MODULE 14: THE SVD

Linear Algebra 4: Orthogonality & Symmetric Matrices and the SVD

TOPIC 1: The Singular Value Decomposition

Singular Values

DEFINITION

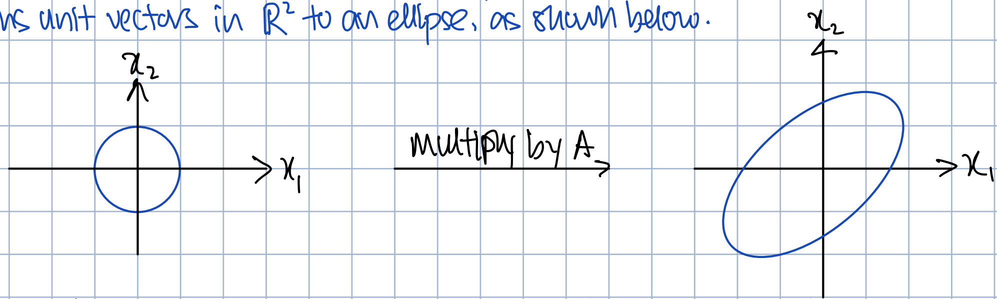
↳ The singular values of any $m \times n$ real matrix A are the square roots of the eigenvalues of A^TA.

Questions → How might this definition connect to some of the other topics we covered?
↳ How might singular values be viewed from a more geometric perspective?
↳ Can the eigenvalues of A^TA be negative?
↳ What problems can singular values be used to solve?

The linear transform whose standard matrix is

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

transforms unit vectors in $\mathbb{R}^2$ to an ellipse, as shown below.



What unit vector, $\vec{v}$, maximizes $\|A\vec{v}\|$? What is $\|A\vec{v}\|$ equal to?

Our goal is to maximize $\|A\vec{v}\|$ subject to $\|\vec{v}\| = 1$. In other words, $\max_{\|\vec{v}\|=1} \|A\vec{v}\|$.

A^TA is symmetric.
We can answer some of our questions with the eigenvalues of A^TA.

We need the eigenvalues of A^TA.

$$A^TA = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \lambda = 8, 2$$

Thus, $\max_{\|\vec{v}\|=1} \|A\vec{v}\|^2 = 8$.

So $\max_{\|\vec{v}\|=1} \|A\vec{v}\| = \sqrt{8}$. Thus, $\sqrt{8}$ is the maximum value that we were seeking.

Next: What is the $\vec{v}$ that corresponds to this maximum?

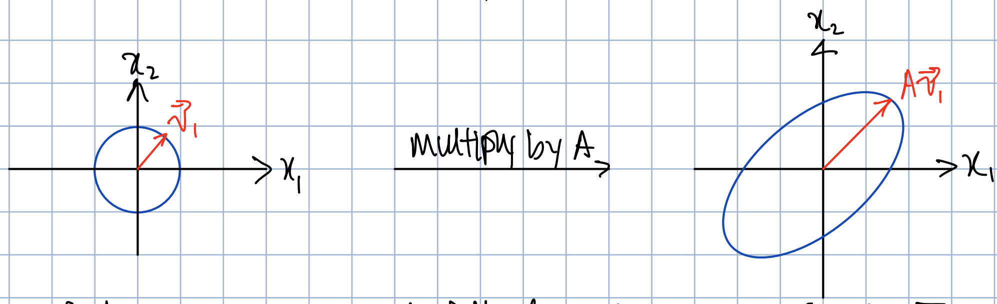
The unit eigenvector corresponding to the largest eigenvalue is the vector $\vec{v}_1$ that maximizes $\|A\vec{v}\|$.

$$A^TA - \lambda I = \begin{pmatrix} 0 & 0 \\ 0 & -6 \end{pmatrix} \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The unit vector that maximizes $\|A\vec{v}\|$ is $\vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Recall that A is the standard matrix for a transform that rotates and scales.

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

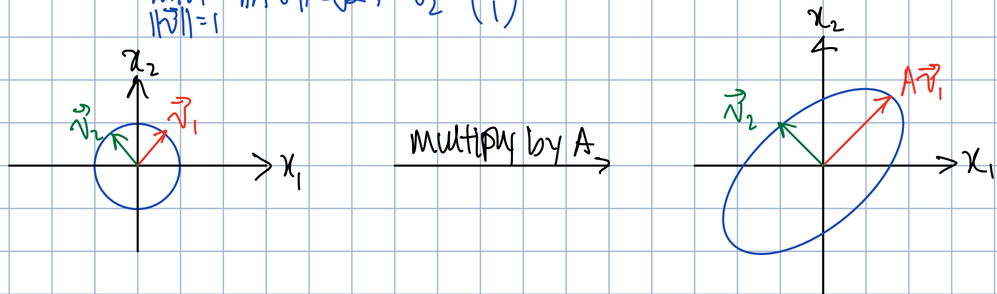


The length of $A\vec{v}_1$ is the square root of the largest eigenvalue of A^TA. $\sqrt{8}$.

