(2) A is <u>orthogonally diagonalizable</u>

i.e. A is diagonalizable and eigenvecs can all be chosen to be orthog. to eachother

i.e. we can find an orthonormal basis of eigenvectors of A for $\mathbb{R}^n$

i.e. $D = Q^{-1}AQ = Q^TAQ$

Ex $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ $a, b \in \mathbb{R}$

$\lambda = a \pm b$, $A = A^T$ $\qquad \underset{\lambda_1 = a+b}{v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \qquad \underset{\lambda_2 = a-b}{v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}}$

$\tilde{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\tilde{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Then $Q = \begin{bmatrix} | & | \\ \tilde{v}_1 & \tilde{v}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ is an <u>orthogonal matrix</u>

$\Rightarrow D = Q^{-1}AQ \Rightarrow \begin{bmatrix} a+b & 0 \\ 0 & a-b \end{bmatrix} = Q^TAQ$

Thm $\underset{n\times n}{A}$ is symetric iff A is orthogonally diagonalizable.

Pf [$\Leftarrow$] $\exists$ D and Q s.t. $D = Q^TAQ$

$\Rightarrow QD = QQ^TAQ \quad (Q^T = Q^{-1})$

$\Rightarrow QD = AQ \Rightarrow QDQ^T = AQQ^T = A \Rightarrow A = QDQ^T$

$\Rightarrow A^T = (QDQ^T)^T = (Q^T)^TD^TQ^T = QDQ^T = A$ ✓

[$\Rightarrow$] $Av_1 = \lambda_1 v_1$, $Av_2 = \lambda_2 v_2$ w/ $\lambda_1 \neq \lambda_2$

Claim: $v_1 \cdot v_2 = 0$. $\lambda_1(v_1\cdot v_2) = \lambda_1 v_1^T v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2$

$= v_1^T(Av_2) = v_1^T(\lambda_2 v_2) = \lambda_2(v_1^Tv_2) = \lambda_2(v_1\cdot v_2)$

$\Rightarrow \lambda_1(v_1\cdot v_2) = \lambda_2(v_1\cdot v_2) \Rightarrow \underbrace{(\lambda_1-\lambda_2)}_{\in\mathbb{R},\ \neq 0}\underbrace{(v_1\cdot v_2)}_{\in\mathbb{R}} = 0 \Rightarrow v_1\cdot v_2 = 0$ $\square$

So, for any $A = A^T$, $\exists$ a diagonal D and orthogonal matrix Q s.t.

$D = Q^TAQ \iff A = QDQ^T \iff$ $\exists$ an orthogonal basis of eigenvecs of A for $\mathbb{R}^n$ $\iff$ Eigenspaces of $A = A^T$ are fundamentally orthogonal space

Energy

Symetric matrices also define "quadratic form."

Let $\underset{n\times n}{A} = A^T$. The Energy of A is a function $E_A : \mathbb{R}^n \to \mathbb{R}$

by $E(\underset{n\times 1}{\vec{x}}) = \underset{1\times n}{x^T}\underset{n\times n}{A}\underset{n\times 1}{x}$ (not a linear transformation)

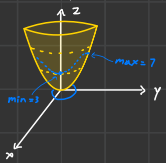
E_A also gets called a quadratic form

quadratic polynomial in 2 variables

↙ Parabaloid

Ex $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ $\quad E_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x+3y & 3x+2y \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 + 6xy + 2y^2$

Ex $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$ $\quad E_A(x,y) = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 3x^2 + 7y^2$



Notice $E_A(x,y)$ is unbounded so no max. But we can ask for a "constraind max."

Set $M_A = \max\{E_A(\vec{x}) \mid \|\vec{x}\| = 1\}$ i.e. if we we only allow for $\|x\| = 1$ what is the biggest output $E_A(X)$.

