HW1 due 9/22
Gradescope code: 9WYBY6

GSIs:
Haozhu Wang (hzwang@umich.edu)
Ankit Bansal (arbansal@umich.edu)

Office Hours:
Salimeh Yasaei Sekeh : Monday 1:00pm-3:00pm in EECS 3115
Mert Pilanci : Tuesday 2:30pm-4:00pm in EECS 4419
Haozhu Wang: Wednesday 6:00pm-9:00pm in GGBL1025
Ankit Bansal: TBA

Review

Linear Algebra
Probability
Optimization

Vectors

Matrices

rank, span

eigenvalue decomposition.

Vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$ $y = \begin{bmatrix} y_1 \\ \vdots \\ y_d \end{bmatrix}$

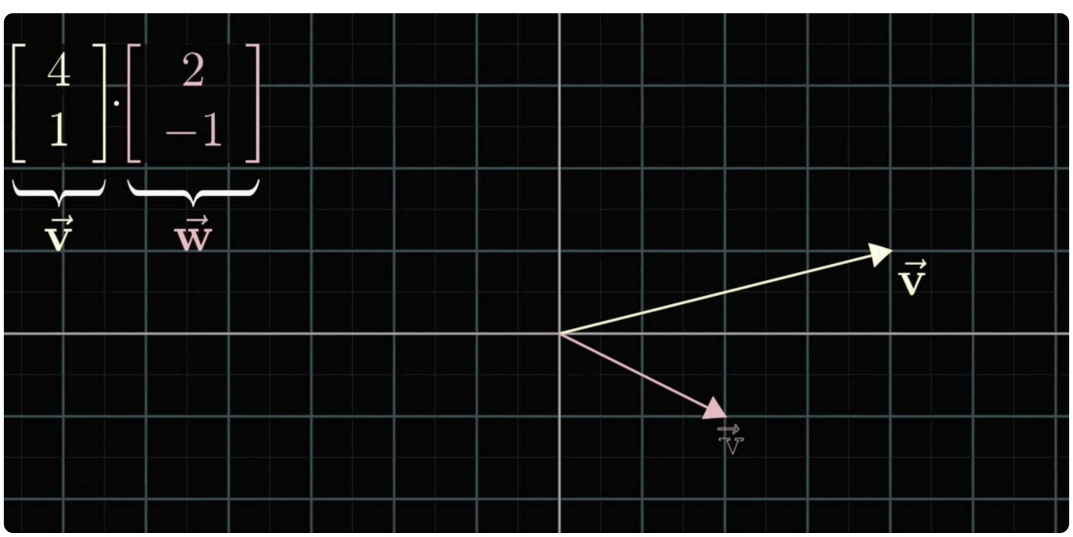
Dot product:

$$x^T y = \langle x, y \rangle = x_1 y_1 + x_2 \cdot y_2 + \cdots x_d \cdot y_d$$

$$= \sum_i x_i y_i$$

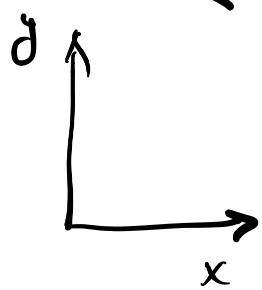
$x \in \mathbb{R}^d \Rightarrow$ inner product

Geometrically



$\underline{x}$ and $\underline{y}$ are orthogonal if

$$\langle x, y \rangle = 0$$



mapping $f(\underline{x}) \longrightarrow y$

linear if additive $f(x+y) = f(x) + f(y)$

homogeneous $f(c\cdot x) = c \cdot f(x)$

$x = \sum c_i \cdot b_i$ ← basis

$\begin{matrix}1\\0\\0\\0\end{matrix} \quad \begin{matrix}0\\1\\0\\0\end{matrix} \quad \begin{matrix}0\\0\\1\end{matrix} \quad \ldots$

$$f(x) = \sum_i f(c_i b_i) = \sum_i c_i f(b_i)$$

Matrix Multiplication

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$y_1 = A_{11}x_1 + A_{12}x_2$$

$$y_2 = A_{21}x_1 + A_{22}x_2$$

Rows of A: $\quad y_i = \sum_j A_{ij} x_j \rightarrow \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \vdots$

$\langle A_{i:}, x \rangle = y_i$

Columns of A:

$$y = Ax$$

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \qquad = \underline{a}_1 x_1 + \underline{a}_2 x_2$$