The Smallest Eigenvalue

A similar process yields that the smallest eigenvalue of A^TA is

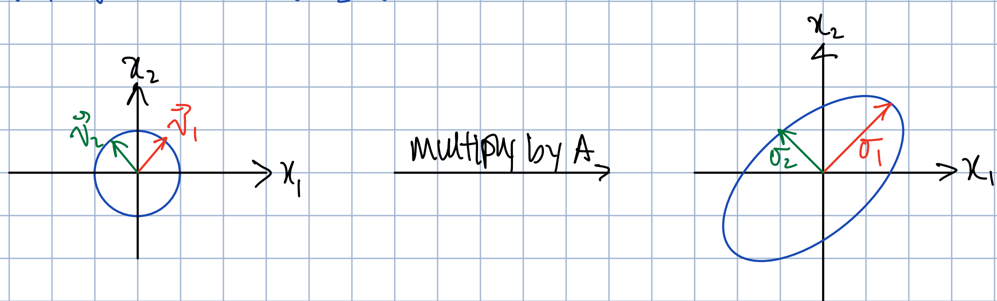
$$\min_{\|\vec{v}\|=1} \|A\vec{v}\| = \sqrt{2}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$



$\|A\vec{v}_2\|$ is the square root of the smallest eigenvalue of A^TA, which is $\sqrt{2}$.

Singular values are the Eigenvalues of A^TA.

Let $\sigma_1 = \sqrt{\lambda_1} = \sqrt{8}$, and $\sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$.



In other words, the length of $A\vec{v}_i$ is the square root of the eigenvalues of A^TA, or $\sqrt{\lambda_i}$.

Eigenvalues of A^TA are Nonnegative

Theorem:
The eigenvalues of A^TA are nonnegative.

Proof: Recall that $\vec{v}_j^T\vec{v}_j = \vec{v}_j \cdot \vec{v}_j = \|\vec{v}_j\|^2 = 1$ because $\vec{v}_j$ are unit eigenvectors of A^TA.

$$\|A\vec{v}_j\|^2 = (A\vec{v}_j)^T A\vec{v}_j = \vec{v}_j^T A^T A \vec{v}_j = \lambda_j \vec{v}_j^T \vec{v}_j = \lambda_j \geq 0.$$

Therefore,
* eigenvalues of A^TA must be real and nonnegative
* singular values of A (ie. square roots of eigenvalues) must also be real and nonnegative.

Note: Singular values are ordered.

Because the eigenvalues of A^TA are nonnegative, the singular values of A must also be real and nonnegative. They can be ordered from largest to smallest.

DEFINITION

↳ The singular values, σ_i, of any $m \times n$ real matrix A are the square roots of the eigenvalues of A^TA. Singular values are arranged in decreasing order.

$$\sigma_1 = \sqrt{\lambda_1} \geq \sigma_2 = \sqrt{\lambda_2} \geq \cdots \geq \sigma_n = \sqrt{\lambda_n}$$

This is a standard convention for singular values.

Singular Values are Lengths

It also follows from the previous slide that the singular values of A represent lengths of vectors in $\mathbb{R}^n$. That is, when showing that the eigenvalues of A^TA are nonnegative, we saw that $\|A\vec{v}_i\|^2 = \lambda_i$.

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

We make use of this relationship when constructing the SVD of a matrix.

Exercise:

Suppose we wish to identify a unit vector $\vec{x}$ for which $A\vec{x}$ has maximum length, where A is the matrix below.

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

The unit vectors that maximize $A\vec{x}$ will have the form given by $\vec{v}_1 = \pm \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

What is x_1 equal to? Assume $x_1 > 0$.

$$A^TA = \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 2 \\ 2 & 5 \end{pmatrix}.$$

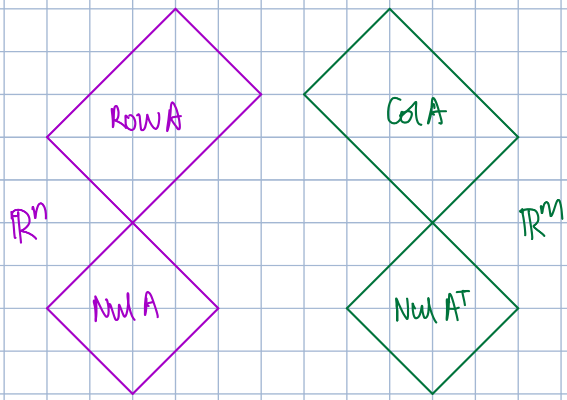
$$\begin{aligned} &\det(A^TA - \lambda I) \\ &= \det\begin{pmatrix} 8-\lambda & 2 \\ 2 & 5-\lambda \end{pmatrix} \\ &= (8-\lambda)(5-\lambda) - 4 = \lambda^2 - 13\lambda + 40 - 4 \\ &= \lambda^2 - 13\lambda + 36 = (\lambda-4)(\lambda-9). \end{aligned}$$

$\Rightarrow$ The eigenvalues of A^TA are $\lambda = 4, 9$. $\Rightarrow \sigma_1 = 3, \sigma_2 = 2$.

