

SVD I

- * Introduction
- * Calculating the SVD
- * SVD & Fundamental Matrix Subspaces
↳ Worksheet!

Helpful resource: https://medium.com/@jonathan_hui/machine-learning-singular-value-decomposition-svd-principal-component-analysis-pca-1d45e885e491

Logistics:

- * MT1 on Friday 7/17, 5-7pm (PST)
- * HW4 due 7/21 (on SVD, PCA, linearization)
- * HW Party moved to F, M 4-6pm
This week! F HW party is at 1-3pm

Review sessions: Thu, Wed 8-9pm

ckts ↗
linear algebra ↖

(I) Introduction to SVD

$$n \times n, \quad \Lambda \quad | \quad | \quad | \quad | \quad |^{-1}$$

$n \times n$ matrices: $A = V \Delta V^{-1}$

If $A^* = A$ (Hermitian),

$$A = V \Delta V^*$$

unitary matrix

diagonal (real-valued)

We factorized A into a product of interesting matrices

\Rightarrow Would be great if we could extend factorization to any $m \times n$ matrix

* SVD *

Viewpoint 1: SVD as a matrix factorization

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

$$A = U \Sigma V$$

\downarrow \uparrow \uparrow \uparrow
 $m \times n$ $m \times m$ $m \times n$ $n \times n$

U, V are square unitary matrices;

$$U^* U = I_{m \times m}$$

$$V^* V = I_{n \times n}$$

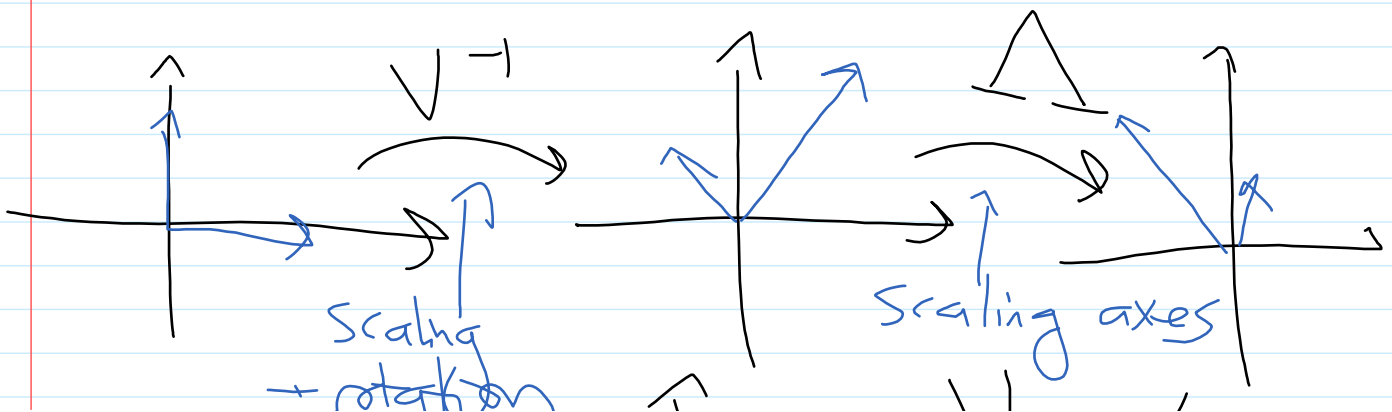
Σ is a rectangular diagonal matrix
w/ diagonal entries $\sigma_i = \sum j_{ii} \geq 0$

($\sigma_i > 0$ for $1 \leq i \leq r$)
where $r = \text{rank } A$)

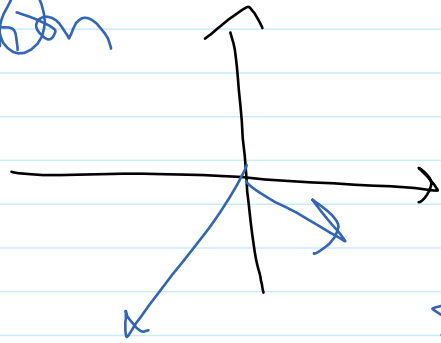
Interpretation?

