EECS16B DIS10B remote

Learning objectives

1. How to compute the Singular Value Decomposition (SVD)
2. What properties of a matrix, A, the SVD of A, $U\Sigma V^T$ can tell you about (~ means related to, below)
   - rank ($\sim\Sigma$)
   - columnspace ($\sim U, \Sigma$)
   - what goes to the columnspace ($\sim\Sigma, V$)
   - nullspace ($\sim\Sigma, V$)
   - what directions can't be reached by $A\vec{x}$ ($\sim U, \Sigma$)
   - what directions shrink or grow the most when you pass it through the matrix ($\sim\Sigma, V$)

$$\vec{y} = \underset{m\times n}{A}\vec{x} \quad \vec{x} \in \mathbb{R}^n \quad \vec{y} \in \mathbb{R}^m$$

1. Computing the SVD: A "Tall" Matrix Example

Define the matrix

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}. \qquad \vec{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \qquad A = \begin{bmatrix} | & | \\ \vec{v} & -\vec{v} \\ | & | \end{bmatrix}$$

(a) In this part, we will find the full SVD of A in steps.

Compute A^TA or AA^T and find its eigendecomposition/diagonalization

(i) **Compute $A^\top A$ and find its eigenvalues.**

$$A^TA = \begin{bmatrix} - & \vec{v}^T & - \\ - & -\vec{v}^T & - \end{bmatrix} \begin{bmatrix} | & | \\ \vec{v} & -\vec{v} \\ | & | \end{bmatrix}$$

$$\vec{v}^T\vec{v} = 1^2 + (-2)^2 + 2^2 = 9$$

$$= \begin{bmatrix} \vec{v}^T\vec{v} & -\vec{v}^T\vec{v} \\ -\vec{v}^T\vec{v} & \vec{v}^T\vec{v} \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

→ columns linearly dependent, has nullspace $\lambda_1 = 0$

Observation: A^TA is symmetric

(ii) **Find orthonormal eigenvectors $\vec{v}_i$ of $A^\top A$ (right singular vectors, columns of V).**

We know $\boxed{\lambda_1 = 0}$ By inspection

$$\begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}\vec{v}_1 = \vec{0} \qquad \boxed{\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \qquad ||\vec{v}_1|| = \sqrt{2}$$

Guess that second eigenvector is

$$\boxed{\vec{v}_2^T\vec{v}_1 = 0} \qquad \boxed{\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}} \qquad ||\vec{v}_2|| = \sqrt{2}$$

using orthogonality of eigenvectors of symmetric A^TA

$$A^TA\vec{v}_2 = \begin{bmatrix} 18 \\ -18 \end{bmatrix} = 18\begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \boxed{\lambda_2 = 18}$$

* $\vec{v}_1, \vec{v}_2$ are not $\vec{v}$ in A

$$\boxed{\vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}}$$

wrote these because wanted normalized eigenvectors

(iii) **Find singular values, $\sigma_i = \sqrt{\lambda_i}$.**

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{0} = \boxed{0}$$

$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{18} = \sqrt{9 \cdot 2} = \boxed{3\sqrt{2}}$$

Q: Why are you using $\vec{v}_2^T\vec{v}_1 = 0$?

A: $A^TA \rightarrow (A^TA)^T = A^T(A^T)^T = A^TA$ ← symmetric

If symmetric can write orthonormal eigenvectors that diagonalize it

I know one eigenvector $\vec{v}_1$

To find $\vec{v}_2$, I can find $\vec{v}_2$ such that $\vec{v}_1^T\vec{v}_2 = 0$ (orthogonal eigenvectors)

$\vec{v}_2^T\vec{v}_1$ same (scalar)

(in this case, inspect for)

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \qquad \sigma_2 = 0 \qquad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\sigma_1 = 3\sqrt{2} \qquad \boxed{\vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}}$$

(iv) **Find the orthonormal vectors $\vec{u}_i$ (and for nonzero σ, you can use $\vec{v}_i$).**
Hint: given $\vec{v}_k$ corresponding to nonzero σ, we can compute $\vec{u}_k = \frac{1}{\sigma_k} A\vec{v}_k$.
Another hint: How can we extend a basis, and why is that needed here? Note what the Jupyter notebook contains.