The corresponding unit eigenvector $\vec{x}$ to $\lambda = 9$ is $\vec{v}_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}$

Singular Vectors

Motivation: The Four Fundamental Subspaces



Theorem (The Four Subspaces):
For any $A \in \mathbb{R}^{m \times n}$, the orthogonal complement of Row A is Nul A, and the orthogonal complement of Col A is Nul A^T.

The eigenvectors of A^TA can be used to construct bases for these subspaces.

Orthogonal Bases for Nul A and Row A

Theorem:
Suppose $\vec{v}_i$ are the n orthogonal eigenvectors of A^TA, ordered so that their corresponding eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Suppose also that A has r nonzero singular values, $r \leq n$. Then the set of vectors

$$\{\vec{v}_{r+1}, \vec{v}_{r+2}, \dots, \vec{v}_n\}$$

is an orthogonal basis for Nul A, and the set

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$$

is an orthogonal basis for Row A, and rank $A = r$.

Proof for Orthogonal Basis for Nul A:

For a set of vectors to form an orthogonal basis for a subspace they must be in that space, span that space, be independent, and mutually orthogonal.

* Each $\vec{v}_i$ is an eigenvector, so none of them are the zero vector.
* $\vec{v}_i$ are orthogonal and span $\mathbb{R}^n$ (they are eigenvectors of a symmetric matrix, A^TA).
* Recall that the lengths of $A\vec{v}_i$ are the singular values of A:

$$\|A\vec{v}_i\| = \sigma_i.$$

* Then if $\|A\vec{v}_i\| = 0$ for $i > r$, then $\vec{v}_i \in$ Nul A for $i > r$.
* Then if $\|A\vec{v}_i\| \neq 0$ for $i \leq r$, then $\vec{v}_i$ cannot be in Nul A for $i \leq r$, they must be in $(\text{Nul } A)^\perp = $ Row A, because $\{\vec{v}_i\}$ is an orthonormal set.

Thus, our basis for Nul A is the set $\{\vec{v}_{r+1}, \vec{v}_{r+2}, \dots, \vec{v}_n\}$ and our basis for Row A is the set $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$.

We must also describe why rank $A = r$.
* There are r vectors in our basis for Row A.
* Recall that dim(Row A) = dim(Col A) = rank A.

Orthogonal Bases for Col A and Nul A^T

Theorem:
Suppose $\vec{v}_i$ are the n orthonormal eigenvectors of A^TA, ordered so that their corresponding eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Suppose also that A has r nonzero singular values. Then

$$\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$$

are an orthonormal basis for Col A.

Proof Outline: * Each $A\vec{v}_i$ is a vector in Col A.
* $A\vec{v}_i$ and $A\vec{v}_j$ are orthogonal:

$$(A\vec{v}_i)\cdot(A\vec{v}_j) = \vec{v}_i^{\,T}A^TA\vec{v}_j = \lambda_j \vec{v}_i \cdot \vec{v}_j = 0.$$

* For $i \leq r = \text{rank } A$, $A\vec{v}_i$ are orthogonal and nonzero. So they must also be independent and form an orthogonal basis for Col A.

Note: For $i > r$, $A\vec{v}_i = \vec{0}$ because $\vec{v}_i \in \text{Nul } A$ for $i > r$.

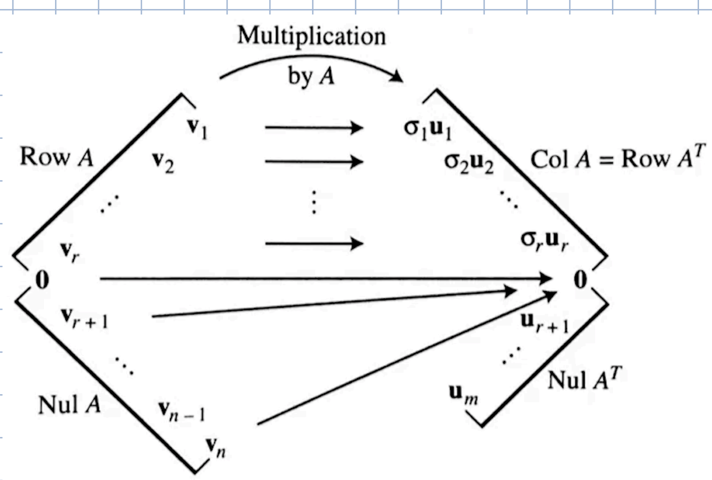
Summary: The Four Fundamental Spaces

Suppose $\vec{i}$ are orthonormal eigenvectors for A^TA, and

$$\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i \quad \text{for } i \leq r = \text{rank } A,\ \sigma_i = \|A\vec{v}_i\|.$$

