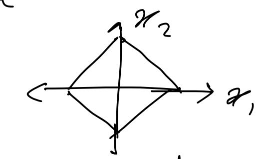
Linear Algebra Basics [Darve Chap 2 + Watkins 2.1]

1) Matrix Norms:

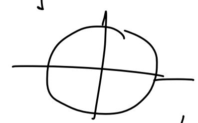
Recall vector norms: $x \in \mathbb{R}^n$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

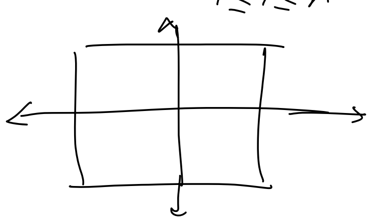


Unit ball

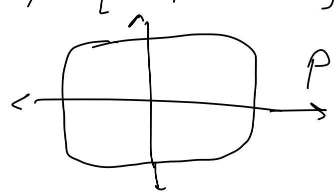
$$\|x\|_2 = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2} = \sqrt{x^T x}$$



$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$



; $\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p}$



$p = 4$

Exercise: Try in Julia: Unit balls as $p \uparrow$?

Matrix Norm : $\|\cdot\| : \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}$ satisfying the
$A \longmapsto \|A\|$

following properties:

For all $A, B \in \mathbb{R}^{n\times n}$ & all $\alpha \in \mathbb{R}$

1) $\|A\| > 0$ if $A \neq 0$
2) $\|\alpha A\| = |\alpha| \, \|A\|$
3) $\|A+B\| \le \|A\| + \|B\|$

New $\Rightarrow$ 4) $\|AB\| \le \|A\| \, \|B\|$ (submultiplicativity)

p-norm of a matrix A : (Induced norm / Operator Norm)

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

Frobenius Norm: $\|A\|_F = \left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right]^{1/2}$

Claim 1: $\|Ax\|_p \leq \|A\|_p \|x\|_p$

Proof: $\dfrac{\|Ax\|_p}{\|x\|_p} \leq \sup_{x \neq 0} \dfrac{\|Ax\|_p}{\|x\|_p} = \|A\|_p$

For $x = 0$, equality holds trivially.

Claim 2: $\|A\|_p$ is a matrix norm:

~~Property 3~~

$$\|A+B\|_p = \max_{x\neq 0} \frac{\|(A+B)x\|_p}{\|x\|_p} = \max_{x \neq 0} \frac{\|Ax + Bx\|_p}{\|x\|_p}$$

$$\leq \max_{x\neq 0} \frac{\|Ax\|_p + \|Bx\|_p}{\|x\|_p}$$

$$\leq \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} + \max_{x\neq 0} \frac{\|Bx\|_p}{\|x\|_p}$$

$$= \|A\|_p + \|B\|_p$$

Other properties: exercise

FACT: $\|A\|_1 = \max_{1\leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$

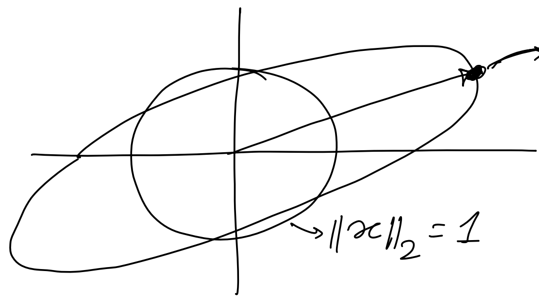
$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