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

Similar: $A \vec{x} = V \Lambda V^{-1} \vec{x}$



scaling + rotation

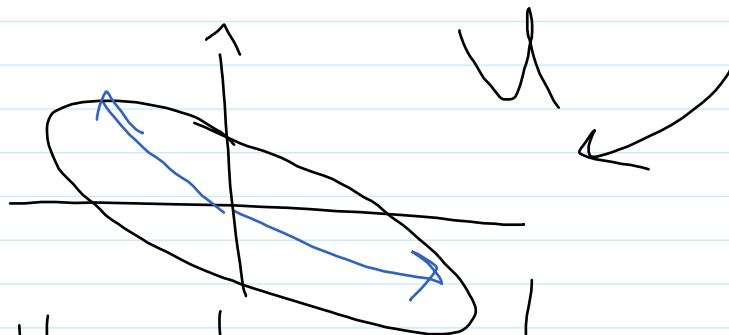
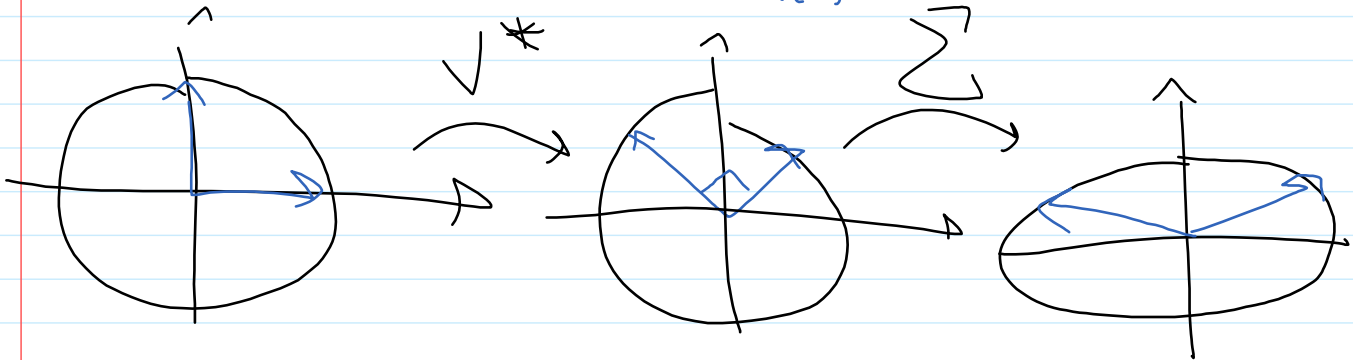


rotating axes

scaling + rotation

$$\text{SVD: } A = U \Sigma V^*$$

unitary matrix rotation
rect diag scaling the axes
unitary matrix rotation



More on this tomorrow!
(Geometric interpretation)

Viewpoint 2: SVD as a sum of rank 1 matrices

$$A = U \Sigma V^* = \begin{pmatrix} | & | & \dots & | \\ u_1 & u_2 & \dots & u_m \\ | & | & \dots & | \end{pmatrix} \left(\begin{array}{ccc|c} \sigma_1 & 0 & & 0 \\ & \ddots & & \\ 0 & & \sigma_m & \\ \hline & & & 0 \end{array} \right)$$

(Assumed $n > m$ but end result will be the same for $n < m$)

$$\times \begin{pmatrix} | & | & \dots & | \\ v_1^* & v_2^* & \dots & v_n^* \\ | & | & \dots & | \end{pmatrix}$$

$$= \begin{pmatrix} | & | & \dots & | \\ \sigma_1 u_1 & \dots & \sigma_m u_m & \\ | & | & \dots & | \end{pmatrix} \begin{pmatrix} | & | & \dots & | \\ v_1^* & \dots & & \\ | & | & \dots & | \\ \vdots & & & \\ v_n^* & & & \end{pmatrix}$$

"outer products"

$$x y^T = \begin{pmatrix} | \\ \dots \\ | \end{pmatrix} \begin{pmatrix} | & \dots & | \end{pmatrix} \quad \begin{matrix} m \times 1 \\ 1 \times n \end{matrix}$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1^* & \dots & y_n^* \end{pmatrix} = \begin{pmatrix} x_1 y_1^* & & \\ & \ddots & \\ x_m y_1^* & & \\ & & \ddots & \\ & & x_m y_n^* & \end{pmatrix}$$

= $m \times n$ rank 1 matrix

rows all lin. dependent!

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^* + \sigma_2 \vec{u}_2 \vec{v}_2^* + \dots + \sigma_m \vec{u}_m \vec{v}_m^*$$

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^* + \sigma_2 \vec{u}_2 \vec{v}_2^* + \dots + \sigma_m \vec{u}_m \vec{v}_m^*$$

Recall that $\sigma_i = 0$ for $i > r$

$$A = \sum_{i=1}^r \sigma_i \vec{u}_i \vec{v}_i^*$$

Compact
SVD

Sum of r rank-1 matrices

Benefit of this viewpoint?