$$\underset{3\times 2}{A} = \underset{3\times 3}{U}\,\underset{3\times 2}{\Sigma}\,\underset{2\times 2}{V^T} \quad \text{SVD of } A$$

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{3\sqrt{2}\sqrt{2}} \begin{bmatrix} 2 \\ -4 \\ 4 \end{bmatrix} = \boxed{\frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}}$$

$$\|\vec{u}_1\| = 1$$

$$A\vec{v}_1 = \underbrace{\begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - \vec{v}_1^T - \\ - \vec{v}_2^T - \end{bmatrix}}_{V^T} \vec{v}_1$$

$$= \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1$$

Try $A\vec{v}_2 = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{0}$

To find other $\vec{u}_i$'s use GS.

Q: Is the output of Gram Schmidt unique?

$\vec{v}_1, \vec{v}_2, \vec{v}_3$, $\vec{v}_2, \vec{v}_1, \vec{v}_3$

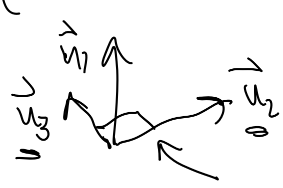
GS on these ↑ GS on these ↑

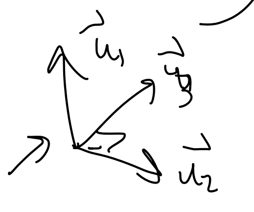
A: Not the same!

SVD is not unique

① GS

② A^TA can have eigenvectors of $\vec{v}_i$ or $-\vec{v}_i$





$$U = \begin{bmatrix} \frac{1}{3} & \frac{4}{3\sqrt{2}} & 0 \\ -\frac{2}{3} & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \leftarrow \text{GS with } \vec{u}_1, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$\sigma_1 = 3\sqrt{2}$ $\quad \vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ $\quad \vec{u}_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$ $\quad \vec{u}_2 = \begin{bmatrix} \frac{4}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \end{bmatrix}$ $\quad \vec{u}_3 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

$\sigma_2 = 0$ $\quad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

(v) **Use the previous parts to write the full SVD of A.**

$A = U \Sigma V^T$ "full SVD"

$m \times n$

$$= \underbrace{\begin{bmatrix} \frac{1}{3} & \frac{4}{3\sqrt{2}} & 0 \\ -\frac{2}{3} & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\substack{U \\ m \times m \\ 3 \times 3}} \underbrace{\begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\substack{\Sigma \\ m \times n \\ 3 \times 2}} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\substack{V^T \\ n \times n \\ 2 \times 2}}$$

(vi) **Use the Jupyter notebook to run the code cell that calls `numpy.linalg.svd` on A. What is the result? Does it match our result above?**

```
A = np.array([[1, -1], [-2, 2], [2, -2]])
print_full_svd(A)
```

```
U:
[[-0.33333333  0.66666667 -0.66666667]
 [ 0.66666667  0.66666667  0.33333333]
 [-0.66666667  0.33333333  0.66666667]]
Sigma:
[[4.24264069 0.        ]
 [0.         0.        ]]
V:
[[-0.70710678  0.70710678]
 [ 0.70710678  0.70710678]]
```

☆ Non-uniqueness
frame picture

$\Sigma = \begin{bmatrix} \sim 4.24 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\quad V^T = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

$U = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ $\quad V = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

Σ is always the same

$\sigma_i \geq 0, \quad \sigma_1 \geq \sigma_2 \geq \cdots$

(b) **Find the rank of A.** - # of non-zero singular values

(c) **Find a basis for the range (or column space) of A.**

(d) **Find a basis for the null space of A.**

$$A = \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \boxed{3\sqrt{2}} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2 & - \end{bmatrix}$$

rank$(A) = 1$

$C(A) = \text{Span}\{\vec{u}_1\}$ ← $\{\vec{u}_1\}$ is the basis

vectors in U that correspond to nonzero singular values

$A\vec{v}_1 = \sigma_1 \vec{u}_1$

$A\vec{v}_2 = \sigma_2 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$

rank$(A) = \dim(C(A))$ ← columnspace

$N(A) = \text{Span}\{\vec{v}_2\}$ ← $\{\vec{v}_2\}$ is the basis

vectors in V that correspond to $\sigma_i = 0$ or don't have a singular value matching