<u>Span:</u> Consider a collection of vectors

$$\underline{a}_1, \ldots, \underline{a}_n \in \mathbb{R}^d$$

a <u>linear combination</u> is any vector of the form

$$\sum_{j=1}^{n} x_j \underline{a}_j$$

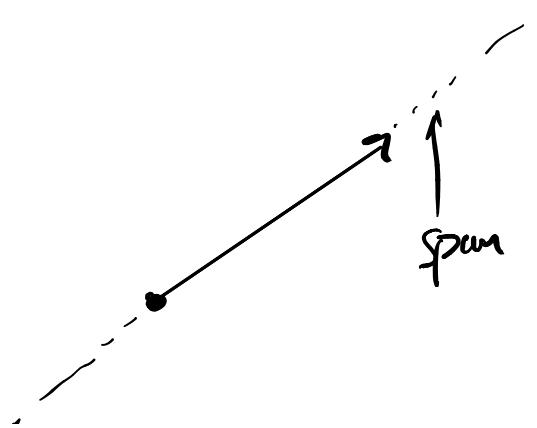
span is the set of all linear combinations

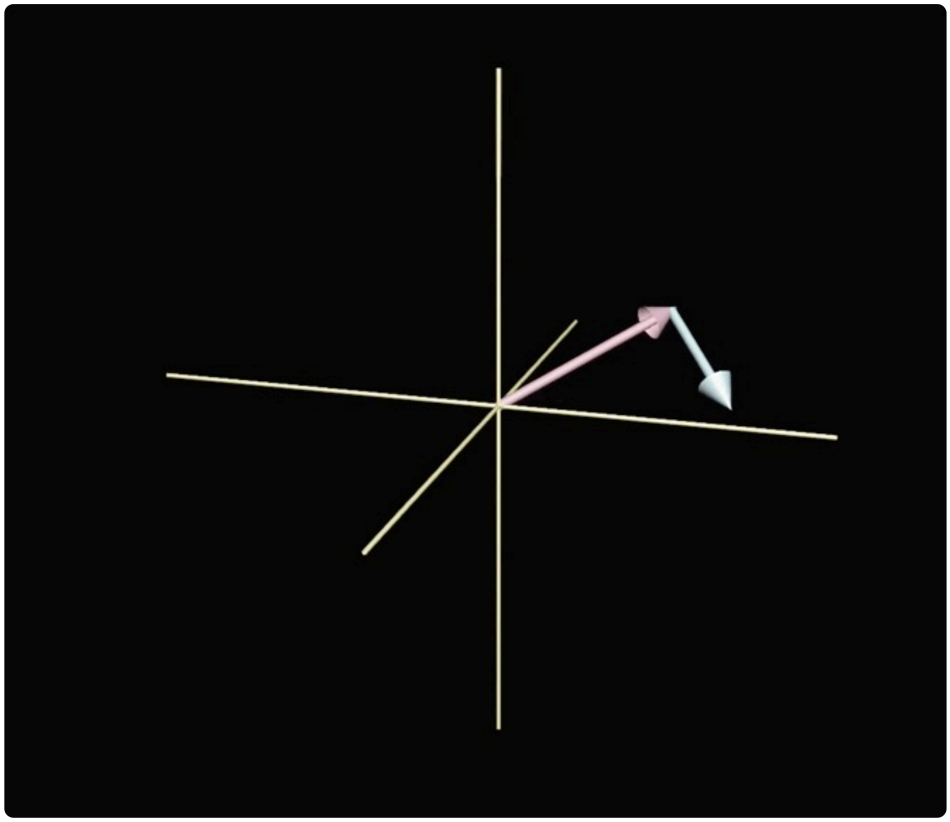
$$A = \begin{bmatrix} \underline{a}_1 & \cdots & \underline{a}_n \end{bmatrix}$$

$$\text{span}\{\underline{a}_1, \ldots, a_n\} = \{ y = Ax : x \in \mathbb{R}^d \}$$

$$= \text{colspan}(A)$$

"image of A"

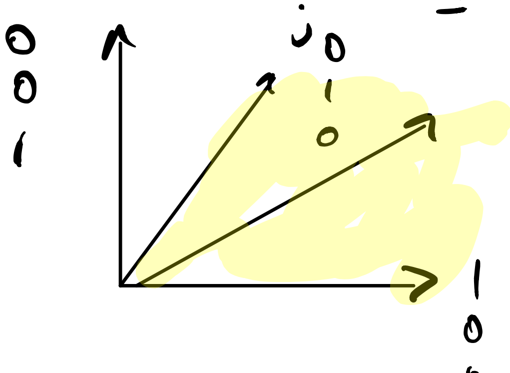




Linear independence

Vectors $\underline{a}_1, \ldots \underline{a}_n$ are linearly independent

iff $\sum_j x_j \underline{a}_j = 0 \Rightarrow x_j = 0 \ \forall j$



linearly dependent?

