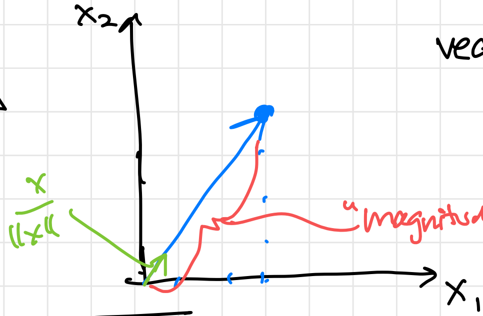
1/12/2024 Section

- Vector $v = \begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix} \in \mathbb{R}^d$ ← how many numbers in v

↖ each element of v is real number

$x = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

vector has <u>magnitude</u> and <u>direction</u>



"magnitude" = length = $\|x\|$

$$\|x\| = \sqrt{\sum_{i=1}^{d} x_i^2} \quad (\text{"Euclidean norm"})$$

<u>Unit vector</u>: vector of magnitude 1

e.g. $\frac{x}{\|x\|}$ is unit vector in same direction as x

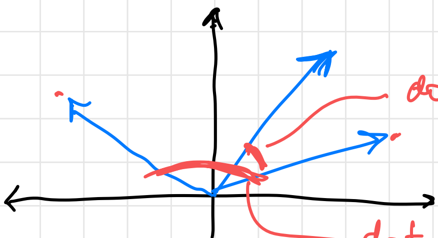
↳ e.g. $\frac{x}{\|x\|} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$

- Dot product ("inner product")

For 2 vectors $v, u \in \mathbb{R}^d$, dot product is

$$v^T u = v_1 u_1 + v_2 u_2 + \cdots + v_d u_d = \sum_{i=1}^{d} v_i u_i$$

Fact: $v^T u = \|v\| \cdot \|u\| \cdot \cos\theta$ ↖ Angle between v & u



dot product > 0 (point in similar direction)

dot product < 0 (point in opposite direction)

If $v^T u = 0$, then v & u are "perpendicular" or "orthogonal"

Fact: $v^T v = \|v\|^2$

Matrix $(m \times n)$

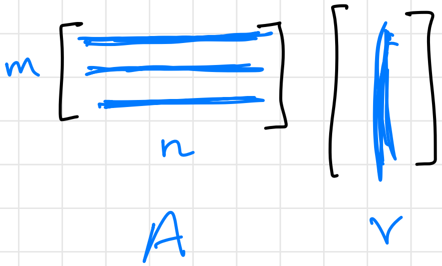
$\uparrow$ #rows $\quad \uparrow$ #columns

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & & & \\ \vdots & & \ddots & \\ A_{m1} & & & A_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

- Matrix-vector product:

For $A \in \mathbb{R}^{m \times n}$, vector $v \in \mathbb{R}^{n}$, then $Av \in \mathbb{R}^{m}$

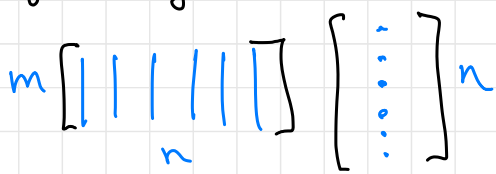
must match



$$= \begin{bmatrix} a_1^T v \\ a_2^T v \\ \vdots \\ a_m^T v \end{bmatrix}$$

where a_i is ith row of A

Equivalently:



$$= v_1 \alpha_1 + \cdots + v_n \alpha_n$$

where α_i is ith column of A

Quick proof for 2×2

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} A_{11} v_1 + A_{12} v_2 \\ A_{21} v_1 + A_{22} v_2 \end{bmatrix} = v_1 \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} + v_2 \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}$$

For matrix $A \in \mathbb{R}^{m \times n}$

the transpose $A^T \in \mathbb{R}^{n \times m}$

Note: Dot product is special case of matrix-vector multiplication

For $v, u \in \mathbb{R}^d$,
view v as column vector $\approx$ $d \times 1$ matrix $\begin{array}{c} \\ d \end{array}\left[\quad \right]$ 1

So v^T is row vector $\approx$ $1 \times d$ $1[\underset{d}{\quad}]$

so $v^T u$ yields "1-d vector" as result $\approx$ scalar

Gradients:

For function $F : \mathbb{R}^d \rightarrow \mathbb{R}$

input = d-dim. vector

output = real number

$$\nabla_x F(x) = \begin{bmatrix} \frac{dF}{dx_1} \\ \vdots \\ \frac{dF}{dx_d} \end{bmatrix}$$

Examples:

$$\nabla_x v^T x = ?$$

$$\frac{d}{dx_i} v^T x = \frac{d}{dx_i}\left[\cancel{v_1 x_1} + \cdots + \boxed{v_i x_i} + \cdots + \cancel{v_d x_d}\right]$$

constant constant

$$= \frac{d}{dx_i} v_i x_i = v_i$$

$$= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix} = v$$

Now you can use this formula!

Normal derivative rules apply:

$$\nabla_x F(x) + G(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} F(x) + G(x) \\ \vdots \\ \frac{\partial}{\partial x_d} F(x) + G(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial F}{\partial x_1} + \frac{\partial G}{\partial x_1} \\ \vdots \end{bmatrix}$$

$$= \left(\nabla_x F(x)\right) + \left(\nabla_x G(x)\right)$$

Same applies to Chain Rule

e.g. $\nabla_x (v^T x)^3 = 3(v^T x)^2 \cdot v$

because $\frac{\partial}{\partial x_i}(v^T x)^3 = \underbrace{\frac{\partial ((v^T x)^3)}{\partial (v^T x)}}_{3(v^T x)^2} \cdot \underbrace{\frac{\partial (v^T x)}{\partial x_i}}_{v_i}$

$3(v^T x)^2 \cdot v_i$