

# Math 524: Linear Algebra

## Notes #7.2 — Operators on Inner Product Spaces

Peter Blomgren  
(blomgren@sdsu.edu)

Department of Mathematics and Statistics  
Dynamical Systems Group  
Computational Sciences Research Center  
San Diego State University  
San Diego, CA 92182-7720

<http://terminus.sdsu.edu/>

Fall 2021

(Revised: December 7, 2021)



# Outline

- 1 Student Learning Targets, and Objectives
  - SLOs: Operators on Inner Product Spaces
- 2 Positive Operators and Isometries
  - Positive Operators
  - Isometries
- 3 Polar Decomposition and Singular Value Decomposition
  - Polar Decomposition
  - Singular Value Decomposition
- 4 Problems, Homework, and Supplements
  - Suggested Problems
  - Assigned Homework
  - Supplements

# Student Learning Targets, and Objectives

## Target Positive Operators

**Objective** Be able to characterize Positive Operators, and in particular construct the Unique Positive Square Root Operator.

## Target Isometries

**Objective** Be able to state the definition of, and characterize Isometries

## Target Polar Decomposition

**Objective** Be able to *abstractly construct*\* the Polar Decomposition of an Operator, through Identification of the appropriate Isometry and Postive Operator.

## Target Singular Value Decomposition

**Objective** Be able to *abstractly construct*\* the Singular Value Decomposition of an Operator, by Identifying the Singular Values and Orthonormal Bases.

---

\* Generally practical constructions must be addressed with computational tools from [MATH 543].

# Positive Operators

## Definition (Positive Operator)

An operator  $T \in \mathcal{L}(V)$  is called **positive** if  $T$  is self-adjoint and

$$\langle T(v), v \rangle \geq 0$$

$\forall v \in V.$

If  $V$  is a complex vector space, then the requirement that  $T$  is self-adjoint can be dropped from the definition above:

Rewind (Over  $\mathbb{C}$ ,  $\langle T(v), v \rangle \in \mathbb{R} \forall v \in V$  Only for Self-Adjoint Operators [NOTES#7.1])

Suppose  $V$  is a complex inner product space and  $T \in \mathcal{L}(V)$ . Then  $T$  is self-adjoint **if and only if**

$$\langle T(v), v \rangle \in \mathbb{R}$$

$\forall v \in V.$



# Positive Operators

## Example (Positive Operators)

- If  $U$  is a subspace of  $V$ , then the orthogonal projections  $P_U$  and  $P_{U^\perp}$  are positive operators
- If  $T \in \mathcal{L}(V)$  is self-adjoint and  $b, c \in \mathbb{R}$  are such that  $b^2 < 4c$ , then  $(T^2 + bT + cI)$  is a positive operator, as shown by the proof of [INVERTIBLE QUADRATIC (OPERATOR) EXPRESSIONS (NOTES#7.1)]

## Rewind (Invertible Quadratic (Operator) Expressions [NOTES#7.1])

Suppose  $T \in \mathcal{L}(V)$  is self-adjoint, and  $b, c \in \mathbb{R} : b^2 < 4c$ , then

$$T^2 + bT + cI$$

is invertible.

# Square Root

## Definition (Square Root)

An operator  $R$  is called a **square root** of an operator  $T$  if  $R^2 = T$ .

## Example (Square Root)

If  $T \in \mathcal{L}(\mathbb{F}^3)$  is defined by  $T(z_1, z_2, z_3) = (z_3, 0, 0)$ , then the operator  $R \in \mathcal{L}(\mathbb{F}^3)$  defined by  $R(z_1, z_2, z_3) = (z_2, z_3, 0)$  is a square root of  $T$ :

$$R^2(z_1, z_2, z_3) = R(z_2, z_3, 0) = (z_3, 0, 0) = T(z_1, z_2, z_3)$$

## Example ( $n$ -th Roots?)

If  $T \in \mathcal{L}(\mathbb{F}^{n+1})$  is defined by  $T(z_1, \dots, z_{n+1}) = (z_{n+1}, 0, \dots, 0)$ , then the operator  $R \in \mathcal{L}(\mathbb{F}^{n+1})$  defined by  $R(z_1, \dots, z_{n+1}) = (z_2, z_3, \dots, z_{n+1}, 0)$  is an  $n$ th root of  $T$ :

$$\begin{aligned} R^n(z_1, \dots, z_n) &= R^{n-1}(z_2, z_3, \dots, z_{n+1}, 0) = R^{n-2}(z_3, z_4, \dots, z_{n+1}, 0, 0) \\ &= \dots = (z_{n+1}, 0, \dots, 0) = T(z_1, \dots, z_{n+1}) \end{aligned}$$