Nullspace: Nullspace of a matrix $A \in \mathbb{R}^{n \times d}$

$$N(A) = \{ x \in \mathbb{R}^d \mid Ax = 0 \}$$

$\underline{a}_1, \ldots, \underline{a}_n$ are linearly independent
iff $N(A) = \{0\}$

<u>Subspace</u>: set of vectors closed under addition and scalar multiplication

subspace

<u>Basis</u>: A basis for a subspace $B \subseteq S$ is a finite collection of vectors satisfying

(1) $\quad \mathrm{span}(B) = S$

(2) $\quad B$ is linearly independent

<u>Ex</u>:

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

Basis for column span of A.

$\underline{a}_1 = \begin{matrix} 1 \\ 0 \\ 0 \end{matrix}$ and $\underline{a}_2 = \begin{matrix} 0 \\ 1 \\ 0 \end{matrix}$ are a basis

for the image of A.

$$\underline{\underline{A}}\underline{x} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + x_3 \underline{a}_3 = (x_1 + x_3)\underline{a}_1 + (x_2 + x_3)\underline{a}_2$$

$$\in \mathrm{span}(\underline{a}_1, \underline{a}_2)$$

linearly independent? (

$$\underline{a}_1 \cdot x_1 + \underline{a}_2 \cdot x_2 = 0 \Rightarrow \begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array}$$

Dimension of a subspace is the cardinality of a basis.

\# elements

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

$$\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & \vdots \\ \vdots & \vdots \end{array}$$

Rank: $\operatorname{rank}(A) = \dim(\operatorname{colspan}(A))$

$A \in \mathbb{R}^{n \times d}$

Nullity: $\operatorname{nullity}(A) = \dim(N(A))$

Theorem: Rank-Nullity $\quad A \in \mathbb{R}^{n \times d}$

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = d$$

Orthogonal Complement:

Given a subspace $S \subseteq \mathbb{R}^d$

its orth. complement

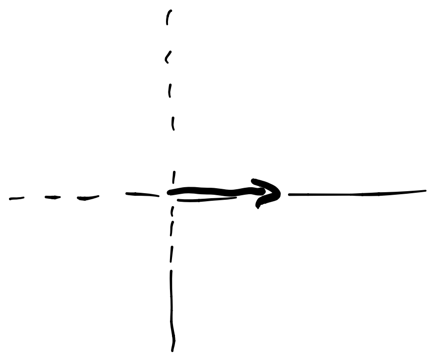
$$S^{\perp} = \{ x : \mathbb{R}^d \mid \langle x, y \rangle = 0 \ \forall y \in S \}$$

$\mathbb{R}^{-} = S \oplus S^{-}$

i.e., $\forall u \in \mathbb{R}^{d}$ $\quad u = v + w$

$v \in S$

$w \in S^{\perp}$



Norms A mapping $g(x)$ $\mathbb{R}^d \to \mathbb{R}_+$ is said to define a norm on $\mathbb{R}^d$ if

(1) definiteness $\forall x \in \mathbb{R}^d$ $g(\underline{x}) = 0 \iff \underline{x} = 0$

(2) homogeneity $g(\alpha \cdot \underline{x}) = |\alpha| g(x)$

(3) triangle inequality $\forall x, y \in \mathbb{R}^d$

$$g(x+y) \leq g(x) + g(y)$$

typically $\|x\|$

Euclidean Norm: $\|x\|_2 = \sqrt{\sum_j x_j^2}$

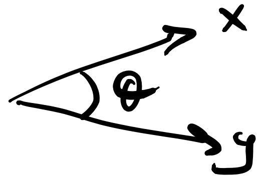
$$= \langle \underline{x}, \underline{x} \rangle^{1/2}$$