\Rightarrow low rank matrix approx

$$A \in \mathbb{C}^{m \times n} \Rightarrow mn \text{ entries}$$

$$A = \sum \sigma_i \vec{u}_i \vec{v}_i^* \Rightarrow r(m+n) \text{ entries}$$

Note if

$$r(m+n) \ll mn$$

compress

Can further approximate / ~~compress~~ to

a good degree if only a few singular values are significant:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$$

only take these values!
"truncated SVD"

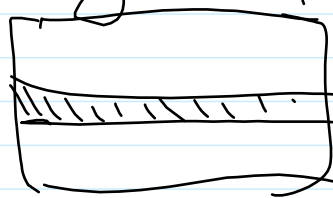
What is the "best" low rank matrix approx?

⇒ SVD

Eckart - Young - (Mirsky) Thm

- won't prove here
- Wikipedia's proof incorrect?

EX: image compression



$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Viewpoint 3: Abstract SVD

(HW3, Q6)

Matrix Form: $A = U \Sigma V^*$

$$V^* V = I$$

$$U^* U = I$$

$$\Rightarrow AV = U \Sigma$$

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

In general, SVD of a linear map

$$f: V \rightarrow W$$

is comprised of:

1) orthonormal basis for V

(right singular vecs)

2) " " for W

(left singular vecs)

$$3) \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

(r positive singular values)

where $f(\vec{v}_i) = \sigma_i \vec{u}_i$

II. Constructing the SVD

\Rightarrow a lot of inspiration from spectral theorem!

A^*A : - Hermitian

- eigenvalues are real

- orthogonally diagonalizable

- Corollary: for $B = A^*A$,
 $\lambda \geq 0$ (dis 3D)

\Rightarrow construct SVD from here?

Using $A^*A \in \mathbb{C}^{n \times n}$

① Find eigenvals λ_i of A^*A .

Order them s.t.:

$$\lambda_1 \geq \dots \geq \lambda_r > 0$$

$$\lambda_{r+1} = \dots = \lambda_n = 0$$

② Find orthonormal eigenvectors of A^*A
for $i = 1 \dots r$

$$\Rightarrow (A^* A) \vec{v}_i = \lambda_i \vec{v}_i$$

($i = 1 \dots r$)

\Rightarrow right singular vectors of A

③ Set $\sigma_i = \sqrt{\lambda_i}$ for $i = 1 \dots r$

④ Left singular vecs:

Note $A \vec{v}_i = \sigma_i \vec{u}_i$

$$\Rightarrow \vec{u}_i = \frac{A \vec{v}_i}{\sigma_i} \quad (i = 1 \dots r)$$

⑤ Complete the \perp bases for U and V

\hookrightarrow m vectors \hookrightarrow n vectors
by end

(Gram-Schmidt)

Note: $\vec{u}_i = \frac{A \vec{v}_i}{\sigma_i}$ are \perp

$$\begin{aligned}
 \vec{u}_i &= A \vec{v}_i \\
 \langle \vec{u}_i, \vec{u}_j \rangle &= \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \\
 \vec{u}_j^* \cdot \vec{u}_i &= \frac{\vec{v}_j^* (A^* A) \vec{v}_i}{\sigma_i \sigma_j} = \frac{\lambda_i}{\sigma_i \sigma_j} \vec{v}_j^* \cdot \vec{v}_i \\
 &= \frac{\sigma_i^2}{\sigma_i \sigma_j} \langle \vec{v}_i, \vec{v}_j \rangle \\
 &= \frac{\sigma_i}{\sigma_j} \langle \vec{v}_i, \vec{v}_j \rangle = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}
 \end{aligned}$$

Something interesting...

Essentially solving:

$$\begin{cases}
 A = U \Sigma V^* \\
 U^* U = I \\
 V^* V = I
 \end{cases}$$

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

$$\begin{aligned}
 A^*A &= V \Sigma^* U^* U \Sigma V^* \\
 &= V \hat{\Sigma}^2 V^*
 \end{aligned}
 \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

diagonalization!