# “Positive” vs “Non-Negative” vs “Semi-Positive”

## Comment (“Positive” vs “Non-Negative” vs “Semi-Positive”)

The positive operators correspond to the numbers  $[0, \infty)$ , so a more precise terminology would use the term **non-negative** instead of positive.

However, operator-theorists consistently call these the positive operators.

Restricted to the Matrix-Vector “universe” we tend to talk about (strictly) *Positive Definite* and *Positive Semi-Definite* Matrices (“Matrix-Operators,” if you want).

# Characterization of Positive Operators

## Theorem (Characterization of Positive Operators)

Let  $T \in \mathcal{L}(V)$ , then the following are equivalent

- (a)  $T$  is positive
- (b)  $T$  is self-adjoint and all the eigenvalues of  $T$  are non-negative
- (c)  $T$  has a positive square root
- (d)  $T$  has a self-adjoint square root;
- (e) there exists an operator  $R \in \mathcal{L}(V)$  such that  $T = R^*R$

MATRICES: CHOLESKY FACTORIZATION; OR “HERMITIAN LU-FACTORIZATION”



## Characterization of Positive Operators

## Proof (Characterization of Positive Operators)

(a) $\Rightarrow$ (b)  $T$  is positive ( $\langle T(v), v \rangle \geq 0$ ), and by ( $\mathbb{R}$ :definition or  $\mathbb{C}$ :[NOTES#7.1](#)) we also have  $T = T^*$ ); suppose  $\lambda$  is an eigenvalue of  $T$  and  $v$  the corresponding eigenvector, then

$$0 \leq \langle T(v), v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle$$

$$\Rightarrow \lambda \in [0, \infty)$$

$$\Rightarrow \text{(b)}$$

## Characterization of Positive Operators

## Proof (Characterization of Positive Operators)

(b) $\Rightarrow$ (c)  $T$  is self-adjoint ( $T = T^*$ ) and  $\lambda(T) \in [0, \infty)$ . By [COMPLEX SPECTRAL THEOREM (NOTES#7.1)] or [REAL SPECTRAL THEOREM (NOTES#7.1)], there is an orthonormal basis  $v_1, \dots, v_n$  of  $V$  consisting of eigenvectors of  $T$ ; let  $\lambda_k : T(v_k) = \lambda_k v_k$ ; thus  $\lambda_k \in [0, \infty)$ . Let  $R \in \mathcal{L}(V)$  such that

$$R(v_k) = \sqrt{\lambda_k} v_k, \quad k = 1, \dots, n$$

$R$  is a positive operator, and  $R^2(v_k) = \lambda_k v_k = T(v_k)$ ,  $k = 1, \dots, n$ ; i.e.  $R^2 = T$ .

Thus  $R$  is a positive square root of  $T$ .  $\Rightarrow$  (c)

# Characterization of Positive Operators

## Proof (Characterization of Positive Operators)

(c) $\Rightarrow$ (d) By definition, every positive operator is self-adjoint.

(d) $\Rightarrow$ (e) Assume  $\exists R \in \mathcal{L}(V)$  so that  $R = R^*$  and  $R^2 = T$ :

Then  $T = R^*R$

$\Rightarrow$  (e)

(e) $\Rightarrow$ (a) Suppose  $\exists R \in \mathcal{L}(V) : T = R^*R$ , then  $T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$ . (which makes  $T$  self-adjoint).

Also,

$$\langle T(v), v \rangle = \langle (R^*R)(v), v \rangle = \langle R(v), R(v) \rangle \geq 0$$

$\forall v \in V$ , hence  $T$  is positive.

$\Rightarrow$  (a)

We now have (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a).  $\checkmark$

## Uniqueness of the Square Root

Theorem (Each Positive Operator Has Only One Positive Square Root)

*Every positive operator on  $V$  has a unique positive square root.*

Comment (“Positive Operators Act Like Real Numbers”)

Each non-negative number has a unique non-negative square root. Again, positive operators have “real” properties.