$m_A = \min\{E_A(\vec{x}) \mid \|\vec{x}\| = 1\}$

Thm $A = A^T$, $E_A(x) = x^T A x$, m_A, M_A

Then ① M_A = largest λ, ② m_A = smallest λ

Pf Since A is symetric, you can orthogonally diaganalize it.

i.e $\exists\, Q, D$ s.t. $A = QDQ^T$

$E_A(x) = x^T A x = x^T Q D Q^T x = (Q^T x)^T D Q^T x = y^T D y = E_D(y)$

In particular, $M_A = M_D$, $m_A = m_D$

Ex $n = 3$ suppose $a \geq b \geq c$ are eigenvalues of A

$\Rightarrow D = \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix}$, set $Q = \begin{bmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{bmatrix}$

For any $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$, $ay_1^2 = ay_1^2$, $by_2^2 \leq ay_2^2$, $cy_3^3 \leq ay_3^2$

If $\|y\| = 1$, then $E_D(y) = y^T D y$

$= ay_1^2 + by_2^2 + cy_3^2$

$\leq ay_1^2 + ay_2^2 + ay_3^2$

$= a(y_1^2 + y_2^2 + y_3^2)$

$= a\|y\|^2 = a$

Furthermore, $E_D(\vec{e}_1) = a$

$E_D(y) = ay_1^2 + by_2^2 + cy_3^2 = a$

$a = M_A$

?

Positive definite / semidefinite matrices

Def Let $S = S^T$ symetric

(1) S is called positive definite (PD) if all eigenvals. of S are positive

(2) S is positive semidefinite (PSD), if all eigenvals are ≥ 0

Ex $\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ PD $\quad \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ PSD $\quad \begin{bmatrix} 2 & -3 \\ -3 & 2 \end{bmatrix}$ $\lambda_1 = 2 - 3 = -1 < 0$, $\lambda_2 = 2 + 3 = 5$ neither

Let $S = S^T$, S is PSD iff $E_S(x) \geq 0 \quad \forall\, x \in \mathbb{R}^n$

i.e. if the minimum of $E_S(x)$, m_S, $\geq 0 \iff E_S(x) \geq 0$

"Test 3" S is PSD iff $S = A^T A$ for some A

Pf $[\Leftarrow]$ suppose $S = A^T A$ for some $\underset{m\times n}{A}$

Energy test: Compute $E_S(x)$

Let $x \in \mathbb{R}^n$. $E_S(x) = x^T S x = x^T A^T A x = (Ax)^T A x$

$= (Ax)\cdot(Ax) = \|Ax\|^2 \geq 0 \Rightarrow S$ is PSD! $\Rightarrow \lambda \geq 0$

$[\Rightarrow]$ For any $\underset{m\times n}{A}$, $A^T A$ is PSD i.e. all evals are ≥ 0

Suppose S is PSD. eg/ $n = 3$

Orthogonally diagonalize S:

$$S = \underbrace{\begin{bmatrix} - q_1 - \\ - q_2 - \\ - q_3 - \end{bmatrix}}_{Q^T} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix}}_{Q}$$

each $\lambda_i \geq 0 \Rightarrow \sqrt{\lambda_i} \in \mathbb{R}$

Write "$D = \sqrt{D} \cdot \sqrt{D}$" $= \underbrace{\begin{bmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \sqrt{\lambda_3} \end{bmatrix}}_{\text{Def } \sqrt{D}}^2$

Write S as $S = Q^T D Q = Q^T \sqrt{D}\sqrt{D}\, Q = \underbrace{(\sqrt{D})^T}_{A^T} \underbrace{\sqrt{D}\, Q}_{A}$

Recall second derivative test Calc 1

$f(x)$ has a minimum at x_0 if $f'(x_0) = 0$ and $f''(x_0) > 0$

SDT for higher dimensions

$f(x, y)$ has min at (x_0, y_0) if $\left.\frac{df}{dx}\right|_{(x_0,y_0)} = \left.\frac{df}{dy}\right|_{(x_0,y_0)} = 0$

and $\underline{H} = \underbrace{\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}}_{E_H > 0}$ is a PD matrix at (x_0, y_0)

Hessian matrix

Singular Value Decomposition

Given any $m \times n$ matrix A, we will study the two square symetric positive semidefinite matrices associated to A

$A^T A$
$n \times n$
symetric
PSD
$\lambda \geq 0$

$A A^T$
$m \times m$
symetric
PSD
$\lambda \geq 0$

A does not have eigenvalues but $A^T A$ and AA^T do.

