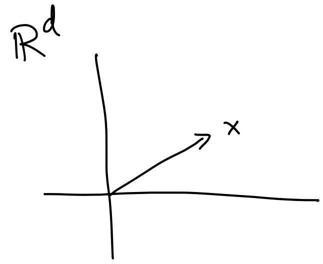
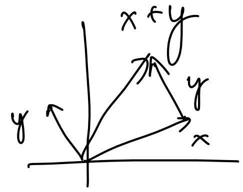
Linear Algebra Background

$$x \in \mathbb{R}^d, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}, \quad x^T = [x_1, x_2, \dots x_d]$$





Vector Norm $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}_+$

1. $\|x\| \geq 0$ & $\|x\| = 0 \Leftrightarrow x = \underline{0}$ — Non-negativity
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ — Homogeneity
3. $\|x+y\| \leq \|x\| + \|y\|$ — Triangle Inequality

$\longrightarrow$ $\|x\|_1 = |x_1| + |x_2| + \dots + |x_d|$ — L_1-norm

$\longrightarrow$ $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots |x_d|^2}$ — L_2-norm

$\|x\|_\infty = \max_{1 \leq i \leq d} |x_i|$ — L_∞-norm

$\|x\|_p = \left(|x_1|^p + |x_2|^p + \dots |x_d|^p\right)^{1/p}$ — L_p-norm

$\hat{x} = \dfrac{x}{\|x\|_2}$ — unit vector in the direction of x

$\hat{x}\hat{x}^T$ is a $d \times d$ matrix

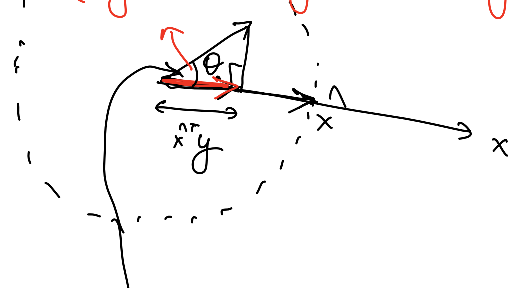
$$\hat{x}\hat{x}^T = \begin{bmatrix} \\ \hat{x} \\ \\ \end{bmatrix} [\ \hat{x}^T\ ]$$

$\downarrow$

orthogonal projection onto the vector x (or $\hat{x}$)

$$(\hat{x}\hat{x}^T)\, y = \underbrace{\hat{x}}_{\text{unit vector}} \underbrace{(\hat{x}^T y)}_{\text{scalar}}$$

$(\hat{x}^T y)\hat{x} = \hat{x}(\hat{x}^T y) = (\hat{x}\hat{x}^T)y$



$$\hat{x}^T y = \left(\frac{x}{\|x\|_2}\right)^T y = \boxed{\frac{x^T y}{\|x\|_2}} = \frac{x^T y}{\sqrt{x^T x}}$$

$$|x^T y| \le \|x\|_2 \|y\|_2$$ Cauchy Schwarz Inequality

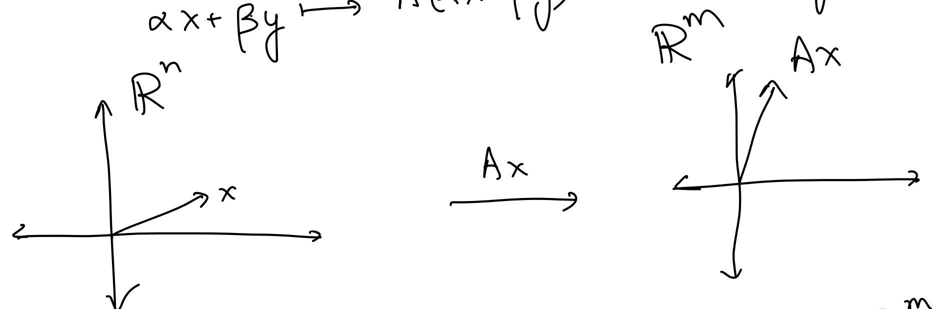
$$\frac{|x^T y|}{\|x\| \cdot \|y\|} \le 1$$

$$-1 \le \overbrace{\frac{x^T y}{\|x\| \, \|y\|}}^{\cos\theta} \le 1$$

x and y are perpendicular to each other $\iff$ $x^T y = 0$

$$A \in \mathbb{R}^{m \times n} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\begin{aligned} A : \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto Ax \\ \alpha x + \beta y &\mapsto A(\alpha x + \beta y) = \alpha A x + \beta A y \end{aligned}$$



Matrix Norms, $A \in \mathbb{R}^{m \times n}$ $\qquad \|\cdot\| : \mathbb{R}^{m \times n} \to \mathbb{R}_+$

1. $\|A\| \ge 0$, & $\|A\| = 0 \iff A = 0$ - Non-negativity
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$ - Homogeneity
3. $\|A + B\| \le \|A\| + \|B\|$ - Triangle Inequality