Comment (What is Unique?)

A positive operator can have infinitely many square roots; **only one of them can be positive.**

# Uniqueness of the Square Root

## Proof (Each Positive Operator Has Only One Positive Square Root)

Suppose  $T \in \mathcal{L}(V)$  is positive; let  $t \in V$  be an eigenvector, and  $\lambda^{(T)} \geq 0$ :  $T(t) = \lambda^{(T)}t$ .

Let  $R$  be a positive square root of  $T$ .

NOTE: We show  $R(t) = \sqrt{\lambda^{(T)}}t \Rightarrow$  the action of  $R$  on the eigenvectors of  $T$  is uniquely determined. Since there is a basis of  $V$  consisting of eigenvectors of  $T$  [ $\mathbb{C}/\mathbb{R}$  SPECTRAL THEOREM (NOTES#7.1)], this implies that  $R$  is uniquely determined.

To show that  $R(t) = \sqrt{\lambda^{(T)}}t$ , we use the fact that [ $\mathbb{C}/\mathbb{R}$  SPECTRAL THEOREM (NOTES#7.1)] guarantees an orthonormal basis  $r_1, \dots, r_n$  of  $V$  consisting on eigenvectors of  $R$ . Since  $R$  is a positive operator  $\lambda^{(R)} \geq 0 \Rightarrow \exists \lambda_1^{(R)}, \dots, \lambda_n^{(R)} \geq 0$  such that  $R(r_k) = \lambda_k^{(R)}r_k$  for  $k = 1, \dots, n$ .

→

→

→

→

→



# Uniqueness of the Square Root

## Proof (Each Positive Operator Has Only One Positive Square Root)

Since  $r_1, \dots, r_n$  is a basis of  $V$ , we can write  $t \stackrel{!}{=} (a_1 r_1 + \dots + a_n r_n)$ , for  $a_1, \dots, a_n \in \mathbb{F}$ , thus

$$R(t) = a_1 \lambda_1^{(R)} r_1 + \dots + a_n \lambda_n^{(R)} r_n$$

$$R^2(t) = a_1 (\lambda_1^{(R)})^2 r_1 + \dots + a_n (\lambda_n^{(R)})^2 r_n$$

But  $R^2 = T$  (by assumption, it is a positive square root of  $T$ ), and  $T(t) = \lambda^{(T)} t$ ; therefore, the above implies

$$a_1 \lambda^{(T)} r_1 + \dots + a_n \lambda^{(T)} r_n = a_1 (\lambda_1^{(R)})^2 r_1 + \dots + a_n (\lambda_n^{(R)})^2 r_n$$

$\Rightarrow a_j (\lambda^{(T)} - (\lambda_j^{(R)})^2) = 0, j = 1, \dots, n$  (either  $a_j = 0$ , or  $(\lambda^{(T)} - (\lambda_j^{(R)})^2) = 0$ ).

$$\text{Hence, } t = \sum_{j:a_j \neq 0} a_j r_j \Rightarrow R(t) = \sum_{j:a_j \neq 0} a_j \sqrt{\lambda^{(T)}} r_j = \sqrt{\lambda^{(T)}} t,$$

which is what we needed to show.  $\checkmark$



## Isometries — Norm-Preserving Operators

## Definition (Isometry)

- An operator  $S \in \mathcal{L}(V)$  is called an **isometry** if

$$\|S(v)\| = \|v\|$$

$$\forall v \in V.$$

- “An operator is an isometry if it preserves norms.”

## Rewind (Orthogonal Transformations [MATH-254 (NOTES#5.3)])

A linear transformation  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called orthogonal if it preserves the length of vectors:

$$\|T(\vec{x})\| = \|\vec{x}\|, \quad \forall \vec{x} \in \mathbb{R}^n.$$

If  $T(\vec{x}) = A\vec{x}$  is an orthogonal transformation, we say that  $A$  is an orthogonal (or *unitary*, when it has complex entries) matrix.



## Isometries — Norm-Preserving Operators

## Example

Suppose  $\lambda_1, \dots, \lambda_n$  are scalars with  $|\lambda_k| = 1$ , and  $S \in \mathcal{L}(V)$  satisfies  $S(s_j) = \lambda_j s_j$  for some orthonormal basis  $s_1, \dots, s_n$  of  $V$ .