Proof: Exercise

FACT: $\|A\|_p = \max_{\|x\|_p = 1} \|Ax\|_p$

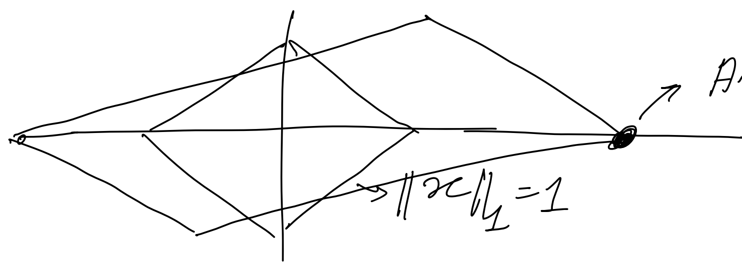
Proof: Exercise

Examples: $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$



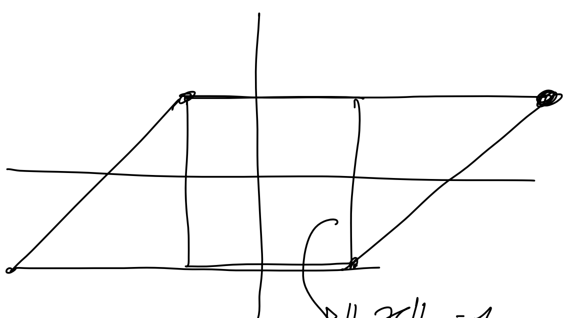
Ax^* $\Rightarrow \|A\|_2 = \|Ax^*\|_2$

$\|x\|_2 = 1$



Ax^* $\Rightarrow \|A\|_1 = \|Ax^*\|_1$

$\|x\|_1 = 1$



Ax^* $\Rightarrow \|A\|_\infty = \|Ax^*\|_\infty$

$\|x\|_\infty = 1$

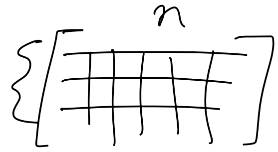
Four Fundamental Subspaces: $A \in \mathbb{R}^{m \times n}$

Col sp: $R(A) = \{Ax \mid x \in \mathbb{R}^n\} \subset \mathbb{R}^m$

$N(A) = \{y \mid Ay = 0\} \subset \mathbb{R}^n$

$N(A^T) = \{x \mid A^T x = 0\} \subset \mathbb{R}^m$

Row sp: $R(A^T) = \{A^T y \mid y \in \mathbb{R}^m\} \subset \mathbb{R}^n$

$A = m$ 

n

m - rows of length n

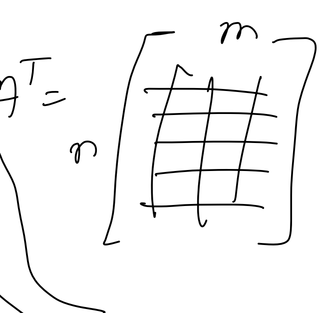
n - cols of length m

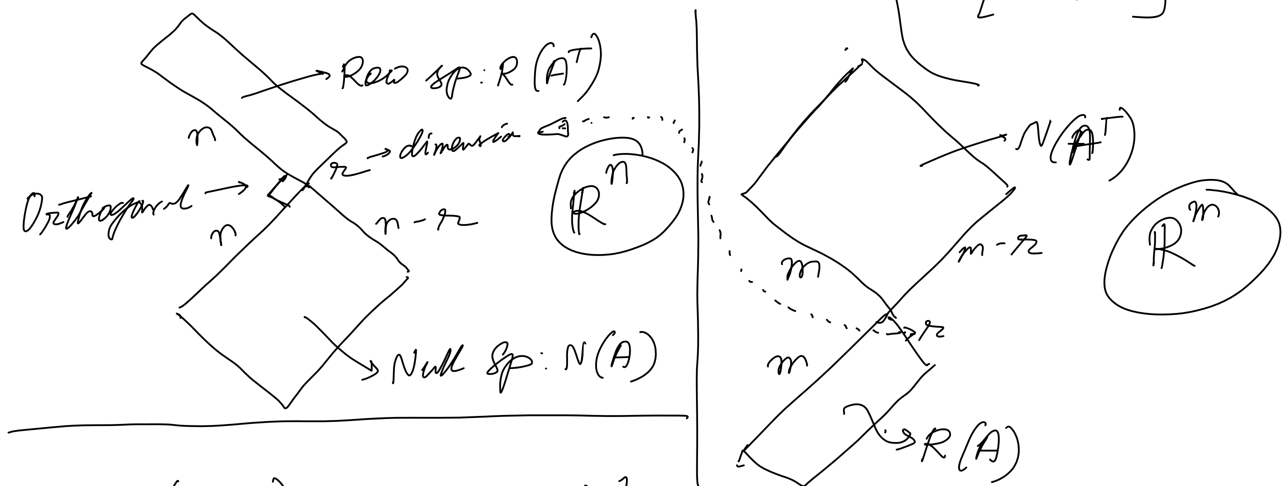
For $m \times n$ matrix A:

$N(A^T) = R(A)^\perp$; $R(A^T) = N(A)^\perp$

~~i.e.~~ $\Rightarrow$ $\dim(R(A)) + \dim[N(A^T)] = m$

$A^T =$





$\dim(R(A)) = \dim[R(A^T)] = r$ in the picture

Col rank = row rank

<u>Eigen-decomposition of matrices</u>

Let Λ be diagonal with the eigenvalues of A on the diagonal.