$V \longrightarrow$ eigenvectors of A^*A
are right singular
vectors

$$\begin{aligned}
 AA^* &= U \Sigma V^* V \Sigma^* U^* \\
 &= U \hat{\Sigma}^2 U^* \quad \hat{\Sigma}^2 = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_m^2 \end{pmatrix}
 \end{aligned}$$

\Rightarrow eigenvectors of AA^* are
left singular vecs of A

Suggests procedure where we
calculate SVD from AA^*
(WS)

III) SVD relation to matrix subspaces
 \rightarrow

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

① Consider $\sigma_i = 0$ for $i > r$
 $\Rightarrow A \vec{v}_i = \vec{0}$

Right singular vcs corresponding
to vanishing singular values
Span $\text{Null}(A)$

② Consider $i = 1 \dots r, \sigma_i \neq 0$
 $A \vec{v}_i = \sigma_i \vec{u}_i$

Recall $\{\vec{v}_i\}$ span \mathbb{C}^N :

Then left singular vectors
corresponding to non zero σ_i
Span $\text{Col}(A)$

What about $\text{Null}(A^*), \text{Col}(A^*)$?

$$A^* = (U \Sigma V^*)^*$$

$$= V \Sigma^* U^*$$

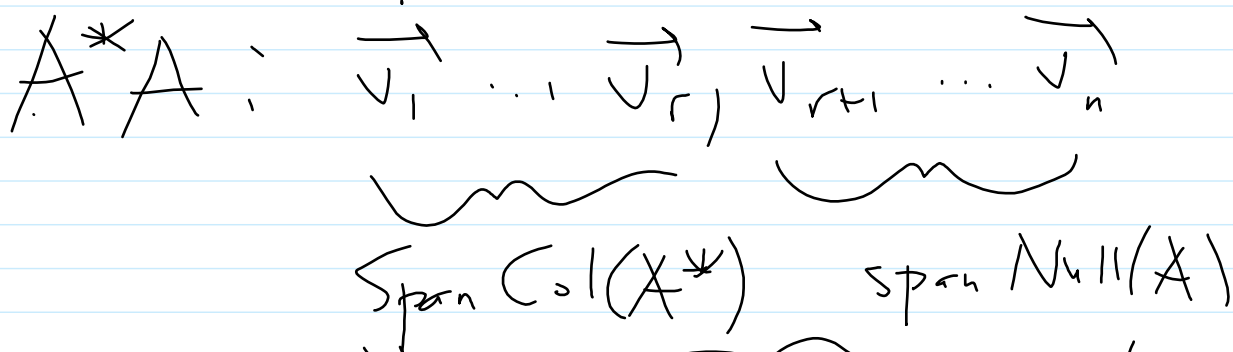
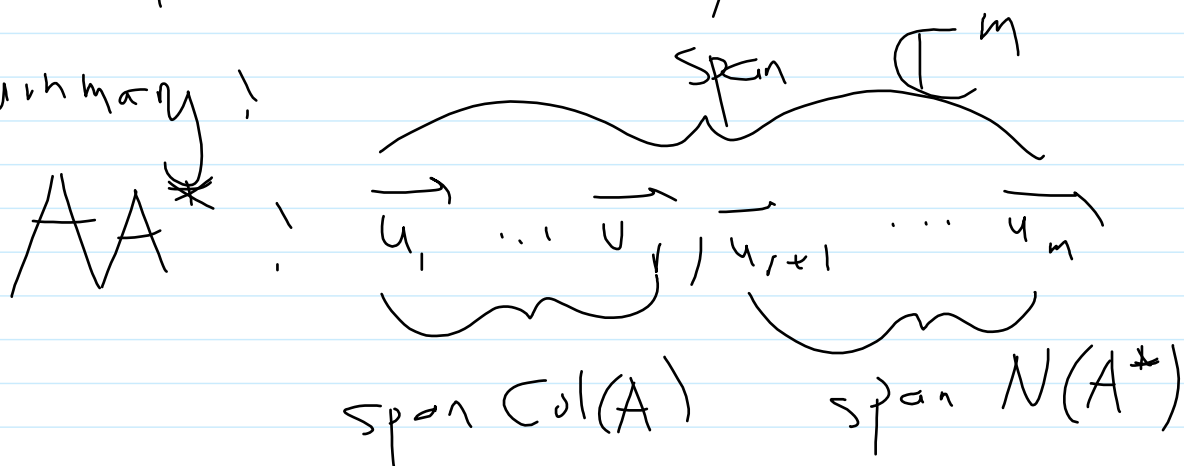
$$= V \Sigma^T U^* \Rightarrow A^* U = V \Sigma^T$$

$$A^* \vec{u}_i = \sigma_i \vec{v}_i$$

① left singular vectors of A
 corresponding to $\sigma_i = 0$
 span $\text{Null}(A^*)$