"$\sqrt{\text{Eigen values of } A^T A}$ = Singular values of A"

we can't diaganalize A, it's not square

Next best thing: Decompose A into $A = Q D P^{-1}$

invertible, diagonal, invertible, $P \neq Q$

Observation

Let $\underset{n \times n}{A}$ with eigenvalue $\lambda \in \mathbb{R}$. Pick an eigenvector $\vec{v} \in \mathbb{R}^n$ for λ

① $Av = \lambda v$

And normalize $\vec{v}$ ② $\|\vec{v}\| = 1$

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

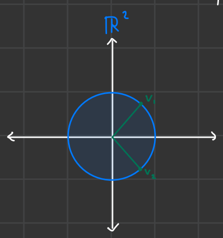
So $|\lambda|$ tells us how much A stretches in the $\vec{v}$ direction

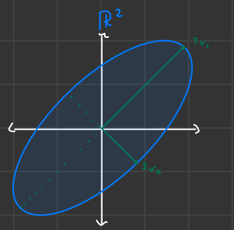
In particular, the evals with largest/smallest magnitudes tell us which directions A stretches the most / least

Ex $A = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$

evals: $\lambda_1 = 7 \longrightarrow v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, $\lambda_2 = 3 \longrightarrow v_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

$Av_1 = 7v_1$ $\qquad$ $Av_2 = 3v_2$





$A = QDQ^{-1}$
$\Rightarrow \det(A) = \det(D)$

Now let A be $m \times n$. We can't ask for evects and evals but the question: "which direction does A stretch the most / least?" still makes perfect sense!

i.e we want to compute

$\max \{ \|Av\| \mid \|v\| = 1 \}$

$\min \{ \|Av\| \mid \|v\| = 1 \}$

Ex $T_A : \mathbb{R}^3 \to \mathbb{R}^2$ given by $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

which direction does A stretch most and by how much?

i.e. $\max \{ \|Av\| \mid \|v\| = 1 \}$?

First: $\|Av\|$ is maximized $\iff$ $\|Av\|^2$ is maximized

$\|Av\|^2 = (Av)\cdot(Av) = (Av)^T Av = v^T (A^T A) v$

So, $\underline{\|Av\|^2 = \text{Energy of } A}$ $\Rightarrow \underset{\|v\|=1}{\|Av\|^2}$ is maximized $\iff$ Energy of $A^T A$ is maximized

By our theorem, the max energy of $A^T A$ is given by the largest eigenvalue and its attaind at a unit eigenvector corresponding to the largest eval.

So, we need to compute eigenvalues and evects for $A^T A$.

$$A^T A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix} \longrightarrow \text{eigenvalues of } A^T A \quad \begin{array}{l} \lambda_1 = 360 \\ \lambda_2 = 90 \\ \lambda_3 = 0 \end{array}$$

So, the max energy of $A^T A$ is $\lambda_1 = 360$ and its attained a unit eigenvect corresponding to $\lambda_1 = 360$

$\Rightarrow$ max value of $\|Av\|^2 = 360$

$\Rightarrow$ max value of $\|Av\| = \sqrt{360} = \sqrt{\lambda_1}$ ← first singular value of A

Recall For any matrix $\underset{m\times n}{A}$, $\underset{n\times n}{A^T A}$ is $\underset{B^T = B}{\text{symmetric}}$, $\underset{\lambda \geq 0}{\text{PSD}}$

