EECS 16B : Module 3 / Lecture 1

Announcements:

- 0.25 EC point for each lecture you attend for rest of the term.

  links.eecs16b.org/lecture-ec

- My OH after lecture today (11:10-12:00 - 299 Cory)

Last time:

- Symmetric real matrices $\longleftrightarrow$ Spectral theorem
  - eigenvalues & eigenvectors of symmetric real matrices
- Min. energy control : motivation

Today:

- Minimum-Energy Control
- Recap of spectral theorem for symmetric matrices
- Singular Value Decomposition (SVD)

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

See Note 14

CONTROL SYSTEMS (Recap)

- Stability (states blow up?)
- Controllability (can I get to where I want)
- Efficiency: (how efficiently can I get from A $\to$ B)

$$\vec{x}[k+1] = A\vec{x}[k] + \vec{b}\,u[k]$$

$$\vec{x}[k] = A^k \vec{x}[0] + A^{k-1}\vec{b}\,u[0] + \ldots + \vec{b}\,u[k-1]$$

Consider $k = 100$

$\vec{x}[0] = 0.$

$$\vec{x}[100] = A^{99}\vec{b}\,u[0] + \ldots + \vec{b}\,u[99]$$

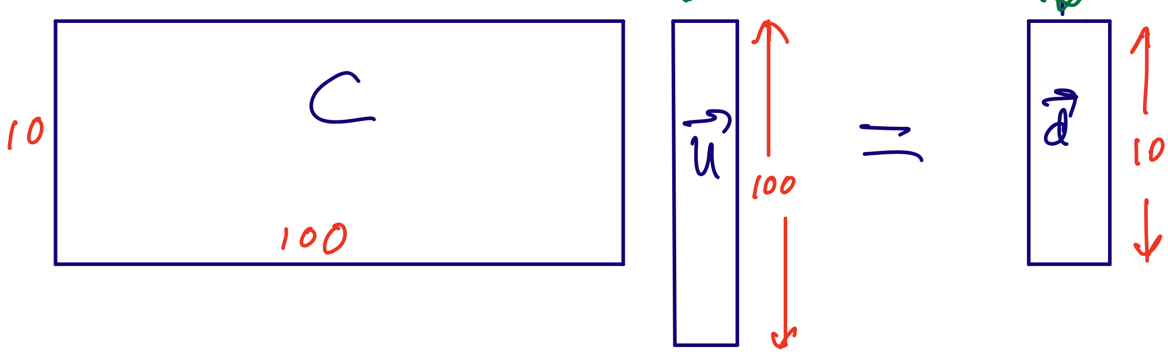
$$\underbrace{\vec{x}[100]}_{\vec{d}} = \underbrace{\left[\, A^{99}\vec{b} \;\middle|\; A^{98}\vec{b} \;\middle|\; \ldots \;\middle|\; A\vec{b} \;\middle|\; \vec{b} \,\right]}_{C} \underbrace{\begin{bmatrix} u[0] \\ \vdots \\ u[99] \end{bmatrix}}_{\vec{u}}$$

$\underbrace{\vec{x}[100]}_{\vec{d}} = C\vec{u}$

$$\min \|\vec{u}\|^2$$
$$\text{s.t. } \vec{x}[100] = Cu$$

"Minimum Energy control"

$$C_{(10\times100)}\ \vec{u}_{(10\times1)} = \vec{d}_{(10\times1)}$$



$$\boxed{\text{Solve } \quad \min \|\vec{u}\|^2 \quad \text{subject to } C\vec{u} = \vec{d}} \quad (*)$$

Solution to (*) is $\vec{u}^*$, called the minimum norm solution.