Then we have the following orthogonal bases for any $m \times n$ real matrix A.
* $\vec{v}_1, \ldots, \vec{v}_r$ is an orthonormal basis for Row A.
* $\vec{v}_{r+1}, \ldots, \vec{v}_n$ is an orthonormal basis for Nul A.
* $\vec{u}_1, \ldots, \vec{u}_r$ is an orthonormal basis for Col A.

If we need a basis for $(\text{Col } A)^{\perp}$, what might we do?
One approach: identify any $m-r$ independent nonzero vectors in $(\text{Col } A)^{\perp}$ and then use Gram-Schmidt to orthogonalize them (more on this later).



The Left and Right Singular Vectors

DEFINITION

↳ The vectors $\{\vec{u}_i\}$ for $i \leq m$ are the left singular vectors of A.

↳ The vectors $\{\vec{v}_i\}$ for $i \leq n$ are the right singular vectors of A.

The reason we refer to these vectors as left and right singular vectors is connected to the singular value decomposition (more on this later).

Example:

Suppose A is a 12x4 real matrix and has 3 nonzero singular values. Indicate whether the following statements are true or false.

① A basis for Col A is given by the vectors $A\vec{v}_1$, $A\vec{v}_2$ and $A\vec{v}_3$.

TRUE. A has rank 3, the three left singular vectors $A\vec{v}_1$, $A\vec{v}_2$ and $A\vec{v}_3$ are independent and are in Col A, so they form a basis for Col A.

② A basis for Nul A is given by $\vec{v}_1, \vec{v}_2, \vec{v}_3$.

FALSE. A has rank 3, so the three right singular vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ form a basis for Row A. The vector $\vec{v}_4$ forms a basis for Nul A.

Exercise:

Suppose $A = \begin{pmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix}$.

Compute the left and right singular vectors of A.

$$A^TA = \begin{pmatrix} 4 & 2 & 0 \\ -2 & -1 & 0 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 20 & -10 \\ -10 & 5 \end{pmatrix}.$$

$\det(A^TA - \lambda I)$

$= \det\begin{pmatrix} 20-\lambda & -10 \\ -10 & 5-\lambda \end{pmatrix} = (20-\lambda)(5-\lambda) - 100$

$= 100 - 25\lambda + \lambda^2 - 100 = \lambda^2 - 25\lambda = \lambda(\lambda - 25)$

$\Rightarrow \lambda_1 = 25, \lambda_2 = 0. \Rightarrow$ singular values: $\sigma_1 = 5$, $\sigma_2 = 0$.

A has 1 nonzero singular value. $\therefore \underline{\underline{r=1}}$

If the unit right singular vector is $\vec{v}_1 = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$, and $k_2 > 0$, what is k_1 equal to?

For $\lambda_1 = 25$, $A^TA - 25I = \begin{pmatrix} -5 & -10 \\ -10 & -20 \end{pmatrix}$.

$x_1 + 2x_2 = 0$. Let $x_1 = -2$, $x_2 = 1$.

A corresponding unit eigenvector is

$$\vec{v}_1 = \begin{pmatrix} -2/\sqrt{(-2)^2+1^2} \\ 1/\sqrt{(-2)^2+1^2} \end{pmatrix} = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}. \quad k_1 = -\frac{2}{\sqrt{5}}.$$

If the unit left singular vector is $\vec{u}_1 = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$, what is h_1 equal to?

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{5}\begin{pmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -10/\sqrt{5} \\ -5/\sqrt{5} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}.$$

The SVD

Theorem: Singular Value Decomposition
Suppose A is an $m \times n$ matrix with singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ and $m \geq n$. Then A has the decomposition $A = U \Sigma V^T$ where

$$\Sigma = \begin{pmatrix} D \\ \vec{0}_{m-n,m} \end{pmatrix}, \quad D = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix}$$

U is an $m \times m$ orthogonal matrix, and V is a $n \times n$ orthogonal matrix. If $m < n$, then $\Sigma = (D \;\; \vec{0}_{m,n-m})$ with everything else is the same.