① $A^T A$ can be orthogonally diagonalized

② eigenvalues of $A^T A$ are all ≥ 0

Def Let $\underset{m\times n}{A}$

Let $\{v_1, v_2, \cdots, v_n\}$ be an orthonormal basis of eigenvectors of $A^T A$ for $\mathbb{R}^n$ corresponding to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n \geq 0$ of $A^T A$. Then, the singular Values of A are $\underline{\sigma_i = \sqrt{\lambda_i} \quad 1 \leq i \leq n}$ arranged in descending order.

Notice $\|Av_i\|^2 = (Av_i)^T Av_i = v_i^T (A^T A v_i) = v_i^T \lambda_i v_i \Rightarrow \underline{\|Av_i\| = \sigma_i}$

Ex $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ $\qquad \lambda_1 = 360 > \lambda_2 = 90 > \lambda_3 = 0$

$\Rightarrow \sigma_1 = \sqrt{\lambda_1} = \sqrt{360} > \sigma_2 = \sqrt{90} > \sigma_3 = 0$

Note: Some singular values might be zero. The non-zero singular values will play an important role.

Def

a possibly non-square matrix is called $\underline{\text{diagonal}}$ if $a_{ij} = 0$ when $i \neq j$

The Singular Value Decomposition (SVD)

Let A be an $m\times n$ matrix with r non-zero singular values $(r \le n)$

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n > 0$$

Then, $A = U\Sigma V^T$ where

- U is $m\times m$ orthogonal (orthonormal cols) matrix
- V is $n\times n$ orthogonal matrix
- Σ is an $m\times n$ diagonal matrix with $\sigma_1, \sigma_2, \cdots, \sigma_r$ along the diagonal.
- The columns of U are called the left singular vectors of A.
- The columns of V are called the right singular vectors of A.

$$\underset{m\times n}{A} = \underbrace{\underset{m\times m}{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{bmatrix}}}_{\text{left singular vectors of } A} \underbrace{\underset{m\times n}{\left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \sigma_2 & & 0 \\ 0 & & \ddots\ \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]}}_{r \text{ non-zero singular values of } A} \underset{n\times n}{\begin{bmatrix} - & v_1 & - \\ - & v_2 & - \\ & \vdots & \\ - & v_n & - \end{bmatrix}} \Bigg\} \text{right singular vectors of } A$$

Note: Since $V^{-1} = V^T \Rightarrow AV = U\Sigma$

$$A\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} = U\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \sigma_2 & & 0 \\ & & \ddots\ \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$$

$$\Rightarrow Av_1 = \sigma_1 u_1$$
$$Av_2 = \sigma_2 u_2$$
$$\vdots$$
$$Av_r = \sigma_r u_r$$
$$Av_{r+1} = 0$$
$$\vdots$$
$$Av_n = 0$$

Ex $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Step 1 Orthogonally Diagonalize A^TA

$\lambda_1 = 360 > \lambda_2 = 90 > \lambda_3 = 0$ $\qquad v_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$, $v_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, $v_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$

Step 2 Set up Σ and V

$$\underbrace{\sigma_1 = \sqrt{\lambda_1} = \sqrt{360} > \sigma_2 = \sqrt{\lambda_2} = \sqrt{90}}_{r=2 \text{ non-zero singular values}} > \sigma_3 = \sqrt{\lambda_3} = 0$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \underbrace{\begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}}_{\text{right singular vectors of } A}$$

Step 3 Find the left singular vectors and set up U

Easy case: If $r=m$, then set $u_i = \frac{1}{\sigma_i} Av_i \quad 1 \le i \le r = m$

For this example, $r = 2 = m$ ✓easy

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{\sqrt{360}}\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} Av_2 = \frac{1}{\sqrt{90}} A \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