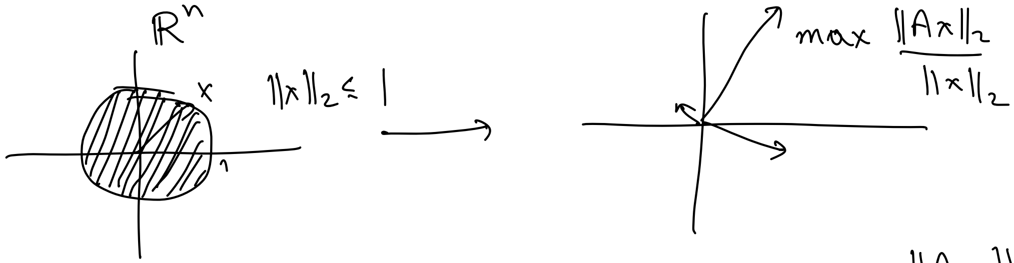
$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |A_{ij}|^2}$$ - Frobenius Norm

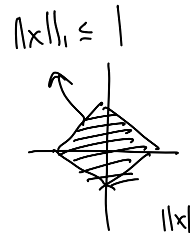
$$\sum_{i=1}^{m} \sum_{j=1}^{n} |A_{ij}|$$ — valid matrix norm

Operator Norms:

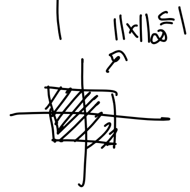
$$\|A\|_2 = \max_{x \ne 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 \le 1} \|Ax\|_2 = \sigma_1$$

↳ maximum singular value of A





$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{\|x\|_1 \le 1} \|Ax\|_1$$



$$\|A\|_\infty = \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{\|x\|_\infty \le 1} \|Ax\|_\infty$$

$$\|A\|_{p,q} = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_q}$$

$K(A)$ - condition number of A

$K(A) = \|A\| \cdot \|A^{-1}\|$

$K_2(A) = \frac{\sigma_1}{\sigma_n}$

$\sigma_1 \to$ max singular value of $A = \|A\|_2$

$\sigma_n \to$ min singular value of $A = \|A^{-1}\|_2$

Eigenvalues / Eigenvectors

$Ax = \lambda x$, $x \neq 0$, $A \in \mathbb{R}^{n \times n}$

$\lambda \to$ eigenvalue

$x \to$ eigenvector

$Ax = \lambda x$

$Ax - \lambda x = 0$

$(A - \lambda I)x = 0$

$A - \lambda I$ is singular and x belongs to its nullspace

$\Rightarrow \det(A - \lambda I) = 0$

$$\det\left(\begin{bmatrix} a_{11}-\lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ & a_{22}-\lambda & & & \vdots \\ & & a_{33}-\lambda & & \vdots \\ & & & \ddots & \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn}-\lambda \end{bmatrix}\right) = 0$$

$\det \to$ is a polynomial of degree n in λ

If A is $n \times n$, then A has n eigenvalues

If $A = A^T$ (A is a symmetric matrix), then all eigenvalues of A are real, and it has a full set (n) of mutually orthogonal eigenvectors

$A = A^T$, $Av_1 = v_1\lambda_1$, $\lambda_i \in \mathbb{R}$

$Av_2 = v_2\lambda_2$, $v_i^T v_j = 0, \ i \neq j$

$\vdots \quad \vdots$, $v_i^T v_i = 1$

$$\underbrace{Av_n = v_n\lambda_n}$$

$V = [v_1\ v_2 \dots v_n]$

$V^T V = I = VV^T$

$$A\underbrace{[v_1\ v_2 \dots v_n]}_{V} = \underbrace{[v_1\ v_2 \dots v_n]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}}_{\Lambda}$$

Orthogonal Matrix

Diagonal matrix

$$AV = V\Lambda$$

$$A\underbrace{VV^T}_{I} = V\Lambda V^T$$

$$\boxed{A = V\Lambda V^T}$$ — Eigenvalue Decomposition

A is positive semi-definite if $x^T A x \geq 0 \ \forall x$ $(A \succeq 0)$

A is positive definite if $x^T A x > 0 \ \forall x$ $(A \succ 0)$

Equivalent to saying that all its eigenvalues ≥ 0

" " " " " " > 0

$\cancel{X^T y}$ $X^T w \approx y$, $\min_w \|X^T w - y\|^2$

$\Rightarrow \boxed{XX^T} w = X^{+} y$

positive semi-definite

$\boxed{A = GG^T}$, $x^T A x = x^T G G^T x = z^T z \geq 0 \quad (z = G^T x)$

any matrix of this form is positive semi-definite

Singular Value Decomposition (SVD)

Eigenvalue Decomposition requires A to be square

$\underbrace{\text{Real, Square, Symmetric matrices}}$ have eigenvalue decomposition:

$$A = V \Lambda V^T$$

SVD exists for ALL matrices

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

SVD of A : $A = U \Sigma V^T$, U & V are orthogonal, Σ is diagonal

$m \geq n$

$$\underset{m \times n}{A} = \underset{m \times m}{\begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}} \underset{m \times n}{\begin{bmatrix} \sigma_1 & & 0 \\ 0 & \sigma_2 & \\ & \cdots & \sigma_n \\ 0 & \cdots & 0 \end{bmatrix}} \underset{n \times n}{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}}$$

$$U \qquad \Sigma \qquad V^T$$

$$UU^T = U^T U = I$$

$$VV^T = V^T V = I$$

Σ is diagonal

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_n \geq 0$$