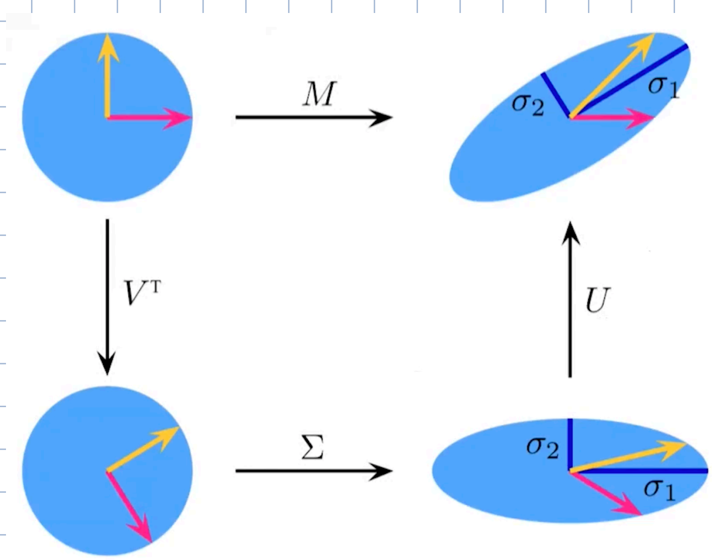
Proof that $A = U\Sigma V^T$: (similar to proof for diagonalization)

We construct $V = (\vec{v}_1, \vec{v}_2 \cdots \vec{v}_n)$ and the set
$\sigma_i \vec{u}_i = A\vec{v}_i$, $\sigma_i = \|A\vec{v}_i\|$.

Thus:
$$\begin{aligned} AV = A(\vec{v}_1, \vec{v}_2 \cdots \vec{v}_n) &= (A\vec{v}_1 \;\; A\vec{v}_2 \cdots A\vec{v}_n) \\ &= (\sigma_1\vec{v}_1 \;\; \sigma_2\vec{v}_2 \cdots \sigma_n\vec{v}_n) \\ &= (\vec{u}_1 \;\; \vec{u}_2 \cdots \vec{u}_n)\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} = U\Sigma. \end{aligned}$$

Thus, $AV = U\Sigma$, or $A = U\Sigma V^T$.

The SVD and Linear Transforms



A Procedure for Constructing the SVD of A

Suppose A is $m \times n$ and has rank r.

① Compute the squared singular values of A^TA, σ_i^2, and construct Σ.
② Compute the unit singular vectors of A^TA, $\vec{v}_i$, use them to form V.
③ Compute an orthonormal basis for Col A using

$$\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i, \quad i = 1, 2, \ldots, r$$

If necessary, extend the set $\{\vec{u}_i\}$ to form an orthonormal basis for $\mathbb{R}^m$ and use the basis to form U.

Example: Construct a singular value decomposition for $A = \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Singular values: $A^TA = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \Rightarrow \lambda_1 = 9, \ \lambda_2 = 4.$

The positive square roots of the eigenvalues are the singular values.

$$\sigma_1 = 3, \ \sigma_2 = 2.$$

REMEMBER: σ_1 is the largest singular value.

Using the singular values we can construct Σ.

$$\sigma_1 = 3, \ \sigma_2 = 2 \Rightarrow \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Next we construct the singular right vectors $\{\vec{v}_i\}$ and form V.

$$A^TA - \lambda_1 I = \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow \vec{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A^TA - \lambda_2 I = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} \Rightarrow \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad \Rightarrow V = (\vec{v}_1 \ \ \vec{v}_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Next, we construct left-singular vectors $\{\vec{u}_i\}$ using $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ for $i = 1, 2, \ldots, r$. Each $\vec{u}_i$ will be a unit vector in $\mathbb{R}^4$.

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3}\begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2 = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

To construct the SVD of A, we must construct the last 2 columns of U.

* In this example, A has rank $r=2$ and U will be a 4×4 orthogonal matrix.
* Because the columns of U must be orthonormal, and $\vec{u}_1$ and $\vec{u}_2$ were standard vectors, by inspection we can set the last two columns to be:

$$\vec{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{u}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that $\vec{u}_3$ and $\vec{u}_4$ are unit vectors, and that $\{\vec{u}_i\}$ are orthonormal.

* We could have chosen other vectors for $\vec{u}_3$ and $\vec{u}_4$.

We have the SVD of A:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Exercise: Construct an SVD for $A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix}$.

$$A^TA = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}.$$

$$\det(A^TA - \lambda I) = (5-\lambda)(8-\lambda) - 4$$
$$= 40 - 13\lambda + \lambda^2 - 4$$
$$= \lambda^2 - 13\lambda + 36 = (\lambda - 4)(\lambda - 9)$$
$$\Rightarrow \lambda_1 = 9, \ \lambda_2 = 4.$$
$$\Rightarrow \sigma_1 = 3, \ \sigma_2 = 2. \Rightarrow \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$$

Constructing V,

$\lambda_1 = 9, \ A^TA - \lambda_1 I = \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \quad 2x_1 - x_2 = 0 \Rightarrow x_2 = 2x_1. \ \vec{v}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$