We demonstrate that  $S$  is an isometry.

Let  $v \in V$ , then

$$\begin{aligned} v &= \langle v, s_1 \rangle s_1 + \cdots + \langle v, s_n \rangle s_n \\ \|v\|^2 &\stackrel{1}{=} |\langle v, s_1 \rangle|^2 + \cdots + |\langle v, s_n \rangle|^2 \\ \hline S(v) &= \langle v, s_1 \rangle S(s_1) + \cdots + \langle v, s_n \rangle S(s_n) \\ &= \lambda_1 \langle v, s_1 \rangle s_1 + \cdots + \lambda_n \langle v, s_n \rangle s_n \\ \|S(v)\|^2 &\stackrel{1}{=} |\lambda_1|^2 |\langle v, s_1 \rangle|^2 + \cdots + |\lambda_n|^2 |\langle v, s_n \rangle|^2 \\ &= |\langle v, s_1 \rangle|^2 + \cdots + |\langle v, s_n \rangle|^2 \end{aligned}$$

<sup>1</sup> [WRITING A VECTOR AS A LINEAR COMBINATION OF ORTHONORMAL BASIS (NOTES#6)]



## Characterization of Isometries

## Theorem (Characterization of Isometries)

Suppose  $S \in \mathcal{L}(V)$ , then the following are equivalent:

- (a)  $S$  is an isometry
- (b)  $\langle S(u), S(v) \rangle = \langle u, v \rangle \quad \forall u, v \in V$
- (c)  $S(u_1), \dots, S(u_n)$  is orthonormal for every orthonormal list of vectors  $u_1, \dots, u_n$  in  $V$
- (d) there exists an orthonormal list of vectors  $u_1, \dots, u_n$  of  $V$  such that  $S(u_1), \dots, S(u_n)$  is orthonormal
- (e)  $S^*S = I$
- (f)  $SS^* = I$
- (g)  $S^*$  is an isometry
- (h)  $S$  is invertible and  $S^{-1} = S^*$



## Some Help for the Proof

## Theorem (The Inner Product on a Real Inner Product Space)

Suppose  $V$  is a real inner product space, then

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

$\forall u, v \in V$ .

## Theorem (The Inner Product on a Complex Inner Product Space)

Suppose  $V$  is a complex inner product space, then

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2}{4}$$

$\forall u, v \in V$ .

The proofs for these identities are by “direct computation” (very similar to what we did in [NOTES#7.1]). The bottom line is that we can express the inner product in terms of the norm.

## Characterization of Isometries

## Proof (Characterization of Isometries)

- (a) $\Rightarrow$ (b) Suppose  $S$  is an isometry; the “help theorems” show that inner products can be computed from norms. Since  $S$  preserves norms,  $\Rightarrow S$  preserves inner products.  $\Rightarrow$  (b)
- (b) $\Rightarrow$ (c) Assume  $S$  preserves inner products, let  $u_1, \dots, u_n$  be an orthonormal list of vectors in  $V$ ;  $S(u_1), \dots, S(u_n)$  must be an orthonormal list of vectors since  $\langle S(u_i), S(u_j) \rangle = \langle u_i, u_j \rangle = \delta_{ij}$ .  $\Rightarrow$  (c)
- (c) $\Rightarrow$ (d)  $\checkmark$

## Characterization of Isometries

## Proof (Characterization of Isometries)

(d) $\Rightarrow$ (e) Let  $u_1, \dots, u_n$  be an orthonormal basis of  $V$  such that  $S(u_1), \dots, S(u_n)$  is orthonormal. Thus

$$\langle S^*S(u_j), u_k \rangle = \langle S(u_j), S(u_k) \rangle = \langle u_j, u_k \rangle$$

All  $v, w \in V$  can be written as unique linear combinations of  $u_1, \dots, u_n$ , therefore  $\langle S^*S(v), w \rangle = \langle v, w \rangle \Rightarrow S^*S = I. \Rightarrow$  (e)

(e) $\Rightarrow$ (f)  $S^*S = I. \Rightarrow \{S^*(SS^*) = S^*, (SS^*)S = S\} \Rightarrow SS^* = I. \Rightarrow$  (f)