For $A \in \mathbb{R}^{n \times n}$, if the eigenvectors of A are lin. ind. then A can be diagonalized. i.e. $\exists$ non-singular $X \in \mathbb{R}^{n \times n}$ s.t. $A = X \Lambda X^{-1}$

$A \in \mathbb{R}^{n \times n}$ is called normal if $A^T A = A A^T$
$\exists$ a unitary matrix Q s.t.
$A = Q \Lambda Q^H$ iff A is normal.

Unitary
$U \in \mathbb{C}^{n \times n}$
$U^H U = U U^H = I$

If $A \in \mathbb{R}^{n \times n}$ & $A = A^T$, then the eigenvalues of A are real and $A = Q \Lambda Q^T$ where Q is real orthogonal.

Jordan form : Any $A \in \mathbb{R}^{n\times n}$ $\exists$ $A = XJX^{-1}$
↳ Jordan form

Schur form: For any $A \in \mathbb{R}^{n\times n}$, $\exists$ a real orthogonal matrix Q and a 2×2 upper block triangular T s.t.

$$A = QTQ^T \qquad T = \begin{bmatrix} \boxed{2\times 2} & & & \\ \begin{matrix}0 & 0\\ 0 & 0\end{matrix} & \boxed{2\times 2} & & \\ & 0 \quad 0 & \boxed{1\times 1} & \\ 0 & 0 & 0 & \boxed{1\times 1} \end{bmatrix}$$

For diagonalizable matrices $\begin{bmatrix} X \end{bmatrix} \begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix} \begin{bmatrix} X^{-1} \end{bmatrix}$

For normal matrices $\begin{bmatrix} Q \end{bmatrix} \begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix} \begin{bmatrix} Q^{H} \end{bmatrix}$

For <u>any sq. matrix</u> $\begin{bmatrix} X \end{bmatrix} \underbrace{\begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix}}_{J} \begin{bmatrix} X^{-1} \end{bmatrix}$

For any sq. matrix $\begin{bmatrix} Q \end{bmatrix} \underbrace{\begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix}}_{T} \begin{bmatrix} Q^{T} \end{bmatrix}$

<u>Singular Value Decomposition</u> : Consider $A \in \mathbb{R}^{m\times n}$ & $p = \min(m, n)$. There exist two real orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{m\times n}$ with real non-negative entries s.t.

$$A = U\Sigma V^T$$

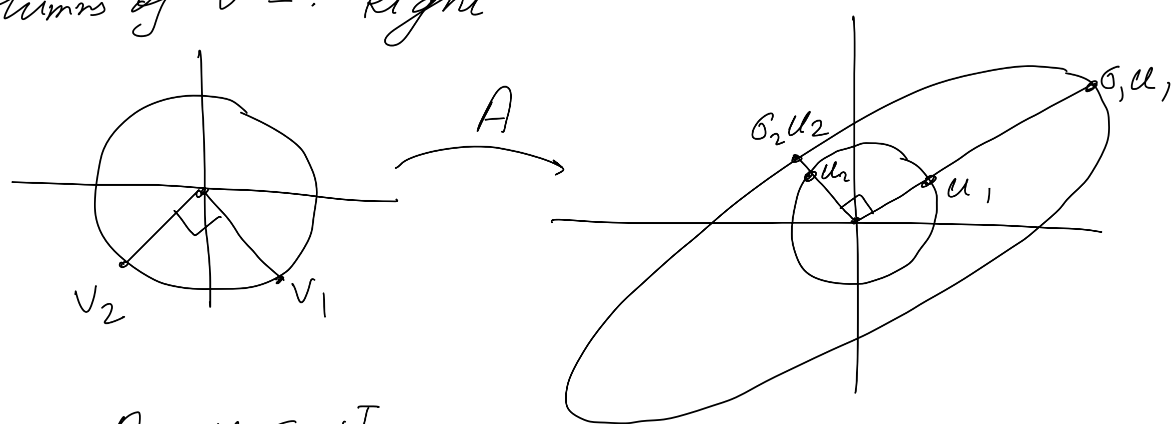
Shapes of SVD: $m > n$ | $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$

$$\begin{bmatrix} \\ A \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \quad U \quad \\ \\ \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \\ \hline & 0 & \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

$m < n$ $\qquad\qquad\qquad \Sigma$

$$\begin{bmatrix} \quad A \quad \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_p & \end{array}\right] \begin{bmatrix} \\ \quad V^T \quad \\ \\ \end{bmatrix}$$