$\lambda_2 = 4, \ A^TA - \lambda_2 I = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad x_1 + 2x_2 = 0. \ \vec{v}_2 = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}. \quad \Rightarrow V = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$

Constructing U,

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{3\sqrt{5}} \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix}$$

$$\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2 = \frac{1}{2\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \frac{1}{2\sqrt{5}} \begin{pmatrix} 0 \\ -4 \\ 2 \end{pmatrix}$$

$\vec{u}_1 \cdot \vec{u}_3 = 0. \quad 5u_{3\alpha} + 2u_{3\beta} + 4u_{3\gamma} = 0.$

$\vec{u}_2 \cdot \vec{u}_3 = 0 \quad -4u_{3\beta} + 2u_{3\gamma} = 0 \Rightarrow u_{3\gamma} = 2u_{3\beta}$

The SVD of a 3X2 Matrix with Rank 1

Example: Construct the singular value decomposition of $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$.

Singular values:

$$A^TA = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}. \quad \det(A^TA - \lambda I) = (9-\lambda)^2 - 81 = \lambda^2 - 18\lambda = \lambda(\lambda - 18)$$

$$\Rightarrow \lambda_1 = 18, \lambda_2 = 0 \quad \Rightarrow \sigma_1 = 3\sqrt{2}, \sigma_2 = 0.$$

Don't forget that

* singular matrices have eigenvalue 0 and the trace of a matrix is the sum of its eigenvalues.
* The positive square roots of the eigenvalues are the single values.
* σ_i is the largest singular value.

Using the singular values we can construct Σ. $\quad \Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

Next we construct the right-singular vectors $\{\vec{v}_i\}$ and form V.

$$\left.\begin{aligned} A^TA - \lambda_1 I &= \begin{pmatrix} -9 & -9 \\ -9 & -9 \end{pmatrix} \Rightarrow \vec{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ A^TA - \lambda_2 I &= \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix} \Rightarrow \vec{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}\right\} \Rightarrow V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Next we construct left-singular vectors $\{\vec{u}_i\}$. The rank of A is $r=1$, so we may use

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

for $i=1$. Vector $\vec{u}_1$ will be a unit vector in $\mathbb{R}^3$.

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

How can we construct the remaining left-singular vectors to construct U?

* In this example, A has rank $r=1$, and U will be a 3x3 orthogonal matrix.
* By inspection, two vectors orthogonal to $\vec{u}_1$ are

$$\vec{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \vec{x}_3 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad \text{recall: } \vec{u}_1 = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

$\vec{x}_3 \cdot \vec{u}_2 = -2(2) + 0(1) + 1(0) = -4.$

$\vec{u}_2 \cdot \vec{u}_2 = 2^2 + 1^2 + 0^2 = 5$

* Because $\vec{u}_1 \in \text{Col } A$, these two vectors are in $(\text{Col } A)^\perp$.
* But U is an orthogonal matrix, so how might we create an orthogonal basis for $(\text{Col } A)^\perp$?
  ↳ Gram-Schmidt.

$$\vec{u}_2 = \vec{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}. \quad \vec{u}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{-4}{5}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{5}\left[\begin{pmatrix} -10 \\ 0 \\ 5 \end{pmatrix} + \begin{pmatrix} 8 \\ 4 \\ 0 \end{pmatrix}\right] = \frac{1}{5}\begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix}$$

Normalizing these vectors yields the remaining left-singular vectors.

$$\vec{u}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

Thus, $A = U\Sigma V^T$, where

$$U = \begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Check:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix}\begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$= \frac{1}{3\sqrt{2}}\begin{pmatrix} 1 & 6/\sqrt{5} & -2/\sqrt{5} \\ -2 & 3/\sqrt{5} & 4/\sqrt{5} \\ 2 & 0 & 5/\sqrt{5} \end{pmatrix}\begin{pmatrix} 3\sqrt{2} & -3\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \frac{1}{\sqrt{5}}\begin{pmatrix} \sqrt{5} & 6 & -2 \\ -2\sqrt{5} & 3 & 4 \\ 2\sqrt{5} & 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \sqrt{5} & -\sqrt{5} \\ -2\sqrt{5} & 2\sqrt{5} \\ 2\sqrt{5} & -2\sqrt{5} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \text{ (QED).}$$

Exercise: Suppose $A = \begin{pmatrix} 8 & 0 \\ 0 & 0 \\ 6 & 0 \end{pmatrix}$. Construct an SVD for A.