(f) $\Rightarrow$ (g)  $SS^* = I$ , let  $v \in V$ , then

$$\|S^*(v)\|^2 = \langle S^*(v), S^*(v) \rangle = \langle SS^*(v), v \rangle = \langle v, v \rangle = \|v\|^2$$

$\Rightarrow S^*$  is an isometry.  $\Rightarrow$  (g)

## Characterization of Isometries

## Proof (Characterization of Isometries)

(g) $\Rightarrow$ (h)  $S^*$  is an isometry. We can apply the previously shown parts of the theorem, in particular (a) $\Rightarrow$ (e), and (a) $\Rightarrow$ (f) to  $S^*$  (with  $(S^*)^*$ ). This gives  $S^*S = SS^* = I$ , which means that  $S$  is invertible, and  $S^{-1} = S^*$ .  $\Rightarrow$  (h)

(h) $\Rightarrow$ (a)  $S$  is invertible, and  $S^{-1} = S^*$ ; let  $v \in V$ , then

$$\|S(v)\|^2 = \langle S(v), S(v) \rangle = \langle (S^*S)(v), v \rangle = \langle v, v \rangle = \|v\|^2$$

that is  $S$  is an isometry.  $\Rightarrow$  (a)

We now have (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (g) $\Rightarrow$ (h) $\Rightarrow$ (a).  $\checkmark$

Description of Isometries when  $\mathbb{F} = \mathbb{C}$ Theorem (Description of Isometries when  $\mathbb{F} = \mathbb{C}$ )

Suppose  $V$  is a complex inner product space and  $S \in \mathcal{L}(V)$ . Then the following are equivalent:

- (a)  $S$  is an isometry
- (b) There is an orthonormal basis of  $V$  consisting of eigenvectors of  $S$  whose corresponding eigenvalues all have absolute value 1

Proof (Description of Isometries when  $\mathbb{F} = \mathbb{C}$ )

The example on slide 16 shows (b) $\Rightarrow$ (a). To show (a) $\Rightarrow$ (b), we assume  $S$  is an isometry and use [COMPLEX SPECTRAL THEOREM (NOTES#7.1)] to guarantee an orthonormal basis  $s_1, \dots, s_n$  of  $V$  consisting of eigenvectors of  $S$ . Let  $\lambda_1, \dots, \lambda_n$  be the corresponding eigenvalues. Then

$$|\lambda_j| = \|\lambda_j s_j\| = \|S(s_j)\| = \|s_j\| = 1,$$

that is  $|\lambda_j| = 1$   $j = 1, \dots, n$ .  $\checkmark$

Upcoming: [DESCRIPTION OF ISOMETRIES WHEN  $\mathbb{F} = \mathbb{R}$  (NOTES#7.2-PREVIEW)].

## “Preview”

Preview (Description of Isometries when  $\mathbb{F} = \mathbb{R}$ )

Suppose  $V$  is a real inner product space and  $S \in \mathcal{L}(V)$ . Then the following are equivalent:

- (a)  $S$  is an isometry
- (b) There is an orthonormal basis of  $V$  with respect to which  $S$  has a block-diagonal matrix such that each block on the diagonal is either a  $(1 \times 1)$  matrix containing 1 or  $-1$ , or is a  $(2 \times 2)$  matrix of the form

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \theta \in (0, \pi)$$

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 7C-**{1, 6}**



---

**7C-1:** Prove or give a counterexample: If  $T \in \mathcal{L}(V)$  is self-adjoint and there exists an orthonormal basis  $u_1, \dots, u_n$  of  $V$  such that  $\langle T(u_k), u_k \rangle \geq 0 \forall k$ , then  $T$  is a positive operator.

---

We have no theorem that helps us, so therefore we suspect the statement is false.

✱

## Constructing a Counter-Example

✱

Consider  $V = \mathbb{R}^2$ , with the standard inner product, and standard basis. Let  $T(x_1, x_2) = (x_2, x_1)$ .