L_∞ norm: $\|x\|_\infty = \max_j |x_j|$

$\|\mathbb{1}\|_2 = \sqrt{2}$

$\|\mathbb{1}\|_\infty = 1$

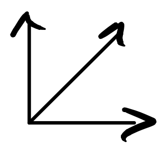
Angle



$$\cos\theta = \frac{\langle x, y\rangle}{\|x\|_2 \cdot \|y\|_2}$$

Orthogonal set $\{\underline{x}_1, \dots, \underline{x}_n\}$

if $\langle x_i, x_j\rangle = 0 \quad \forall\, i \neq j$

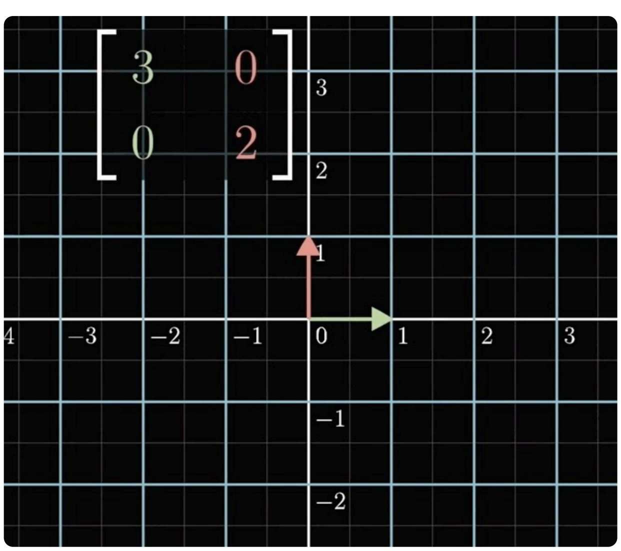


Orthonormal set: in addition $\|x_i\|_2 = 1 \quad \forall i$

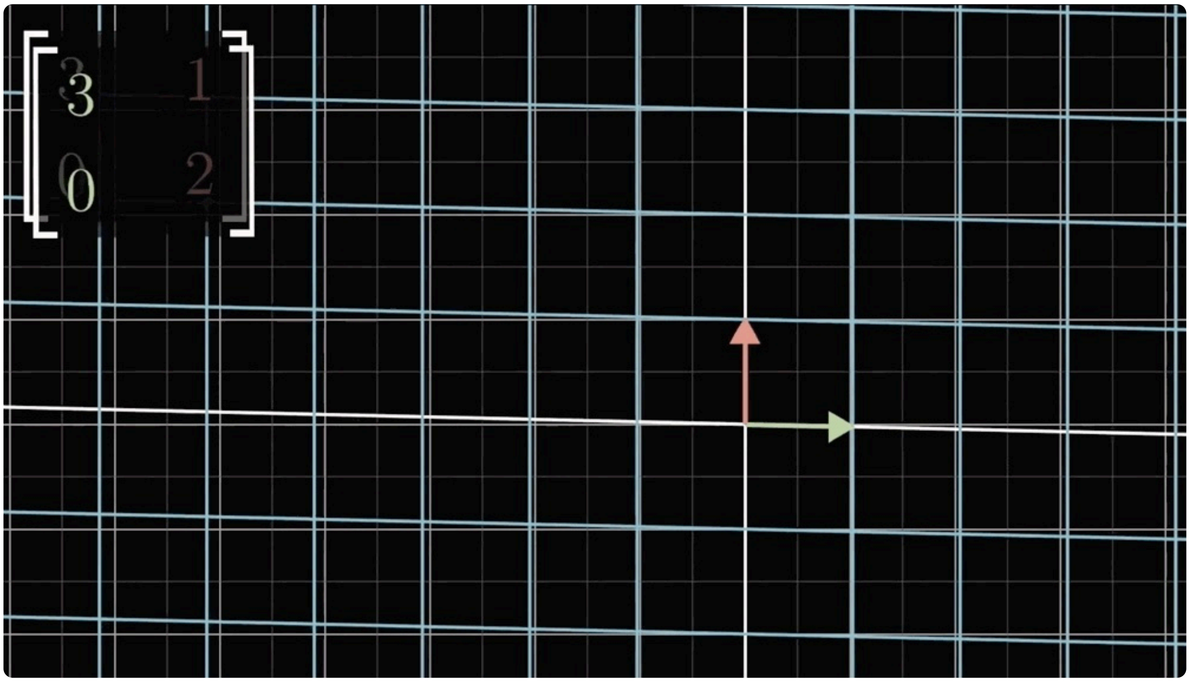
$$I = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Eigenvalues and Eigenvectors