(e) We now want to create the SVD of $A^\top$. Rather than repeating all of the steps in the algorithm, feel free to use the jupyter notebook for this subpart (which defines a `numpy.linalg.svd` command). **What are the relationships between the matrices composing A and the matrices composing $A^\top$?**

$$A^T = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix}$$

$$A^T = (\underbrace{U_A \Sigma_A V_A^T}_{\text{SVD of } A})^T = V_A \Sigma_A^T U_A^T = U_{A^T} \Sigma_{A^T} V_{A^T}^T$$

$U_A = V_{A^T}$

$V_A = U_{A^T}$

$\boxed{\Sigma_A^T = \Sigma_{A^T}}$ always true

$$A^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \; & \; & \; \end{bmatrix}^T$$

$A^T\vec{u}_3 = \vec{0}\ (\mathbb{R}^2)$ $A^T\vec{u}_2 = \vec{0}\ (\mathbb{R}^2)$

Fill this in w/ $\vec{u}_1, \vec{u}_2, \vec{u}_3$

2. **Understanding the SVD**

We can compute the SVD for a wide matrix A with dimension $m \times n$ where $n > m$ using $A^\top A$ with the method covered in lecture. However, when doing so, you may realize that $A^\top A$ is much larger than $AA^\top$ for such wide matrices. This makes it more efficient to find the eigenvalues for $AA^\top$. In this question, we will explore how to compute the SVD using $AA^\top$ instead of $A^\top A$.

(a) **What are the dimensions of $AA^\top$ and $A^\top A$?**

A
$m \times n$

$A \; A^T = AA^T$ ($m \times m$)

A^T ($n \times m$) A ($m \times n$) $= A^TA$ ($n \times n$)

When computing SVD, compute the smaller one!

(b) **Given that the SVD of A is $A = U\Sigma V^\top$, find a symbolic expression for $AA^\top$ in terms of U, Σ, $V^\top$.** Simplify where possible!

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$

$$\Sigma\Sigma^T = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & \ddots & \\ & & & \sigma_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & \sigma_m \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \; \sigma_m^2 \end{bmatrix}$$

$$= U\Sigma V^T V \Sigma^T U^T = \boxed{U\Sigma\Sigma^T U^T} = U \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \; \sigma_m^2 \end{bmatrix} U^T$$

$U \quad \Sigma \quad V^T \quad V \quad \Sigma^T \quad U^T$

(c) Using the solution to the previous part, **how can we find a U and Σ from $AA^\top$?** *Hint: first, think about matrix dimensions. Next, consider the properties of the SVD, and what each matrix signifies. Another Hint: you may want to compute for yourself, based on the structure of Σ, what $\Sigma^\top\Sigma$ and $\Sigma\Sigma^\top$ are.*

U's columns are eigenvectors of AA^T

$\Sigma\Sigma^T$ has as diagonal entries eigenvalues of AA^T

$\lambda_i = \sigma_i^2 \qquad \sigma_i = \sqrt{\lambda_i}$

$AA^T = U\Lambda U^{-1} \Rightarrow \Lambda = \Sigma\Sigma^T$

diagonal matrix of eigenvalues

(d) Now that we have found the singular values σ_i and the corresponding vectors $\vec{u}_i$ in the matrix U, **can you find the corresponding vectors $\vec{v}_i$ in V?** *Hint: Apply the definition of an eigenvector. What do the $\vec{v}_i$ vectors signify with regards to $A^\top A$?*

$$\underset{m\times n}{A} = U\Sigma V^T \qquad \underset{m\times 1}{\vec{u}_i}$$

$$\cancel{\underset{m\times n}{A}\ \underset{n\times 1}{\vec{v}_i}\ ?}$$

$$\Sigma^T = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \sigma_m \\ & & 0 & \end{bmatrix}$$

$$A^T\vec{u}_i = (U\Sigma V^T)^T\vec{u}_i$$

$$= V\Sigma^T \underbrace{U^T\vec{u}_i}$$

$$= V\Sigma^T \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{i-th location}$$

$$= V\sigma_i \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{ith position}$$