✱

 $T$  is Self-Adjoint:

✱

$$\begin{aligned}\langle (x_1, x_2), T^*(y_1, y_2) \rangle &= \langle T(x_1, x_2), (y_1, y_2) \rangle = \langle (x_2, x_1), (y_1, y_2) \rangle \\ &= x_2 y_1 + x_1 y_2 = \langle (x_1, x_2), (y_2, y_1) \rangle\end{aligned}$$

---


$$T^*(y_1, y_2) = (y_2, y_1)$$

✱

However,  $T$  is not a Positive Operator

✱

$$\langle T(1, 0), (1, 0) \rangle = 0 \quad \text{OK}$$

$$\langle T(0, 1), (0, 1) \rangle = 0 \quad \text{OK}$$

$$\langle T(1, -1), (1, -1) \rangle = -2 \quad \text{NOT } \geq 0$$

## Analogies: $\mathbb{C}$ and $\mathcal{L}(V)$

$\mathbb{C}$	$\mathcal{L}(V)$
$z$	$T$
$z^*$	$T^*$
$z = \Re(z) \geq 0$ (non-negative)	$\langle T(v), v \rangle \geq 0$ (positive)
$z^*z =  z ^2 = 1$ (unit circle)	$T^*T = I$ (isometry)

Any complex  $z \in \mathbb{C} \setminus \{0\}$  can be written in the form

$$z = \left( \frac{z}{|z|} \right) |z| = \left( \frac{z}{|z|} \right) \sqrt{z^*z},$$

where, of course

$$w = \left( \frac{z}{|z|} \right) \in \{\text{unit circle}\},$$

# Polar Decomposition

Notation ( $\sqrt{T}$ , The Square Root of  $T$ )

If  $T$  is a positive operator, then  $\sqrt{T}$  is the unique positive square root of  $T$ .

**Note:**  $T^*T$  is a positive operator for every  $T \in \mathcal{L}(V)$

$$\langle (T^*T)(v), v \rangle = \langle T(v), T(v) \rangle = \|T(v)\|^2 \geq 0,$$

therefore  $\sqrt{T^*T}$  is always well defined.

Theorem (Polar Decomposition)

Suppose  $T \in \mathcal{L}(V)$ . Then there exists an isometry  $S \in \mathcal{L}(V)$  such that

$$T = S\sqrt{T^*T}$$



## Polar Decomposition

## Why Should We Care???

The [POLAR DECOMPOSITION THEOREM] shows that we can write **any** operator on  $V$  as the product of an isometry, and a positive operator.

The characterization of the positive operators is given by the [ $\mathbb{C}/\mathbb{R}$  SPECTRAL THEOREM (NOTES#7.1)]; and

- we have characterized the isometries over  $\mathbb{C}$  in [DESCRIPTION OF ISOMETRIES WHEN  $\mathbb{F} = \mathbb{C}$ ]; and
- have “previewed” the characterization over  $\mathbb{R}$  [DESCRIPTION OF ISOMETRIES WHEN  $\mathbb{F} = \mathbb{R}$  (NOTES#7.2–PREVIEW)].

Thus, the [POLAR DECOMPOSITION THEOREM] provides us with a “complete” characterization of all operators in the sense of the [ $\mathbb{C}/\mathbb{R}$  SPECTRAL THEOREM (NOTES#7.1)] and the matching [DESCRIPTION OF ISOMETRIES WHEN  $\mathbb{F} = \mathbb{C}$ , OR  $\mathbb{F} = \mathbb{R}$ ] results.

I do daresay, this is quite a major result, indeed.



# Polar Decomposition

## Proof (Polar Decomposition)

Let  $v \in V$ , then

$$\begin{aligned} \|T(v)\|^2 &= \langle T(v), T(v) \rangle &= \langle (T^*T)(v), v \rangle \\ &= \langle (\sqrt{T^*T})(\sqrt{T^*T})(v), v \rangle &= \langle (\sqrt{T^*T})(v), (\sqrt{T^*T})(v) \rangle \\ &= \|(\sqrt{T^*T})(v)\|^2 \end{aligned}$$

Thus

$$\|T(v)\| = \|(\sqrt{T^*T})(v)\|, \quad \forall v \in V. \quad (\text{PD-1})$$

We define a linear map  $S_1 : \text{range}(\sqrt{T^*T}) \mapsto \text{range}(T)$  by

$$S_1((\sqrt{T^*T})(v)) = T(v) \quad (\text{PD-2})$$

The goal is to extend  $S_1$  to an isometry  $S \in \mathcal{L}(V)$  so that  $T = S\sqrt{T^*T}$ ...