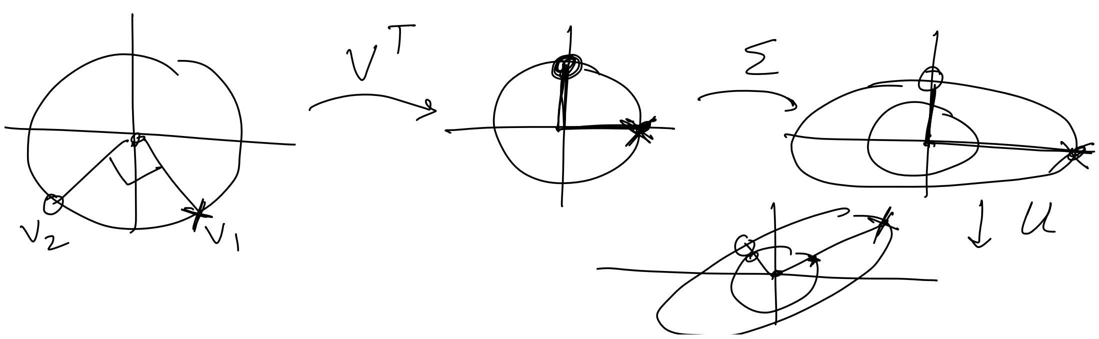
Columns of U =: Left Singular Vectors
Columns of V =: Right " "



$$A = U \Sigma V^T$$

or $\quad AV = U\Sigma$

$$\begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 u_1 & \sigma_2 u_2 \end{bmatrix}$$



If $\text{Rank}(A) = r$, $\sigma_i = 0$ for $r < i \leq p$

No of non-zero singular values = rank

$\|A\|_2 = \sigma_1(A)$, $\|A\|_F = \sqrt{\sum_{i=1}^{p} \sigma_i^2}$

<u>Proof</u>: Exercise

<u>Thm</u>: Let A be a sq. matrix of size n with singular values σ_i. Then

$$[A^TA]v_i = \sigma_i^2 v_i$$
$$[AA^T]u_i = \sigma_i^2 u_i$$

Note $AA^T \neq A^TA$

$\|A\|_2 = \sqrt{\lambda_{max}(A^TA)} = \sqrt{\lambda_{max}(AA^T)}$

<u>Thin SVD</u> For any $A \in \mathbb{R}^{m \times n}$, $\exists$ are matrices s.t. $\hat{U}^T\hat{U} = I$ and $\hat{V}^T\hat{V} = I$ and a diagonal $\hat{\Sigma}$ s.t.

$$A = \underbrace{\hat{U}}_{m \times p}\ \underbrace{\hat{\Sigma}}_{p \times p}\ \underbrace{\hat{V}^T}_{p \times n} \qquad \text{Let } p = \min(m, n)$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} A \end{bmatrix} \end{array} = \begin{bmatrix} \hat{U} \end{bmatrix}\begin{bmatrix} \diagdown \end{bmatrix}\begin{bmatrix} \hat{V}^T \end{bmatrix}$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} A \end{bmatrix} \end{array} = \begin{bmatrix} \hat{U} \end{bmatrix}\begin{bmatrix} \diagdown \end{bmatrix}\begin{bmatrix} \hat{V}^T \end{bmatrix}$$

Four Fundamental Subspaces & SVD

Let $A \in \mathbb{R}^{m \times n}$ with rank r. Let

$$A = \underbrace{\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ - & & - & & - & - \\ & & 0 & & & \end{bmatrix}}_{m \times n} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}}_{n \times n}$$

picture might be diff for $m < n$

Then
$$N(A) = \operatorname{span}\{\overset{\in \mathbb{R}^n}{v_{r+1}}, \dots, v_n\} \subset \mathbb{R}^n$$
$$R(A) = \operatorname{span}\{\underset{\in \mathbb{R}^m}{u_1}, \dots, u_r\} \subset \mathbb{R}^m$$

$$N(A^T) = \operatorname{span}\{u_{r+1}, \dots, u_m\}$$
$$R(A^T) = \operatorname{span}\{v_1, \dots, v_r\}$$

Rank r approximation

Let $A \in \mathbb{R}^{m \times n}$ & $r < \min(m,n)$. with $A = U \Sigma V^T$
with $U = [u_1 \cdots u_m]$, $\Sigma = \operatorname{diag}\{\sigma_i\}$, $V = [v_1 \cdots v_n]$
Let $B = \sum_{i=1}^{r} \sigma_i u_i v_i^T$. Then

$$\|A - B\|_2 = \sigma_{r+1}$$
$$\|A - B\|_F = \sum_{i=r+1}^{\min(m,n)} \sigma_i^2$$

Also, B minimizes $\|A-B\|_2$ & $\|A-B\|_F$ over the set of matrices of rank at most r.