$$A^TA = \begin{pmatrix} 100 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \det(A^TA - \lambda I) = (100-\lambda)(-\lambda) = \lambda(\lambda - 100).$$

$$\Rightarrow \lambda_1 = 100, \lambda_2 = 0. \Rightarrow \sigma_1 = 10, \sigma_2 = 0.$$

$$\Rightarrow \underline{\Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}.$$

$$\left.\begin{aligned} A^TA - \lambda_1 I &= \begin{pmatrix} 0 & 0 \\ 0 & -100 \end{pmatrix}. \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \\ A^TA - \lambda_2 I &= \begin{pmatrix} 100 & 0 \\ 0 & 0 \end{pmatrix}. \Rightarrow \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}\right\} \Rightarrow \underline{V = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}$$

$$\vec{u}_1 = \frac{1}{\sigma_1}A\vec{v}_1 = \frac{1}{10}\begin{pmatrix} 8 & 0 \\ 0 & 0 \\ 6 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 8 \\ 0 \\ 6 \end{pmatrix}.$$

$\vec{x}_3 \cdot \vec{u}_2 = 0.$

$\vec{u}_2 \cdot \vec{u}_2 = \frac{1}{10}(36 + 64) = 10.$

Two vectors orthogonal to $\vec{u}_1$ are $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\frac{1}{10}\begin{pmatrix} 6 \\ 0 \\ -8 \end{pmatrix}$.

Let $\vec{u}_2 = \frac{1}{10}\begin{pmatrix} 6 \\ 0 \\ -8 \end{pmatrix}$ and $\vec{x}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Using Gram-Schmidt,

$$\vec{u}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{0}{10} \times \frac{1}{10}\begin{pmatrix} 6 \\ 0 \\ -8 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \Rightarrow U = \frac{1}{10}\begin{pmatrix} 8 & 6 & 0 \\ 0 & 0 & 10 \\ 6 & -8 & 0 \end{pmatrix}$$

$$\underline{U = \frac{1}{5}\begin{pmatrix} 4 & 3 & 0 \\ 0 & 0 & 5 \\ 3 & -4 & 0 \end{pmatrix}}.$$

$$\therefore A = U\Sigma V^T = \frac{1}{5}\begin{pmatrix} 4 & 3 & 0 \\ 0 & 0 & 5 \\ 3 & -4 & 0 \end{pmatrix}\begin{pmatrix} 10 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

TOPIC 2: Applications of the SVD

The SVD has been applied to many modern applications in CS, engineering, and mathematics.
→ estimating the rank and condition number of a matrix
→ constructing bases for the four fundamental spaces
→ computing the pseudoinverse of a matrix
→ linear least squares problems
→ machine learning and data mining
→ facial recognition
→ principle component analysis
→ image compression

The Condition Number of a Matrix

If A is an inverse $n \times n$ matrix, the ratio $\frac{\sigma_1}{\sigma_n}$ is the condition number of A.

Example: Suppose $A = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$. We found that $\sigma_1 = \sqrt{8}$, $\sigma_2 = \sqrt{2}$.

Therefore, the condition number of A is $\frac{\sigma_1}{\sigma_n} = \frac{\sqrt{8}}{\sqrt{2}} = 2$.

Notes on the Condition Number:
* In some applications of linear algebra, entries of A and contain errors. The condition number of a matrix describes the sensitivity that any approach to determining solutions to $A\vec{x} = \vec{b}$ might have errors in A.
* The larger the condition number, the more sensitive the system is to errors.
* We could define the condition number for a rectangular matrix, but that would go beyond the scope of this course.

Exercise: What is the condition number of the matrix $A = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$?

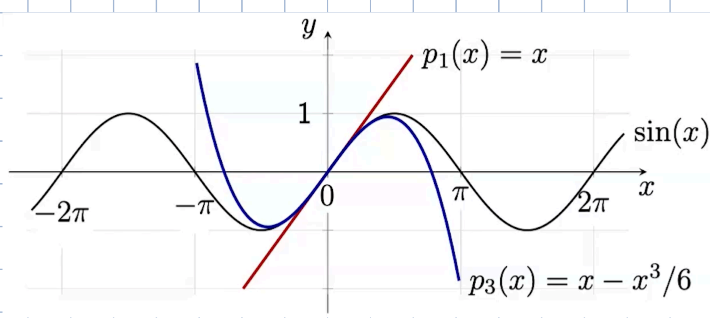
$A^TA = \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix}$. $\det(A^TA - \lambda I) = (4-\lambda)(16-\lambda)$. $\Rightarrow \lambda_1 = 16, \lambda_2 = 4$.
$\Rightarrow \sigma_1 = 4, \sigma_2 = 2$.

$\therefore$ condition number of $A = \frac{4}{2} = 2$.

The SVD and The Spectral Decomposition of a Matrix

Motivation: Approximation

Recall that from calculus, Taylor expansions and Taylor polynomials, can be used to approximate functions near a point.



Can we use expansions to approximate matrices?