# Polar Decomposition

## Proof (Polar Decomposition)

First, we make sure  $S_1$  is well defined: let  $v_1, v_2 \in V$  such that  $\sqrt{T^*T}(v_1) = \sqrt{T^*T}(v_2)$ . For (PD-2) to make sense, we need  $T(v_1) = T(v_2)$ .

$$\begin{aligned} \|T(v_1) - T(v_2)\| &= \|T(v_1 - v_2)\| \stackrel{\text{(PD-1)}}{=} \|\sqrt{T^*T}(v_1 - v_2)\| \\ &= \|\sqrt{T^*T}(v_1) - \sqrt{T^*T}(v_2)\| = 0 \end{aligned}$$

Hence  $T(v_1) = T(v_2)$ , and  $S_1$  is well-defined (we leave the verification of the basic linear mapping properties as an “exercise.”)

By definition (PD-2)  $S_1 : \text{range}(\sqrt{T^*T}) \mapsto \text{range}(T)$ ; together with (PD-1), we have that

$$\|S_1(u)\| = \|u\|, \quad \forall u \in \text{range}(\sqrt{T^*T})$$



# Polar Decomposition

## Proof (Polar Decomposition)

Now, we extend  $S_1$  to an isometry  $S$  on all of  $V$ :

By construction  $S_1$  is injective ( $\text{null}(S_1) = \{0\}$  by norm-preservation), so the [FUNDAMENTAL THEOREM OF LINEAR MAPS (NOTES#3.1)] gives

$$\dim(\text{range}(\sqrt{T^*T})) = \dim(\text{range}(T))$$

By [DIMENSION OF THE ORTHOGONAL COMPLEMENT (NOTES#6)]

$$\dim(\text{range}(\sqrt{T^*T})^\perp) = \dim(\text{range}(T)^\perp)$$

Let  $e_1, \dots, e_m$  be an orthonormal basis of  $(\text{range}(\sqrt{T^*T}))^\perp$ , and  $f_1, \dots, f_m$  be an orthonormal basis of  $(\text{range}(T))^\perp$ .

*Both bases have the same length.*



# Polar Decomposition

## Proof (Polar Decomposition)

Now, we define linear map  $S_2 : (\text{range}(\sqrt{T^*T}))^\perp \mapsto (\text{range}(T))^\perp$  by

$$S_2(a_1e_1 + \cdots + a_me_m) = a_1f_1 + \cdots + a_mf_m$$

[THE NORM OF AN ORTHONORMAL LINEAR COMBINATION (NOTES#6)] guarantees  $\|S_2(w)\| = \|w\|$ ,  $\forall w \in (\text{range}(\sqrt{T^*T}))^\perp$ .

Due to [DIRECT SUM OF A SUBSPACE AND ITS ORTHOGONAL COMPLEMENT (NOTES#6)] any  $v \in V$  can be uniquely written in the form

$$v = u + w, \quad u \in \text{range}(\sqrt{T^*T}), \quad w \in (\text{range}(\sqrt{T^*T}))^\perp \quad (\text{PD-3})$$

## Polar Decomposition

## Proof (Polar Decomposition)

Now, we define  $S(v)$  by

$$S(v) = S_1(u) + S_2(w), \quad u \in \text{range}(\sqrt{T^*T}), \quad w \in (\text{range}(\sqrt{T^*T}))^\perp$$

$\forall v \in V$  we have

$$S(\sqrt{T^*T}(v)) = S_1(\sqrt{T^*T}(v)) = T(v)$$

so  $T = S\sqrt{T^*T}$ . We must show that  $S$  is an isometry; with the decomposition (PD-3)  $v = u + w$  ( $u \perp w$ ), we can use the [PYTHAGOREAN THEOREM ( $\approx 500$  BC)]:

$$\begin{aligned} \|S(v)\|^2 &= \|S_1(u) + S_2(w)\|^2 \stackrel{\text{PT}^*}{=} \|S_1(u)\|^2 + \|S_2(w)\|^2 \\ &= \|u\|^2 + \|w\|^2 \stackrel{\text{PT}}{=} \|v\|^2 \end{aligned}$$

$\stackrel{\text{PT}^*}{=}$  holds since  $S_1(u) \in \text{range}(T)$ , and  $S_2(w) \in \text{range}(T)^\perp$

# Polar Decomposition

## Comment (Diagonalizability)

When  $\mathbb{F} = \mathbb{C}$  let  $T = S\sqrt{T^*T}$  be the Polar Decomposition of an operator  $T \in \mathcal{L}(V)$ , where  $S$  is an isometry.