The $\vec{u}^*$ that solves (*) is written as:

$$\boxed{\vec{u}^* = \underset{\vec{u}}{\arg\min} \|\vec{u}\|^2 \ \text{ s.t. } C\vec{u} = \vec{d}}$$

Ex.

$$x[k+1] = \underbrace{A}_{1}x[k] + \underbrace{B}_{1}u[k]$$

$$\boxed{\begin{aligned} x[2] &= x[1] + u[1] \\ &= x[0] + u[0] + u[1] \end{aligned}}$$

$$\Rightarrow \underbrace{\begin{array}{cc} AB & B \\ \left[\begin{matrix} 1 \end{matrix}\right. & \left.\begin{matrix} 1 \end{matrix}\right] \end{array}}_{C} \underbrace{\begin{bmatrix} u_0 \\ u_1 \end{bmatrix}}_{\vec{u}} = \underbrace{2}_{d}$$

If $x[0]=0$ and $x[2]=2$, how should we assert our control actions, $u[0]$ & $u[1]$ s.t. the exerted energy $|u[0]|^2 + |u[1]|^2$ is minimum?

$$\boxed{\vec{u^*} = \underset{\vec{u}}{\arg\min}\ \|u\|^2 \quad \text{s.t. } u_0 + u_1 = 2}$$

Given $\boxed{u_0 + u_1 = 2}$, what is the soln. (u_0, u_1) that min. $|u_0|^2 + |u_1|^2$?

One equation, two unknowns:

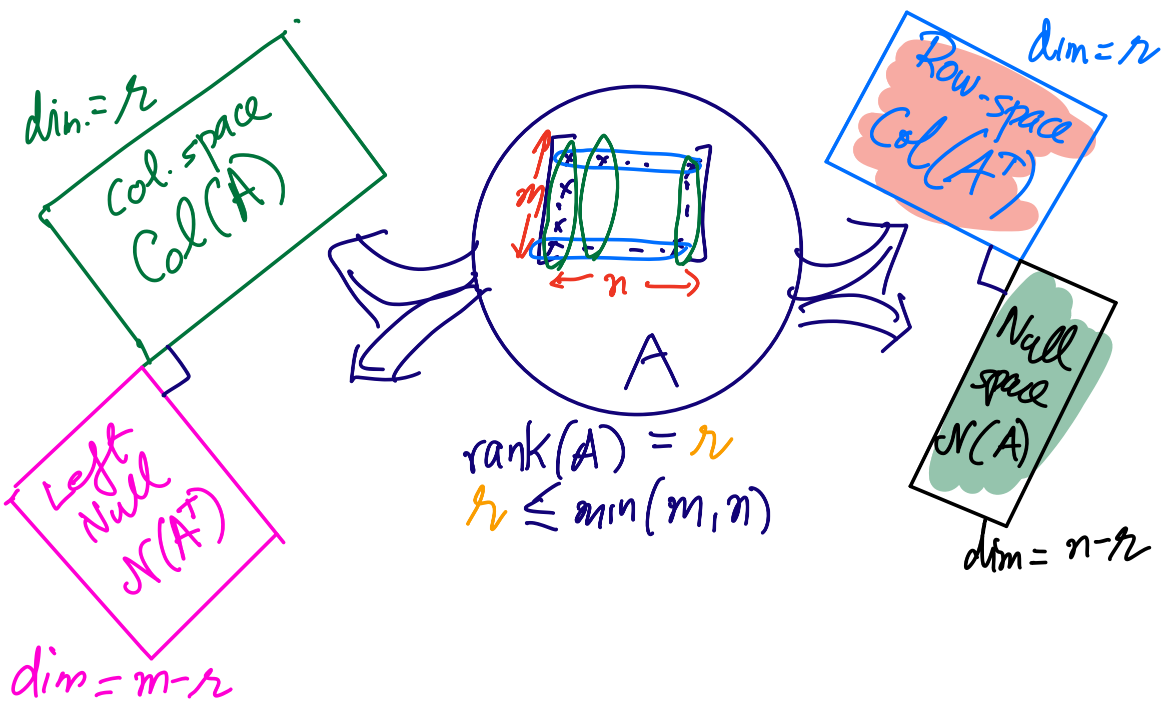
Q) How many solutions for (u_0, u_1)?