Set $U = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}$

So, the SVD of A is

$$\underset{2\times 3}{A} = \underset{2\times 2}{U}\ \underset{2\times 3}{\Sigma}\ \underset{3\times 3}{V^T}$$

$$= \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}\begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

SVD Facts

Let $A = U\Sigma V^T$ be an SVD of A. $\underbrace{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r}_{r \text{ non-zero singular values}} > \underbrace{\sigma_{r+1} = \cdots = \sigma_n = 0}_{n-r \text{ zero singular values}}$

(0) $A = U\Sigma V^T \Rightarrow AV = U\Sigma$

$$
\begin{aligned}
Av_1 &= \sigma_1 u_1 & \qquad Av_{r+1} &= 0 \\
Av_2 &= \sigma_2 u_2 & &\vdots \\
&\vdots & & \\
Av_r &= \sigma_r u_r & Av_n &= 0
\end{aligned}
$$

(1) $\{u_1, u_2, \cdots, u_r\}$ is an orthonormal basis for $Col(A)$

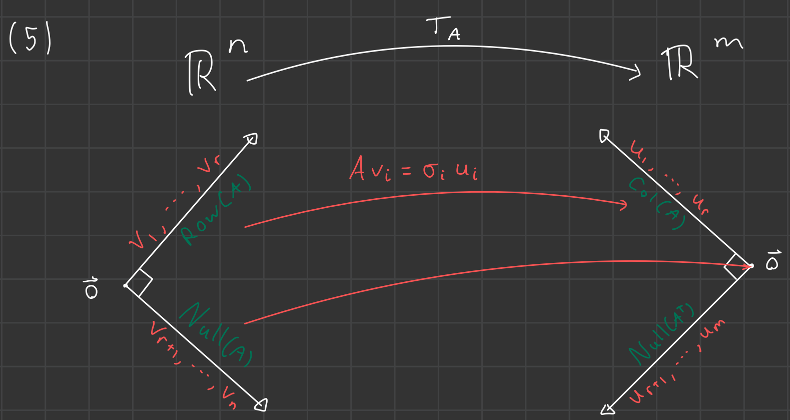
$\Rightarrow \text{rank } A = \dim Col(A) = \underline{r}$ ← # of non-zero singular values

(2) $\{v_{r+1}, v_{r+2}, \cdots, v_n\}$ forms an ONB for $Null(A)$ $\Rightarrow \dim Null(A) = n - r$

(3) $\{v_1, v_2, \ldots, v_r\}$ forms an ONB for Row A $\Rightarrow \dim \text{Row } A = r = \text{rank}(A)$

(4) $\{u_{r+1}, \cdots, u_m\}$ forms an ONB for $Null(A^T) \Rightarrow \dim Null(A^T) = m - r$ ✓

(5)



(6) If A is $n \times n$, then A is invertible iff $\sigma = 0$ is not a singular value of A

Principle Component Analysis

PCA ≈ SVD

Goal Given some data, find a new way to view the data (basis) which extracts as much info as possible in as few dimensions as possible

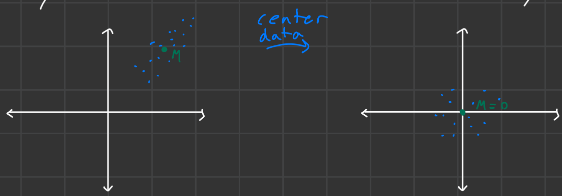
Setup Given vectors $\overbrace{X_1, X_2, \ldots, X_N}^{\text{"filled w/ data"}} \in \mathbb{R}^p$

The mean of $X_1, X_2, \ldots, X_N$ $\quad M = \frac{1}{N}(x_1 + x_2 + \cdots + x_N) \in \mathbb{R}^p$