$$= \sigma_i\vec{v}_i$$

$$\Rightarrow \quad \vec{v}_i = \frac{1}{\sigma_i}A^T\vec{u}_i$$

(e) Now we have a way to find the vectors $\vec{v}_i$ in matrix V! **Use the fact that the vectors $\underbrace{\vec{u}_i, \vec{u}_j}$ are orthonormal to show that $\vec{v}_i, \cancel{\vec{v}_j}$ $\vec{v}_j$ in V (corresponding to nonzero σ_i, σ_j and $i, j \le n$) are orthonormal by direct computation.**

$$\underbrace{\vec{v}_i^T\vec{v}_j} = \left(\frac{1}{\sigma_i}A^T\vec{u}_i\right)^T\left(\frac{1}{\sigma_j}A^T\vec{u}_j\right)$$

$$= \frac{1}{\sigma_i\sigma_j}\vec{u}_i^T\underbrace{AA^T\vec{u}_j}$$

$$= \frac{1}{\sigma_i\sigma_j}\vec{u}_i^T\sigma_j^2\vec{u}_j$$

$$= \frac{\sigma_j^2}{\sigma_i\sigma_j}\vec{u}_i^T\vec{u}_j \leftarrow \begin{matrix} i=j \rightarrow 1 \\ i\neq j \rightarrow 0 \end{matrix}$$

(f) **[Practice]** Given that $A = U\Sigma V^\top$, **verify that the vectors after the first** ~~n~~ m **vectors in** V **are in the nullspace of** A**.**

Sorry, didn't complete these during, but notes are here.

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T} = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \Bigg[\underbrace{\begin{matrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \; \sigma_m \end{matrix}}_{m\times m} \; \underbrace{\begin{matrix} | & & | \\ \vec{0} & \cdots & \vec{0} \\ | & & | \end{matrix}}_{m\times n-m} \Bigg] \begin{bmatrix} - \vec{v}_1^T - \\ - \vec{v}_2^T - \\ \vdots \\ - \vec{v}_n^T - \end{bmatrix}$$

$$A\vec{v}_i = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & | & & | \\ & \sigma_2 & & \vec{0} & \cdots & \vec{0} \\ & & \ddots \; \sigma_m & | & & | \end{bmatrix} \begin{bmatrix} - \vec{v}_1^T - \\ - \vec{v}_2^T - \\ \vdots \\ - \vec{v}_n^T - \end{bmatrix} \vec{v}_i$$

$$= \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & | & & | \\ & \sigma_2 & & \vec{0} & \cdots & \vec{0} \\ & & \ddots \; \sigma_m & | & & | \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \vec{v}_i \\ \vec{v}_2^T \vec{v}_i \\ \vdots \\ \vec{v}_n^T \vec{v}_i \end{bmatrix}$$

I went ham on calculations but you don't have to go as intense. The main reason I do this is for visibility/understanding

$$= \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & | & & | \\ & \sigma_2 & & \vec{0} & \cdots & \vec{0} \\ & & \ddots \; \sigma_m & | & & | \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{ith position}$$

$$= \begin{cases} \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \left(\begin{bmatrix} 0 \\ \vdots \\ \sigma_i \\ \vdots \\ 0 \end{bmatrix} \cdot 1 + \begin{bmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdot 0 + \begin{bmatrix} 0 \\ \sigma_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdot 0 + \cdots + \begin{bmatrix} \vec{0} \end{bmatrix} \cdot 0 + \cdots \begin{bmatrix} \vec{0} \end{bmatrix} \cdot 0 \right) & \text{if } i \leq m \\ \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \left(\begin{bmatrix} \sigma_1 \\ \\ \\ \end{bmatrix} \cdot 0 + \begin{bmatrix} \\ \sigma_2 \\ \\ \end{bmatrix} \cdot 0 + \cdots + \begin{bmatrix} \\ \\ \sigma_m \end{bmatrix} \cdot 0 + \begin{bmatrix} \vec{0} \end{bmatrix} \cdot 0 + \cdots + \begin{bmatrix} \vec{0} \end{bmatrix} \cdot 1 + \cdots \right) & \text{if } i > m \end{cases}$$

$$= \begin{cases} \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \sigma_i \\ \vdots \\ 0 \end{bmatrix} = \sigma_i \vec{u}_i & \text{if } i \leq m \\ \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \\ | & | & & | \end{bmatrix} \vec{0} = \vec{0} & \text{if } i > m \end{cases}$$

$$\Rightarrow A\vec{v}_i = \vec{0} \text{ if } i > m$$

$$\Rightarrow \vec{v}_i \in N(A) \; \forall i > m$$

(g) **[Practice]** Using the previous parts of this question and what you learned from lecture, **write out a procedure on how to find the SVD for *any* matrix.**