A) Infinitely many!

e.g. $(u_0, u_1) = (2, 0), (1.5, 0.5), \ldots (0, 2)$

BASICS: Fundamental theorem of Linear Algebra:

<u>Reminder</u> (4 fundamental spaces)



- $\{\vec{x} = A\vec{w}\ (\forall \vec{w} \in \mathbb{R}^n)\} \in Col(A)$
- $\{\vec{w} : A\vec{w} = 0\} \in Null(A)$
- $\{\vec{x} = A^T\vec{w}\ (\forall \vec{w} \in \mathbb{R}^m)\} \in$ Row-space(A)
- $\{\vec{w} : A^T\vec{w} = 0\} \in Null(A^T)$

Gilbert Strang:

Linear Algebra Fact:

$$\boxed{\text{Nullspace}(C) \perp \text{Row-space}(C)}$$

Why?

Null-space(C): all $\vec{w}$ s.t. $C\vec{w} = 0$

$$C\vec{w} = \begin{bmatrix} - & \vec{c}_1^{\,T} & - \\ - & \vec{c}_2^{\,T} & - \\ & \vdots & \\ - & \vec{c}_m^{\,T} & - \end{bmatrix} \vec{w} = \begin{bmatrix} \langle \vec{c}_1, \vec{w} \rangle \\ \langle \vec{c}_2, w \rangle \\ \vdots \\ \langle \vec{c}_m, \vec{w} \rangle \end{bmatrix} = 0$$

$\Rightarrow$ If $\vec{w} \in N(C)$, then $\vec{w} \perp$ each row of C

$$\Rightarrow \underbrace{\vec{w}}_{N(C)} \perp \underbrace{\text{any weighted combo of rows of } C}_{\text{Row-space}(C)}$$

$$\boxed{\vec{u}^* = \underset{\vec{u}}{\arg\min}\ \|u\|^2 \quad \text{s.t.} \quad u_0 + u_1 = 2}$$

$$\underbrace{\begin{bmatrix} 1 & 1 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} u_0 \\ u_1 \end{bmatrix}}_{\vec{u}} = \underbrace{2}_{d}$$

min-norm soln. is the $\vec{u}^*$ that is "smallest" in length; i.e. "closest" to the origin.

From linear algebra fact, $\underbrace{\text{Row-space}(C)}_{\alpha\begin{bmatrix} 1 & 1 \end{bmatrix}} \perp \underbrace{N(C)}_{\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0}$

$u_0 + u_1 = 0$

$\Rightarrow \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \beta \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

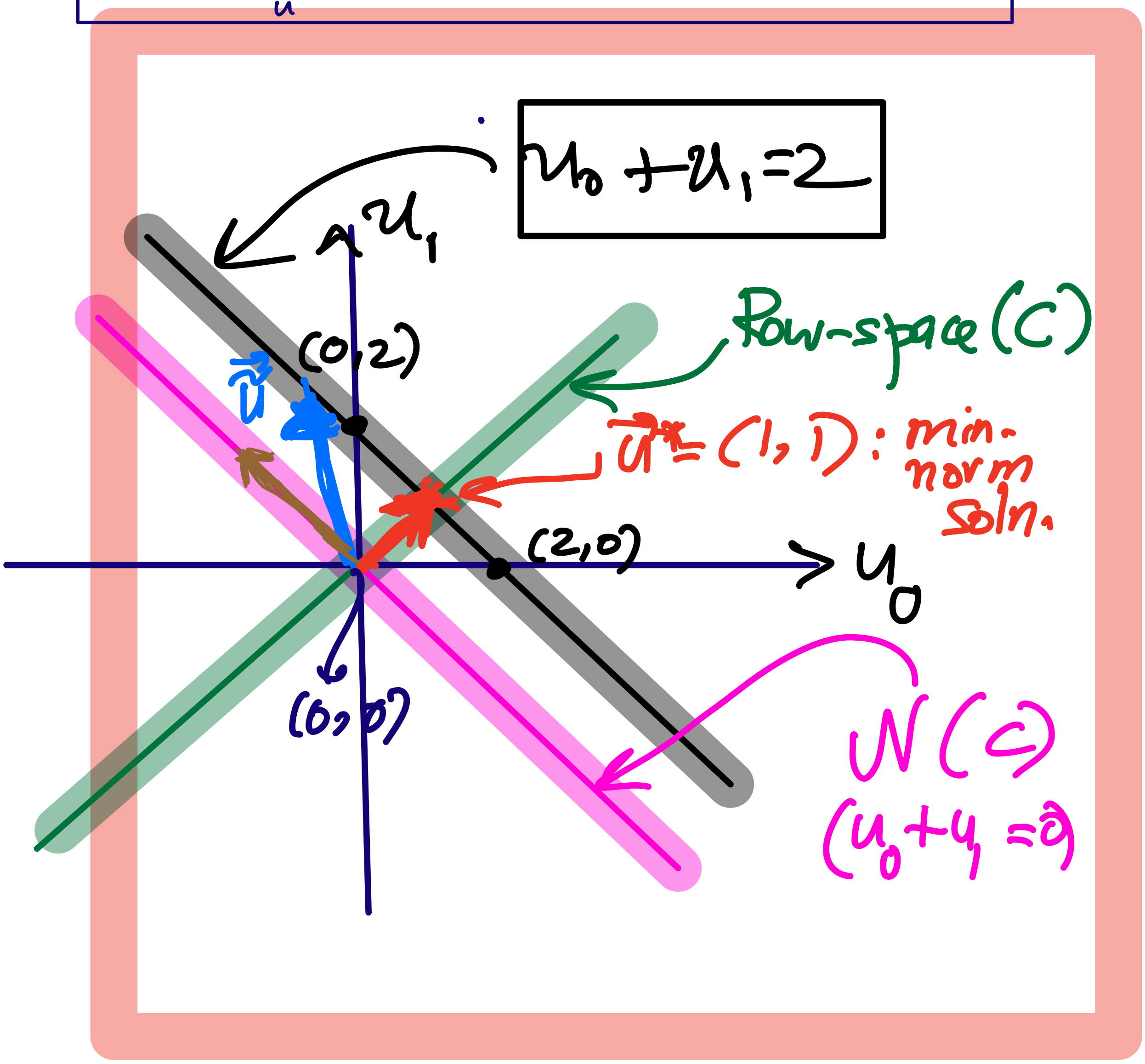
$\vec{u}^*$ has no component in the Null-space (C).