This calculates the average of each row of the matrix:

$$\underset{N \times P}{\begin{bmatrix} | & | & & | \\ X_1 & X_2 & \cdots & X_N \\ | & | & & | \end{bmatrix}} \quad \text{ie} \quad M = \begin{bmatrix} \text{avg } R_1 \\ \vdots \\ \text{avg } R_p \end{bmatrix}$$

Assume/Force the data to be centered, i.e. $M = 0$



Assuming $M = 0$, we say $\underset{P \times N}{B} = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_N \\ | & & | \end{bmatrix}$ is the data matrix in mean-deviation form

Set $\quad S = \frac{1}{N-1} BB^T \quad \leftarrow$ Covariance matrix

The trace of S, $tr(S)$ is called <u>total variation</u> of the data. It's a measure of how spread the data is.

PCA Let $B = \underset{P \times N}{\begin{bmatrix} | & & | \\ x_1 & \cdots & x_N \\ | & & | \end{bmatrix}}$, w/ $M = 0$, $\quad \underset{N \times P}{A} = \frac{1}{\sqrt{N-1}} B^T$

$A^T A = \frac{1}{N-1} BB^T = S$

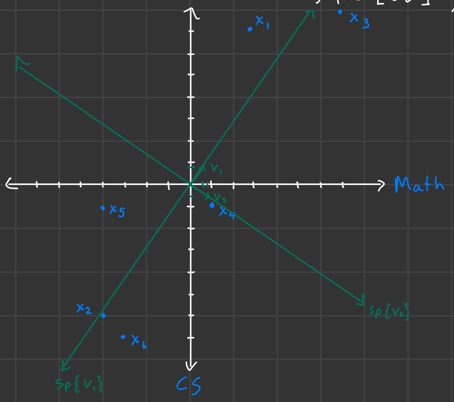
Apply SVD to A

The right singular vectors of A are called <u>the principle components of the data</u>. V_1 = first principle component (largest σ), V_2 = second principle component (second largest σ), $\ldots$, V_p.

Ex Math/CS

$$\begin{array}{c} \text{Math} \\ \text{CS} \end{array}\begin{bmatrix} 88 & 81 & 92 & 86 & 81 & 82 \\ 92 & 79 & 93 & 84 & 84 & 78 \end{bmatrix}$$
students $\rightarrow$ $x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6$

First center the data: $M = \begin{bmatrix} 85 \\ 85 \end{bmatrix}$, $B = \begin{array}{c} \text{Math} \\ \text{CS} \end{array}\begin{bmatrix} 3 & -4 & 7 & 1 & -4 & -3 \\ 7 & -6 & 8 & -1 & 1 & -7 \end{bmatrix}$ $\leftarrow$ mean-deviation form
$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6$



PCA $\quad A = \frac{1}{\sqrt{6-1}} B^T = \frac{1}{\sqrt{5}} B^T$

$A^TA = S = \begin{bmatrix} 20 & 25 \\ 25 & 40 \end{bmatrix}$

An SVD of A (with rounding) is:

$$\underset{6\times 2}{A} \approx \underset{6\times 6}{U} \underset{6\times 2}{\begin{bmatrix} \sqrt{57} & 0 \\ 0 & \sqrt{3} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}} \overset{V^T}{\underset{2\times 2}{\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}}} \begin{array}{l} \leftarrow 1^{st} \text{ p.c.} \\ \leftarrow 2^{nd} \text{ p.c.} \end{array}$$

$v_1 = \begin{bmatrix} .6 \\ .8 \end{bmatrix}$ explains the most dominant direction of the data

$v_2 = \begin{bmatrix} .8 \\ -.6 \end{bmatrix}$

* v_1 can be interpretted as a new way to compare students.