Diagonal matrix



General Matrix



Let $A \in \mathbb{R}^{n \times n}$ Square matrices

If $Ax = \lambda x$ for some $\lambda \in \mathbb{R}$

$x \in \mathbb{R}^n$

we say λ is an eigenvalue of A

and x is a corresponding eigenvect

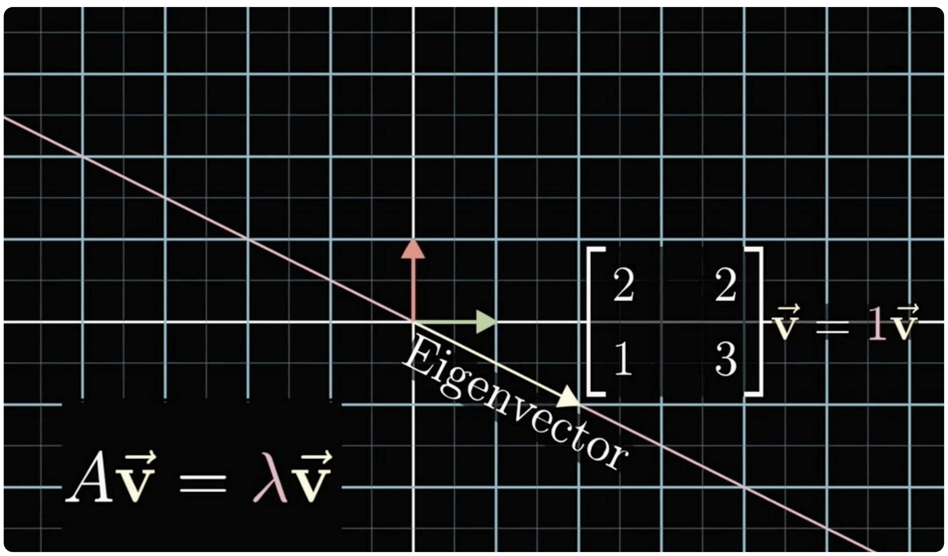
If $A = A^T$ (symmetric) we can characterize A using its e. values and e. vectors

Theorem: $(A=A^T)$: $\boxed{A = U \Lambda U^T}$

where U is an orthogonal matrix ($U^TU = UU^T = I$)

and Λ is a diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$



$A = U \Lambda U^T$ $\qquad$ $Ax = \underline{U \Lambda} \underline{U^T x}$

Multiply with U on the right

$$AU = U \Lambda \overbrace{U^T \cdot U}^{I}$$
$$= \boxed{U\Lambda}$$

$$U = \begin{bmatrix} \underline{u_1} & \underline{u_2} & \cdots & u_d \end{bmatrix}$$

$A \cdot u_i = \lambda \cdot u_i$ $\qquad$ $\lambda_1(A) \geq \lambda_2(A) \cdots \geq \lambda_d(A)$

Singular Value Decomposition (SVD)

$A \in \mathbb{R}^{n \times d}$

rectangular

Singular values : $\boxed{\sigma_i(A) = \lambda_i(A^T A)^{1/2}}$

spectral norm $\|\underline{\underline{A}}\|_2 = \max_i \sigma_i(A)$

$$= \lambda_{max}(A^T A)^{1/2}$$

The compact SVD of a matrix $A \in \mathbb{R}^{n \times d}$ with $\text{rank}(A) = r \leq \min(n,d)$ can be written

$$\underline{\underline{A}} = \underline{\underline{U_A}}\, \underline{\underline{\Sigma_A}}\, \underline{\underline{V_A}}^T$$

$\Sigma_A = \text{diag}(\sigma_1, \dots, \sigma_r)$

left singular vectors (U_A's columns) are eigenvectors of AA^T

right singular vectors (V_A's columns) are eigenvectors of $A^T A$.

Important identity: $A = U \Sigma V^T$

$$= \sum_i \sigma_i u_i v_i$$

$$\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$$

(symmetric case $\quad A = U \Lambda U^T$

$$A = \sum \lambda_i u_i u_i^T \;)$$