Didn't complete this during but notes are here

A wide

A

A tall

A

① ~~Compute smaller of AA^T, A^TA~~

$A \; A^T = AA^T$, $A^T \; A = A^TA$

⇑ for wide case

$A \; A^T = AA^T$, $A^T \; A = A^TA$

⇓ For tall case

② Compute eigenvalues, eigenvectors

wide: $AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$

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

$= U\Sigma\Sigma^T U^T$

eigenvalues : $\boxed{\lambda_i = \sigma_i^2}$

eigenvectors : $\boxed{\vec{u}_i}$

tall: $A^TA = (U\Sigma V^T)^T (U\Sigma V^T)$

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

$= V\Sigma^T\Sigma V^T$

eigenvalues : $\boxed{\lambda_i = \sigma_i^2}$

eigenvectors : $\boxed{\vec{v}_i}$

③ ~~Compute other vectors~~

wide: Compute $\vec{v}_i$ from $\vec{u}_i$

For $\vec{u}_i \sim \sigma_i \neq 0$

$A^T\vec{u}_i = \sigma_i \vec{v}_i \Rightarrow \boxed{\vec{v}_i = \frac{1}{\sigma_i} A^T \vec{u}_i}$

For other $\vec{v}_i$, use GS to complete

tall: compute $\vec{u}_i$ from $\vec{v}_i$

For $\vec{v}_i \sim \sigma_i \neq 0$

$A\vec{v}_i = \sigma_i \vec{u}_i \Rightarrow \boxed{\vec{u}_i = \frac{1}{\sigma_i} A \vec{u}_i}$

For other $\vec{u}_i$, use GS to complete.

SVD reference sheet

For illustration if A is tall

$$\underset{}{A} = \underset{m\times m}{\underbrace{\begin{bmatrix} | & | & & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_r & \cdots & \vec{u}_m \\ | & | & & | & & | \end{bmatrix}}_{U}} \; \underset{m\times n}{\underbrace{\begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & \ddots \ \sigma_m \\ \hline & & 0_{m-n\times n} & & \end{bmatrix}}_{\Sigma}} \; \underset{n\times n}{\underbrace{\begin{bmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2^T & - \\ & \vdots & \\ - & \vec{v}_r^T & - \\ - & \vec{v}_n^T & - \end{bmatrix}}_{V^T}}$$

For A wide, just transpose everything

These stuff below true for both tall & wide

If $\sigma_i \neq 0$ for $i = 1, 2, 3, \ldots, r$
and $\sigma_i = 0$ for $i = r+1, r+2, \ldots, \min(m,n)$

$C(A) = \text{Span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$

$N(A) = \text{Span}\{\vec{v}_{r+1}, \vec{v}_{r+2}, \ldots, \vec{v}_n\}$

$\text{rank}(A) = r$ (# nonzero σ_i)

Directions not in columnspace: $\text{Span}\{\vec{u}_{r+1}, \vec{u}_{r+2}, \ldots, \vec{u}_m\}$

Directions not in nullspace: $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$
($\vec{x}$ s.t. $A\vec{x} \neq \vec{0}$)

If $\sigma_1 > \sigma_2 > \cdots > \sigma_r > \sigma_{r+1} = 0$

$\|A\vec{v}_i\| = \|\sigma_i \vec{u}_i\| = |\sigma_i| \, \|\vec{u}_i\| = \sigma_i$

So $\|A\vec{v}_1\| > \|A\vec{v}_2\| > \|A\vec{v}_3\| > \cdots$

$\vec{v}_1$ grows the biggest under A

$\vec{v}_2$ grows next biggest

etc.