Strength Index $\quad SI = .6\,\text{Math} + .8\,CS$

Consistency Index $\quad CI = .8\,\text{Math} - .6\,CS$

The "hard" case of SVD ($r < m$)

$$\underset{m\times n}{A} \rightarrow \underset{n\times n}{A^TA} \quad \sigma_i = \sqrt{\lambda_i}, \; v_i = \text{right sing. vec.}$$
$$\underset{m\times n}{A} \rightarrow \underset{m\times m}{AA^T}$$

Thm Let A be $m \times n$

(1) Non-zero eigenvalues A^TA and AA^T are the same!
They are $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$

U (2) The left singular vectors of A is an orthonormal basis of eigenvectors AA^T

Pf of 1 Let $A = U\Sigma V^T$

$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma V^T V \Sigma^T U^T = U \underbrace{\Sigma\Sigma^T}_{m\times m} U^T$

Turns out, $\Sigma\Sigma^T$ is $\begin{bmatrix} \sigma_1^2 & & & & 0 \\ & \sigma_2^2 & & & \\ & & \ddots & & \\ & & & \sigma_r^2 & \\ 0 & & & & \underbrace{0 \; \ddots \; 0}_{m-r} \end{bmatrix}_{m\times m}$ diagonal!

$\Rightarrow AA^T = U(\Sigma\Sigma^T)U^{-1}$ is an orthogonal diagonalization of AA^T

$\Rightarrow$ eigenvals of AA^T are $\underbrace{\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2}_{\text{evals of } A^TA}, 0, 0, \dots, 0$

General process for SVD

$r \le \min\{m, n\}$

Let A be $m \times n$

Step 1 Orthogonally digonalize A^TA

Step 2 Set up Σ and V

Step 3 For $1 \le i \le r$, define $u_i = \frac{1}{\sigma_i} A v_i$

Step 4 Find the remainding $m-r$ lef singular vectors $u_{r+1}, \cdots, u_m$

(4.1) Find a basis for the eigenspace of AA^T corresponding to $\lambda = 0$

i.e. parameterize $\mathrm{Null}(AA^T)$

(4.2) Apply G.S.P to the basis found in 4.1

(4.3) these are the last $m-r$ left singular vectors

$$U = \Big[\underbrace{u_1 \cdots u_r}_{\text{formula}} \; \underbrace{u_{r+1} \cdots u_m}_{\mathrm{Null}(AA^T)}\Big]$$

Low Rank Approximation

Let A be $m \times n$, $A = U \Sigma V^T$

$$\forall\ 1 \le k \le \mathrm{rank}(A) = r \qquad A_k \overset{\text{def}}{=} \underbrace{\overset{U_k}{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_k \\ | & | & & | \end{bmatrix}}}_{m \times k} \underbrace{\overset{\Sigma_k}{\begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 \ddots & \\ 0 & & \sigma_k \end{bmatrix}}}_{k \times k} \underbrace{\overset{V_k^T}{\begin{bmatrix} - v_1 - \\ - v_2 - \\ \vdots \\ - v_k - \end{bmatrix}}}_{k \times n}$$

Thm (Eckart – Young)

The matrix A_k has rank k and it is the best approximation of A by a rank k matrix. $\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2$

i.e. $\forall\ B_{m \times n}$ with rank k, $\|A - A_k\| \le \|A - B\|$

this theorum holds for all notions of $\|A\|$

what could $\|A\|$ mean?

- $\|A\| = \sum a_{ij}^2$ for $A = [a_{ij}]$
- $\|A\| = \max \|Ax\| = \sigma_1$
- $\vdots$

Use Given a huge data matrix $A_{m \times n}$ w/ singular values

$\sigma_1 \ge \sigma_2 \ge \sigma_3 \ge \cdots \ge \sigma_{1000} \ge \cdots \ge \sigma_r$

Pseudo inverse

Let A be $m \times n$ w/ $\mathrm{rank}(A) = r$. Pick $k = \mathrm{rank}(A)$

$$A = A_r = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} - v_1 - \\ \vdots \\ - v_r - \end{bmatrix}$$

$$A = A_r = U_r \Sigma_r V_r^T$$

$$A^+ = V_r \Sigma_r^{-1} U_r^T$$