# Math 221: LINEAR ALGEBRA

## Chapter 8. Orthogonality

## §8-6. Singular Value Decomposition

Le Chen[1]

Emory University, 2021 Spring

(last updated on 01/12/2023)

Creative Commons License
(CC BY-NC-SA)

[1] Slides are adapted from those by Karen Seyffarth from University of Calgary.

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

# Linear Algebra with Applications

## Lecture Notes

**Current Lecture Notes Revision: Version 2018 — Revision B**

These lecture notes were originally developed by Karen Seyffarth of the University of Calgary. Edits, additions, and revisions have been made to these notes by the editorial team at Lyryx Learning to accompany their text Linear Algebra with Applications based on W. K. Nicholson's original text.

In addition we recognize the following contributors. All new content contributed is released under the same license as noted below.

- **Ilijas Farah, York University**

## BE A CHAMPION OF OER!

Contribute suggestions for improvements, new content, or errata:
A new topic
A new example or problem
A new or better proof to an existing theorem
Any other suggestions to improve the material

Contact Lyryx at info@lyryx.com with your ideas.

**License**

Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)
This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

# Singular Value Decomposition

## Definition

Let $A$ be an $m \times n$ matrix. The singular values of $A$ are the square roots of the nonzero eigenvalues of $A^T A$. Singular Value Decomposition (SVD) can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.
Given an $m \times n$ matrix $A$, we will see how to express $A$ as a product

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

where

- $U$ is an $m \times m$ orthogonal matrix whose columns are eigenvectors of $AA^T$.
- $V$ is an $n \times n$ orthogonal matrix whose columns are eigenvectors of $A^T A$.
- $\Sigma$ is an $m \times n$ matrix whose only nonzero values lie on its main diagonal, and are the square roots of the eigenvalues of both $AA^T$ and $A^T A$.

## Theorem

If $A$ is an $m \times n$ matrix, then $A^T A$ and $AA^T$ have the same nonzero eigenvalues.

## Proof.

Suppose $A$ is an $m \times n$ matrix, and suppose that $\lambda$ is a nonzero eigenvalue of $A^T A$. Then there exists a nonzero vector $\vec{x} \in \mathbb{R}^n$ such that

$$(A^T A)\vec{x} = \lambda\vec{x}. \tag{1}$$

Multiplying both sides of this equation by $A$:

$$\begin{aligned} A(A^T A)\vec{x} &= A\lambda\vec{x} \\ (AA^T)(A\vec{x}) &= \lambda(A\vec{x}). \end{aligned}$$

Since $\lambda \neq 0$ and $\vec{x} \neq \vec{0}_n$, $\lambda\vec{x} \neq \vec{0}_n$, and thus by equation (1), $(A^T A)\vec{x} \neq \vec{0}_n$; thus $A^T(A\vec{x}) \neq \vec{0}_n$, implying that $A\vec{x} \neq \vec{0}_m$.

Therefore $A\vec{x}$ is an eigenvector of $AA^T$ corresponding to eigenvalue $\lambda$. An analogous argument can be used to show that every nonzero eigenvalue of $AA^T$ is an eigenvalue of $A^T A$, thus completing the proof. ■

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

# Examples

## Example

Let $A = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}$. Then

$$AA^T = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 5 \\ 5 & 11 \end{bmatrix}.$$

$$A^TA = \begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 2 & 6 \\ 2 & 2 & -2 \\ 6 & -2 & 10 \end{bmatrix}.$$

## Example (continued)

Since $AA^T$ is $2 \times 2$ while $A^TA$ is $3 \times 3$, and $AA^T$ and $A^TA$ have the same nonzero eigenvalues, compute $c_{AA^T}(x)$ (because it's easier to compute than $c_{A^TA}(x)$).

$$\begin{aligned} c_{AA^T}(x) &= \det(xI - AA^T) = \begin{vmatrix} x-11 & -5 \\ -5 & x-11 \end{vmatrix} \\ &= (x-11)^2 - 25 \\ &= x^2 - 22x + 121 - 25 \\ &= x^2 - 22x + 96 \\ &= (x-16)(x-6). \end{aligned}$$

Therefore, the eigenvalues of $AA^T$ are $\lambda_1 = 16$ and $\lambda_2 = 6$.

## Example (continued)

The eigenvalues of $A^T A$ are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in nonincreasing order (i.e., from largest to smallest).