② Right singular vecs of A
 corresponding to $\sigma_i \neq 0$
 span $\text{Col}(A^*)$

Summary!



$$\underbrace{\text{Span}(\text{Col}(A^*)) \quad \text{Span}(\text{Null}(A))}_{\text{Span } \mathbb{C}^n}$$

1 SVD and Fundamental Subspaces

Define the matrix

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

$$A = \sum_{i=1}^r \sigma_i u_i v_i^*$$

a) Find the SVD of A (compact form is fine).

$$A^*A = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

$$\Rightarrow \lambda_1 = 18, \quad \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_2 = 0, \quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\sigma_1 = \sqrt{18} = 3\sqrt{2}$$

$$\sigma_2 = 0$$

$$\begin{aligned} \vec{u}_1 &= \frac{A\vec{v}_1}{\sigma_1} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 2 \\ -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix} \end{aligned}$$

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^* = 3\sqrt{2} \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

b) Find the rank of A .

$$\boxed{\text{rank } A = 1}$$

\hookrightarrow only 1 nonzero σ_i

c) Find a basis for the kernel (or nullspace) of A .

$$A v_i = \sigma_i u_i$$

$$\Rightarrow A \underline{\underline{v_2}} = 0 \underline{\underline{u_2}} = \underline{\underline{0}}$$

$$\boxed{\text{Null}(A) = \text{span} \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}}$$

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

$$A v_i = \sigma_i u_i$$

\Rightarrow nonzero σ_i

$$\Rightarrow \boxed{\text{Col}(A) = \text{span} \{ u_1 \} = \text{span} \left\{ \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} \right\}}$$

e) Repeat parts (a) - (d), but instead, create the SVD of $B = A^*$. What are the relationships between the answers for A and the answers for $B = A^*$?

$$B = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix}$$

$$(A = U\Sigma V^*)^* \Rightarrow B = A^* = V\Sigma^* U^*$$

$$B = 3\sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

$$\text{Rank } A^* = 1$$

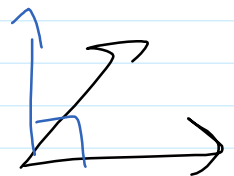
$$\text{Col}(A^*) = \text{span} \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\}$$

Nullspace of A^* ?

$$\lambda_2 = \lambda_3 = 0 \quad (\text{for } AA^*)$$

$$\begin{bmatrix} 2 & -4 & 4 \\ -4 & 8 & -8 \\ 4 & -8 & 8 \end{bmatrix} \vec{v} = \vec{0}$$

$$\text{Try } \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$



Gram-Schmidt!

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} / \sqrt{2}$$

$$\vec{z}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \\ -1/2 \end{bmatrix}$$

$$\text{Null}(A) = \text{span} \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{3\sqrt{2}} \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix} \right\}$$

2 Understanding the SVD

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

a) What are the dimensions of AA^* and A^*A .

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

$$AA^* : m \times m$$

$$A^*A : n \times n$$

b) Given that the $A = U\Sigma V^*$, find a symbolic expression for AA^* .

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

$$= U \Sigma \Sigma^* U^*$$

$$B = AA^* = U \hat{\Sigma}^2 U^*$$

$$\hat{\Sigma}^2 = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_m^2 \end{pmatrix}$$

c) Using the solution to the previous part explain how to find U and Σ from AA^* .

$$AA^* = U \hat{\Sigma}^2 U^* \quad \text{diagonalization!}$$

\Rightarrow find eigenvectors of AA^*
 $\{ \vec{u}_i \}$ and eigenvalues $\{ \lambda_i \}$

$$\Rightarrow \sigma_i = \sqrt{\lambda_i}$$

d) Now that we have found the singular values σ_i and the corresponding vectors \vec{u}_i in the matrix U , devise a way to find the corresponding vectors \vec{v}_i in matrix V .