Why? See BLUE soln. $\vec{u}$ that has larger length than $\vec{u}^*$.

$$\text{Null-Space}(C) = \beta \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$\vec{u}^*$, the min. norm soln. should be such that it has NO PROJECTION on the Null-Space (C).

$$\vec{u}^* = \underset{\vec{u}}{\operatorname{argmin}} \|u\|^2 \quad \text{s.t.} \quad u_0 + u_1 = 2$$



$\vec{u}^*$ (the minimum-norm soln.) (red vector $\vec{u}^*$) has no component in the null-space of C (green line). Any non-zero component in the null-space of C (brown vector) is "wasted" energy for $\vec{u}$.

General min-norm soln:

$\vec{u}^* = \underset{\vec{u}}{\arg\min} \|\vec{u}\|^2$ s.t. $C\vec{u} = \vec{d}$

We saw that we want $\text{Proj}_{Null(C)} \vec{u}^* = 0$

(Why? If $\vec{u}_{soln} = \vec{u}^* + \vec{u}_{null}$, then
$\|\vec{u}_{soln}\|^2 = \|\vec{u}^* + \vec{u}_n\|^2 = \|\vec{u}^*\|^2 + \|\vec{u}_n\|^2 + 2\cancelto{0}{\langle \vec{u}^*, \vec{u}_n \rangle}$
$\geq \|\vec{u}^*\|^2$ with equality iff $\vec{u}_n = 0$)

From Linear Algebra, we know that Row-space(C) is orthogonal to $Null(C)$.

$\Rightarrow \vec{u}^* \in$ Rowspace(C).

$\vec{u}^* = C^T\vec{w}$ for some $\vec{w}$ ($\because$ Row-space (C) = Col-space (C^T))

We also want $C\vec{u}^* = \vec{d}$

$\Rightarrow \quad CC^T\vec{w} = \vec{d} \Rightarrow \vec{w} = (CC^T)^{-1}\vec{d}$

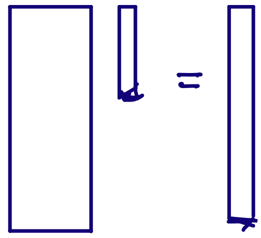
$\Rightarrow \boxed{\vec{u}^* = C^T(CC^T)^{-1}\vec{d}}$ $\circledast$

min-norm soln.