To find the matrix V, find eigenvectors for $A^T A$. Since the eigenvalues of $AA^T$ are distinct, the corresponding eigenvectors are orthogonal, and we need only normalize them.

$\lambda_1 = 16$: solve $(16I - A^T A)\vec{y}_1 = \vec{0}$.

$$\left[\begin{array}{rrr|r} 6 & -2 & -6 & 0 \\ -2 & 14 & 2 & 0 \\ -6 & 2 & 6 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } \vec{y}_1 = \begin{bmatrix} t \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, t \in \mathbb{R}.$$

$\lambda_2 = 6$: solve $(6I - A^T A)\vec{y}_2 = \vec{0}$.

$$\left[\begin{array}{rrr|r} -4 & -2 & -6 & 0 \\ -2 & 4 & 2 & 0 \\ -6 & 2 & -4 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } \vec{y}_2 = \begin{bmatrix} -s \\ -s \\ s \end{bmatrix} = s \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, s \in \mathbb{R}.$$

## Example (continued)

$\lambda_3 = 0$: solve $(-A^T A)\vec{y}_3 = \vec{0}$.

$$\left[\begin{array}{rrr|r} -10 & -2 & -6 & 0 \\ -2 & -2 & 2 & 0 \\ -6 & 2 & -10 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } \vec{y}_3 = \begin{bmatrix} -r \\ 2r \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

Let

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Then

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & -\sqrt{2} & -1 \\ 0 & -\sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & 1 \end{bmatrix}.$$

Also,

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix},$$

and we use A, $V^T$, and $\Sigma$ to find U.

## Example (continued)

Since V is orthogonal and $A = U\Sigma V^T$, it follows that $AV = U\Sigma$. Let $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$, and let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}$, where $\vec{u}_1$ and $\vec{u}_2$ are the two columns of U. Then we have

$$\begin{aligned} A\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}\Sigma \\ \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & A\vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \sigma_1\vec{u}_1 + 0\vec{u}_2 & 0\vec{u}_1 + \sigma_2\vec{u}_2 & 0\vec{u}_1 + 0\vec{u}_2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 & \sigma_2\vec{u}_2 & \vec{0} \end{bmatrix} \end{aligned}$$

which implies that $A\vec{v}_1 = \sigma_1\vec{u}_1 = 4\vec{u}_1$ and $A\vec{v}_2 = \sigma_2\vec{u}_2 = \sqrt{6}\vec{u}_2$. Thus,

$$\vec{u}_1 = \frac{1}{4}A\vec{v}_1 = \frac{1}{4}\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{4\sqrt{2}}\begin{bmatrix} 4 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and

$$\vec{u}_2 = \frac{1}{\sqrt{6}}A\vec{v}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}\frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{3\sqrt{2}}\begin{bmatrix} 3 \\ -3 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

## Example (continued)

Therefore,

$$\mathrm{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and

$$\begin{aligned} \mathrm{A} &= \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \\ &= \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \left( \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 0 & \sqrt{3} \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} \\ -1 & 2 & 1 \end{bmatrix} \right). \end{aligned}$$

## Problem

Find an SVD for $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$.

## Solution

Since A is $3 \times 1$, $A^T A$ is a $1 \times 1$ matrix whose eigenvalues are easier to find than the eigenvalues of the $3 \times 3$ matrix $AA^T$.

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

Thus $A^T A$ has eigenvalue $\lambda_1 = 9$, and the eigenvalues of $AA^T$ are $\lambda_1 = 9$, $\lambda_2 = 0$, and $\lambda_3 = 0$. Furthermore, A has only one singular value, $\sigma_1 = 3$.

To find the matrix V, find an eigenvector for $A^T A$ and normalize it. In this case, finding a unit eigenvector is trivial: $\vec{v}_1 = \begin{bmatrix} 1 \end{bmatrix}$, and

$$V = \begin{bmatrix} 1 \end{bmatrix}.$$

## Solution (continued)

Also, $\Sigma = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$, and we use A, $V^T$, and $\Sigma$ to find U.