Recall: Spectral Decomposition of a Symmetric Matrix

Suppose A can be orthogonally diagonalized as

$$A = PDP^T = (\vec{u}_1 \ \cdots \ \vec{u}_n)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}\begin{pmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{pmatrix}$$

Then A has the decomposition

$$A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \ldots + \lambda_n \vec{u}_n \vec{u}_n^T = \sum_{i=1}^{n} \lambda_n \vec{u}_n \vec{u}_n^T$$

Can we give a more general result using the SVD? YES! :)

The SVD can also be used to construct the spectral decomposition for any matrix with rank r:

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

vectors $\vec{u}_i$, $\vec{v}_i$ are the ith columns of U and V respectively.

* The proof similar to the proof we used for the symmetric case.
* Each term in this sum is a rank 1 matrix.
* In applications of linear algebra, σ_i can become sufficiently small, allowing us to approximate A with a small number of rank 1 matrices.
* For the case when $A = A^T$, we obtain the same spectral decomposition obtained using the orthogonal diagonalization of A, or $A = PDP^T$.

Example: Suppose A has the following SVD.

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The spectral decomposition of A is as follows:

$$A = \sum_{s=1}^{r} \sigma_s \vec{u}_s \vec{v}_s^T = 3\begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}(0 \quad 1) + 2\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}(1 \quad 0) = 3\begin{pmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 2\begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Exercise: Suppose $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

A has the SVD:

$$A = U\Sigma V^T, \quad U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

The spectral decomposition has the form: $A = A_1 + A_2 = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T$.

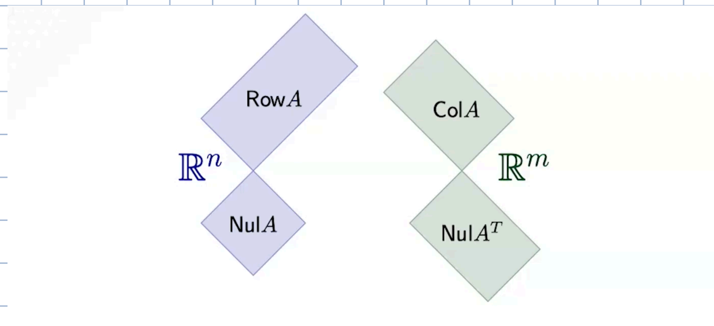
Find A_1.

$$\sigma_1 = \sqrt{2}. \quad \vec{u}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \vec{v}_1^T = (1 \quad 0 \quad 1). \quad A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The SVD and the Four Fundamental Spaces

Suppose $\vec{v}_i$ are orthonormal eigenvectors for A^TA, and

$\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ for $i \leq r = \text{rank } A$, $\sigma_i = \|A\vec{v}_i\|$.



The SVD of A can be used construct bases for these subspaces.

Then we have the following bases for any $m \times n$ real matrix A.

* $\vec{v}_1, \ldots, \vec{v}_r$ is an orthonormal basis for Row A.
* $\vec{v}_{r+1}, \ldots, \vec{v}_n$ is an orthonormal basis for Nul A.
* $\vec{u}_1, \ldots, \vec{u}_r$ is an orthonormal basis for Col A.

We can now also see that because U is an orthogonal matrix, $\vec{u}_{r+1}, \ldots, \vec{u}_m$ give an orthonormal basis for Nul A^T.

Example:
Given the SVD of A, determine rank(A), and bases for Nul A and for Col A.

$$A = U\Sigma V^T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \sqrt{0.8} & 0 & -\sqrt{0.2} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{0.8} & 0 & \sqrt{0.2} \end{pmatrix}$$

Solution $\Rightarrow$ there are exactly 3 nonsingular values $\Rightarrow$ rank $A = 3$
$\Rightarrow$ the last 2 columns of V (or rows of V^T) are a basis for Nul A
$\Rightarrow$ first 3 columns of U are a basis for Col A.

Exercise: Suppose A has the SVD factorization $A = U\Sigma V^T$, where

$$U = (\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3 \ \vec{u}_4), \ \Sigma = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \ V = (\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3 \ \vec{v}_4 \ \vec{v}_5)$$

What set gives a basis for Nul A? $\{\vec{v}_4, \vec{v}_5\}$
3 nonsingular values $\Rightarrow$ Rank 3.
$\Rightarrow$ last 2 columns of V give a basis for Nul A.
$\because \sigma_4 = \sigma_5 = 0 \Rightarrow \|A\vec{v}_4\| = \|A\vec{v}_5\| = 0$.

What set gives a basis for Row A? $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$
$(\text{Nul } A)^\perp =$ Row A. Last 2 columns of V give a basis for $(\text{Row } A)^\perp =$ Nul A

What set gives a basis for Col A? $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$
First r columns of U are independent and in Col A, so they form a basis for Col A.

What set gives a basis for Nul A^T? $\{\vec{u}_4\}$
The remaining column of U gives a basis for $(\text{Col } A)^\perp = \text{Nul } A^T$.