For our ex. $C = [1 \quad 1]$; $d = 2$

$\Rightarrow \vec{u}^* = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{C^T} \cdot \underbrace{\left([1 \; 1]\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1}}_{(CC^T)^{-1}} \cdot \underbrace{2}_{\vec{d}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ✓

Contrast with Least Squares setting:



$C \vec{u} = \vec{d}$

$$\boxed{\vec{u}_{LS} = (C^T C)^{-1} C^T \vec{d}} \circledast$$

Contrast the 2 solns. $\vec{u}_{MN} = \vec{u}^*$ and $\vec{u}_{LS}$

min. norm

Stay tuned for the SVD

we will show that $\vec{u}^* = C^+ \vec{d}$

pseudo-inverse of C

handles both settings simultaneously!

$$C\vec{u} = \vec{d}$$

$$\vec{u} = C^{-1} \vec{d}$$

RECAP

Let's now recall our study of symmetric matrix S:

① Symmetric matrix S can always be diagonalized.

② The diagonalizing basis/matrix V is made up of the eigenvectors of S that are orthonormal.

③ All the eigenvalues of S are real.

$$\underbrace{V^T}_{} S \underset{\substack{\uparrow\\ \text{orthonormal basis}}}{V} = \underset{\text{diagonal}}{\Lambda} \quad \begin{bmatrix} \begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{matrix} & \\ & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix} \end{bmatrix}$$

$$\Rightarrow S = V \Lambda V^T$$

$$S V = V \Lambda$$

$$S \Big[\underbrace{\vec{v}_1 \; \vec{v}_2 \cdots \vec{v}_r}_{\substack{\text{correspond}\\ \text{to non-zero}\\ \text{e-vals}}} \Big| \underbrace{\vec{v}_{r+1} \cdots \vec{v}_n}_{\substack{\text{correspond}\\ \text{to zero}\\ \text{e-vals}}} \Big] = \Big[\vec{v}_1 \cdots \vec{v}_r \Big| \vec{v}_{r+1} \cdots \vec{v}_n \Big] \begin{bmatrix} \begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{matrix} & 0 \\ 0 & \begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix} \end{bmatrix}$$

$$\left.\begin{aligned} S\,\vec{v}_{r+1} &= 0 \\ S\,\vec{v}_{r+2} &= 0 \\ &\vdots \\ S\,\vec{v}_n &= 0 \end{aligned}\right\} \Rightarrow \vec{v}_{r+1}, \vec{v}_{r+2}, \ldots \vec{v}_n$$

form an orthonormal basis for $\underbrace{\text{Null-space}(S)}_{\mathcal{N}(S)}$

Warmup for the SVD:

① $V^{-1}AV = \Lambda$ for (square matrix) A:

$$A = V\Lambda V^{-1}$$

$V = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$ eigenvectors of A

(A: square matrix)

(A: square matrix)

② $U^T A U = T$ upper-triangular

$$A = UTU^T$$

$U = [\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_n]$ ON-basis for $\mathbb{R}^n$

---

③ S: square symmetric matrix ($S = S^T$)

$$Q^T S Q = \Lambda$$

$Q = [\vec{q}_1\ \vec{q}_2 \cdots \vec{q}_n]$ ON-basis + e-vectors of S

Q) What is a good decomposition for a general (non-square) matrix A?

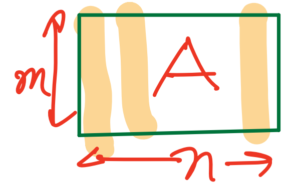
- We love an orthonormal basis and we love diagonalization but don't want to rely on special structures like symmetric matrices or even square matrices!

Let us see how to generalize the concept of EIGENVALUE and EIGENVECTOR for square matrices to a similar concept for rectangular matrices, while insisting that we have orthonormal bases.

Key insight: We need TWO orthonormal bases, one for the Column Space(A) & one for the Row-Space(A).