Now $AV = U\Sigma$, with $V = \begin{bmatrix} \vec{v}_1 \end{bmatrix}$, and $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$, where $\vec{u}_1$, $\vec{u}_2$, and $\vec{u}_3$ are the columns of U. Thus

$$\begin{aligned} A \begin{bmatrix} \vec{v}_1 \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} \Sigma \\ \begin{bmatrix} A\vec{v}_1 \end{bmatrix} &= \begin{bmatrix} \sigma_1 \vec{u}_1 + 0\vec{u}_2 + 0\vec{u}_3 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \vec{u}_1 \end{bmatrix} \end{aligned}$$

This gives us $A\vec{v}_1 = \sigma_1 \vec{u}_1 = 3\vec{u}_1$, so

$$\vec{u}_1 = \frac{1}{3} A\vec{v}_1 = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

## Solution (continued)

The vectors $\vec{u}_2$ and $\vec{u}_3$ are eigenvectors of $AA^T$ corresponding to the eigenvalue $\lambda_2 = \lambda_3 = 0$. Instead of solving the system $(0I - AA^T)\vec{x} = \vec{0}$ and then using the Gram-Schmidt orthogonalization algorithm on the resulting set of two basic eigenvectors, the following approach may be used.

Find vectors $\vec{u}_2$ and $\vec{u}_3$ by first extending $\{\vec{u}_1\}$ to a basis of $\mathbb{R}^3$, then using the Gram-Schmidt algorithm to orthogonalize the basis, and finally normalizing the vectors.

Starting with $\{3\vec{u}_1\}$ instead of $\{\vec{u}_1\}$ makes the arithmetic a bit easier. It is easy to verify that

$$\left\{ \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is a basis of $\mathbb{R}^3$. Set

$$\vec{f}_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

and apply the Gram-Schmidt orthogonalization algorithm to $\{\vec{f}_1, \vec{x}_2, \vec{x}_3\}$.

## Solution (continued)

This gives us

$$\vec{f}_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Therefore,

$$\vec{u}_2 = \frac{1}{\sqrt{18}} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, \vec{u}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

and

$$U = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Finally,

$$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}.$$

## Problem

Find a singular value decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$.

## Solution

$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \left(\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\right)\begin{bmatrix} \sqrt{85} & 0 \\ 0 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{17}}\begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix}\right).$$

## Remark

Since there is only one non-zero eigenvalue, $\vec{u}_2$ (the second column of U) can not be found using the formula $\vec{u}_2 = \frac{1}{\sigma_2}A\vec{v}_2$. However, $\vec{u}_2$ can be chosen to be any unit vector orthogonal to $\vec{u}_1$; in this case, $\vec{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

## Problem

Find a singular value decomposition of $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

## Solution

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\|$$

$$\left(\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}\right)\begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{6}}\begin{bmatrix} 1 & -2 & 1 \\ -\sqrt{3} & 0 & \sqrt{3} \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix}\right)$$

$\blacksquare$

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

# Fundamental Subspaces

Full Singular Value Decomposition

Full Singular Value Decomposition

Copyright

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

# Applications

## Example (Polar Decomposition)

$$a + bi = \underbrace{\sqrt{a^2 + b^2}}_{\text{radius}} \underbrace{e^{i\theta}}_{\text{rotation}} .$$

Similarly, any square matrix

$$A = U\Sigma V^T = \underbrace{U\Sigma U^T}_{\text{nonneg. def.}} \underbrace{UV^T}_{\text{rotation}}$$

## Definition

A real $n \times n$ matrix $G$ is nonnegative definite (or positive in the book) if it is symmetric and for all $\vec{x} \in \mathbb{R}^n$

$$\vec{x}^T G \vec{x} \geq 0.$$

Example (Generalized inverse)

## Example (Image of unit ball under linear transform A)

Let $A = U\Sigma V^T$ be the full SVD for an $m \times n$ matrix A. We will see how the unit ball will be mapped:

$$\{A\vec{x} \mid ||\vec{x}|| \leq 1\}$$

The linear map $\vec{y} = A\vec{x}$ is trying to do the following things:

1. Rotate the n-vector $\vec{x}$ by $V^T$
2. Stretch along axes by $\sigma_i$ with $\sigma_i = 0$ for $i >$ rank (A)
3. Zero-pad for tall matrix (i.e., $m > n$) or truncate for fat matrix (i.e., $m < n >$) to get m-vector
4. Rotate the m-vector by $U^T$

# Example (Image of unit ball under linear transform A – continued)

# Example (Image Compression)

"I think you should be more explicit here in step two."

Image is a A is a $300 \times 300$ matrix.

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