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

$$A^* = (U \Sigma V^*)^* = V \Sigma^* U^*$$

$$A^* = (U \Sigma V^*)^* = V \Sigma^* U^*$$

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

$$\Rightarrow A^* \vec{u}_i = \sigma_i \vec{v}_i \quad (\sigma_i \in \mathbb{R})$$

e) Now we have a way to find the vectors \vec{v}_i in matrix V , verify that they are orthonormal.

$$\begin{aligned} \vec{v}_i &= \frac{A^* \vec{u}_i}{\sigma_i} \\ \vec{v}_j &= \frac{A^* \vec{u}_j}{\sigma_j} \end{aligned} \quad \left. \right\} \langle \vec{v}_i, \vec{v}_j \rangle = \frac{\vec{v}_i^* \vec{v}_j}{\sigma_i \sigma_j} = \frac{\vec{u}_i^* A A^* \vec{u}_j}{\sigma_i \sigma_j} = \frac{\vec{u}_i^* \vec{u}_j}{\sigma_i \sigma_j}$$

Note $\lambda_i = \sigma_i^2$ \rightarrow $\frac{\sigma_i}{\sigma_j} \langle \vec{u}_i, \vec{u}_j \rangle$

if $i=j$: $\langle \vec{v}_i, \vec{v}_i \rangle = \frac{\sigma_i}{\sigma_i} \langle \vec{u}_i, \vec{u}_i \rangle = 1$

if $i \neq j$: $\langle \vec{v}_i, \vec{v}_j \rangle = \frac{\sigma_i}{\sigma_j} \langle \vec{u}_i, \vec{u}_j \rangle = 0$

f) Now that we have found \vec{v}_i you may notice that we only have $m < n$ vectors of dimension n . This is not enough for a basis. How would you complete the m vectors to form an orthonormal basis?

Start: $\{ \vec{v}_1, \dots, \vec{v}_m \} \rightarrow$ need n to span

Start: $\{ \vec{v}_1 \dots \vec{v}_m \}$ \longrightarrow need n to span
 n -dim space!

Use G-S procedure!

- Specifically, need to keep $\vec{v}_1, \dots, \vec{v}_m$ for
SVD, but $\vec{v}_{m+1}, \dots, \vec{v}_n$ will all get
mapped to $\vec{0}$ anyway so they can be
arbitrary as long as $\{ \vec{v}_1, \dots, \vec{v}_m, \vec{v}_{m+1}, \dots, \vec{v}_n \}$
is an orthonormal set

- We also need to ensure that the
set spans the whole space; so
add the entire standard basis in
 $\{ \vec{e}_1, \dots, \vec{e}_n \}$

Thus, we perform G-S on:

$$\{ \vec{v}_1, \dots, \vec{v}_m, \vec{e}_1, \dots, \vec{e}_n \}$$

Any vector that is redundant
will become $\vec{0}$ after G-S

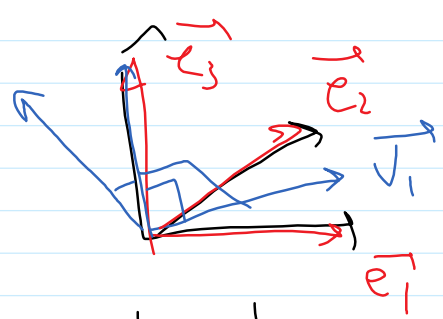
(redundant = lin. dep. on the
already orthonormalized
vectors)

EXAMPLE:

-

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$\begin{aligned} \vec{z}_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \vec{0} \end{aligned}$$



So our basis set becomes:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

remove this

In Summary:

- 1) Add std basis n so our spanning set becomes $\{\vec{v}_1, \dots, \vec{v}_n, \vec{e}_1, \dots, \vec{e}_k\}$
- 2) Orthonormalize w/ G-S, starting

with \vec{e}_i (keep $\vec{v}_1, \dots, \vec{v}_n$ untouched)

3) Remove the m zero vectors that result so you have a basis of n vectors to span an n -dim vector space!

g) 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.

① Pick the smaller of A^*A and AA^*

$$A \in \mathbb{R}^{1000 \times 2}$$

$$A^*A \in \mathbb{R}^{2 \times 2}$$

$$AA^* \in \mathbb{R}^{1000 \times 1000}$$

② Find the eigenvals and eigenvcs

a) if A^*A : find $A^*A \vec{v}_i = \lambda_i \vec{v}_i$

b) if AA^* : find $AA^* \vec{u}_i = \lambda_i \vec{u}_i$

③ $\sigma_i = \sqrt{\lambda_i}$

④ Complete U, V with $G-S$