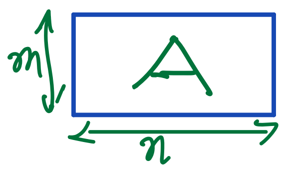
- $A\vec{v}_i = \cancel{\lambda_i \vec{v}_i}$; $(\lambda_i, \vec{v}_i)$ are an e-val/e-vector pair for A

How about $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$



We now have 2 orthonormal bases:

- $\{\vec{u}_i\}$ for the col. space $(A) \in \mathbb{R}^m$ and
- $\{\vec{v}_i\}$ for the row space $(A) \in \mathbb{R}^n$



- Assume $m < n$
- $\text{Rank}(A) = m$

In matrix form:

$$\underbrace{A}_{A_{m\times n}} \underbrace{\begin{bmatrix} \vec{v}_1 \cdots \vec{v}_m \cdots \vec{v}_n \end{bmatrix}}_{V_{n\times n}} = \underbrace{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}}_{U_{m\times m}} \underbrace{\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \sigma_2 \ddots & & 0 \\ 0 & & \sigma_m & \end{array}\right]}_{\Sigma_{m\times n}}$$

$$A\,V = U\,\Sigma$$

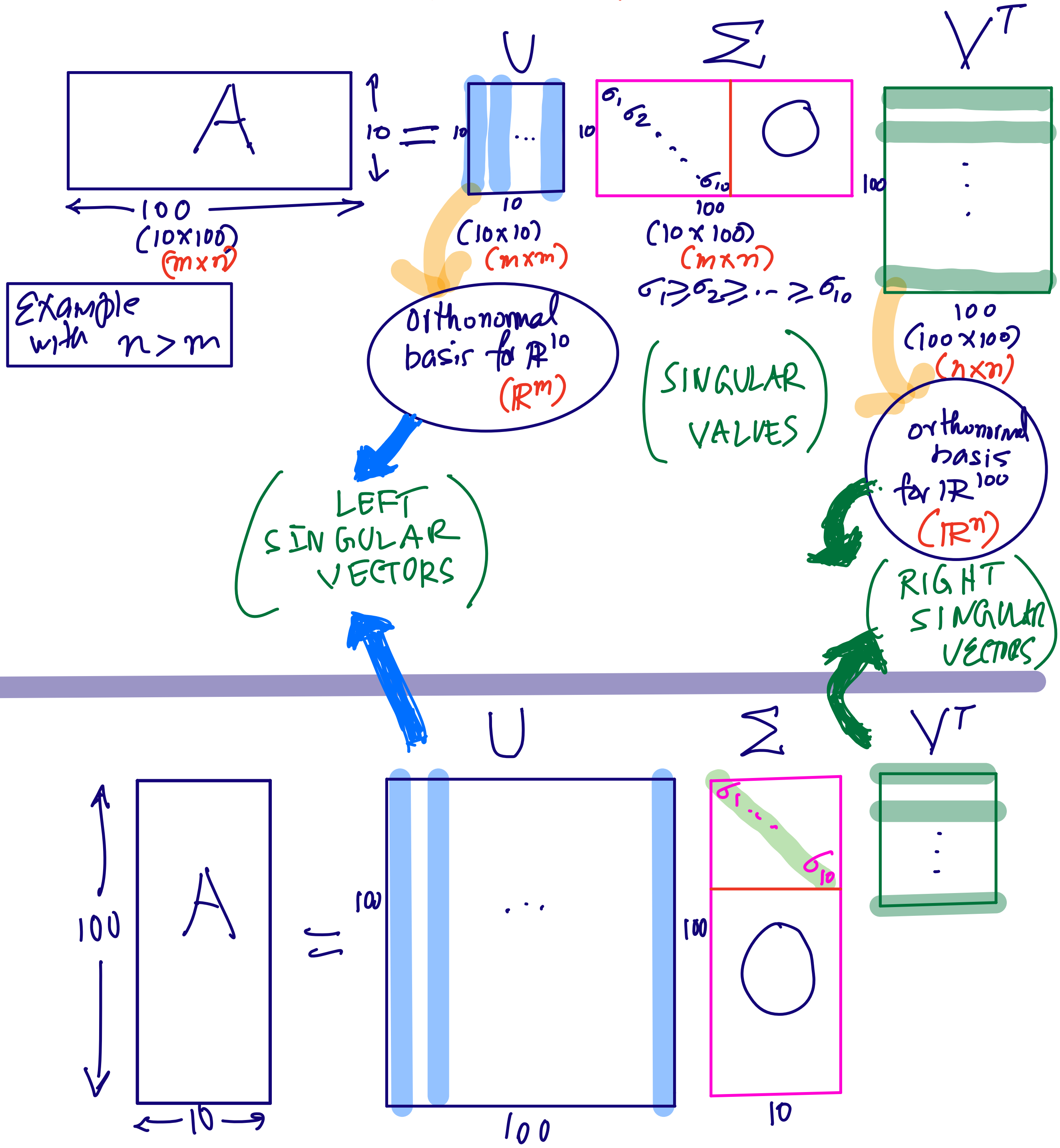
SINGULAR VALUE DECOMPOSITION (SVD)

$$\boxed{A = U \Sigma V^T}$$

$\text{rank}(A) = \min(m, n) = m$

Singular Value Decomposition (SVD)

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



$$A V = U \Sigma$$

$$\Rightarrow \boxed{A = \underbrace{U}_{\text{left singular vectors}} \underbrace{\Sigma}_{\text{singular values along diagonal}} \underbrace{V^T}_{\text{right singular vectors}}}$$

"Full-SVD" of A

Ex. 1

$$A = \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix} = \underbrace{\frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{50} & \\ & 0 \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^T}$$

$\vec{u}_1 = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}$

Two orthonormal bases,

$U = [\vec{u}_1 \ \vec{u}_2]$ for $Col(A) \in \mathbb{R}^2$

and $V = [\vec{v}_1 \ \vec{v}_2]$ for $Row(A) \in \mathbb{R}^2$

Ex. 2

$$\boxed{A = \begin{bmatrix} 1 & 2 & 4 & 5 \\ 2 & 4 & 8 & 10 \end{bmatrix}_{2\times 4}}$$

$$A = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}}_{\substack{U \\ 2\times 2}} \underbrace{\begin{bmatrix} \sqrt{230} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\substack{\Sigma \\ 2\times 4}} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{46}} & \sqrt{\frac{2}{23}} & 2\sqrt{\frac{2}{23}} & \frac{5}{\sqrt{46}} \\ -\frac{5}{\sqrt{26}} & 0 & 0 & \frac{1}{\sqrt{26}} \\ -\frac{2}{\sqrt{273}} & 0 & \sqrt{\frac{13}{21}} & -\frac{10}{\sqrt{273}} \\ -\frac{1}{\sqrt{483}} & \sqrt{\frac{21}{23}} & -\frac{4}{\sqrt{483}} & -\frac{5}{\sqrt{483}} \end{bmatrix}}_{\substack{V^T \\ 4\times 4}}$$

How do we get this decomposition?

<u>Key insight:</u> Convert $A_{m\times n}$ into a square matrix.

How?

Form $A^T A$!

$$\left({}_{100}\overset{10}{\boxed{A^T}} \; {}_{10}\overset{100}{\boxed{A}} = \overset{100}{\boxed{A^TA}}{}_{100} \right)$$

Consider $\boxed{A^T A}$

If we use SVD, $A = U \Sigma V^T$,

then
$$
\begin{aligned}
A^T A &= (U \Sigma V^T)^T U \Sigma V^T \\
&= V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T \\
&= V \underbrace{\Sigma^T \Sigma}_{\Sigma^2} V^T = \boxed{V \Sigma \Sigma V^T}
\end{aligned}
$$

Similarly,
$$
\begin{aligned}
A A^T &= U \Sigma \underbrace{V^T V}_{I} \Sigma^T U^T \\
&= \boxed{U \Sigma \Sigma^T U^T}
\end{aligned}
$$

Suggests that the key to understanding the SVD of $A = U \Sigma V^T$ is to study the square matrix $A^T A$!!!

Consider $C^T C$

Sorry, we will use C and A interchangeably!

Fact: $C^T C$ is symmetric

Proof: $(C^T C)^T = C^T C$ $\square$

Special property: eigenvalues of $S = C^T C$ are always real and non-negative.