Then

- (1) there is an orthonormal basis,  $\mathfrak{B}_1(V)$ , of  $V$  with respect to which  $S$  has a diagonal matrix, and
- (2) there is an orthonormal basis,  $\mathfrak{B}_2(V)$ , of  $V$  with respect to which  $\sqrt{T^*T}$  has a diagonal matrix.

**WARNING:** Usually, there does **not** exist an orthonormal basis that diagonalizes  $\mathcal{M}(S)$ , and  $\mathcal{M}(\sqrt{T^*T})$  at the same time.

# Singular Value Decomposition

So far, we have used the eigenvalues (and eigenvectors) to describe the properties of operators.

## Rewind (Eigenspace, $E(\lambda, T)$ )

Suppose  $T \in \mathcal{L}(V)$  and  $\lambda \in \mathbb{F}$ . The **Eigenspace** of  $T$  corresponding to  $\lambda$  denoted  $E(\lambda, T)$  is defined to be

$$E(\lambda, T) = \text{null}(T - \lambda I)$$

$E(\lambda, T)$  is the set of all eigenvectors of  $T$  corresponding to  $\lambda$ , along with the 0 vector.

We are particularly interested in (obsessed with?) scenarios where we can find orthonormal bases; this is the focus of [SCHUR'S THEOREM (NOTES#6)], [COMPLEX SPECTRAL THEOREM (NOTES#7.1)], and [REAL SPECTRAL THEOREM (NOTES#7.1)]

In [POLAR DECOMPOSITION THEOREM] we needed (in general) 2 orthonormal bases to perform the decomposition. The **Singular Value Decomposition** is an “alternate” way to leverage the use of 2 bases.

# Singular Value Decomposition

## Definition (Singular Values, $\sigma$ )

Suppose  $T \in \mathcal{L}(V)$ . The **singular values** of  $T$  are the eigenvalues, in this context denoted  $\sigma_i$ , of  $\sqrt{T^*T}$ , with each eigenvalue repeated  $\dim(E(\sigma_i, \sqrt{T^*T}))$  times.

In applications, and algorithms, it is customary to sort the singular values in descending order,  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ .

The singular values of  $T$  are all non-negative, because they are the eigenvalues of the positive operator  $\sqrt{T^*T}$ .

## Singular Value Decomposition

Example ( $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$ )

Let  $T \in \mathcal{L}(\mathbb{F}^4)$  be defined by

$$T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$$

we find the singular values.

(1) First we find the eigenvalues,  $\lambda(T)$ ; consider:

$$\lambda(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$$

the only solutions are  $\lambda \in \{0, -3\}$ , and the eigenspaces are given by

$$\begin{cases} E(\lambda = 0, T) &= \text{span}((0, 0, 1, 0)) \\ E(\lambda = -3, T) &= \text{span}((0, 0, 0, 1)) \end{cases}$$

Since  $\dim(E(0, T)) + \dim(E(-3, T)) = 2 < 4 = \dim(\mathbb{F}^4)$  we cannot fully diagonalize the operator with an eigenbasis.

$\mathbb{F}^4 \neq E(-3, T) \oplus E(0, T) \Rightarrow$  No Diagonalization.

## Singular Value Decomposition

Example ( $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$ )

(2) Next, we find the adjoint  $T^*$ ;  $T^*T$ , and  $\sqrt{T^*T}$ :

$$\begin{aligned}\langle z, T^*(w) \rangle &= \langle T(z), w \rangle = \langle (0, 3z_1, 2z_2, -3z_4), (w_1, w_2, w_3, w_4) \rangle \\ &= 3z_1w_2 + 2z_2w_3 - 3z_4w_4 \\ &= \langle (z_1, z_2, z_3, z_4), (3w_2, 2w_3, 0, -3w_4) \rangle\end{aligned}$$

$$T^*(w) = (3w_2, 2w_3, 0, -3w_4)$$

$$T^*T(z) = T^*(0, 3z_1, 2z_2, -3z_4) = (9z_1, 4z_2, 0, 9z_4)$$

$$\sqrt{T^*T}(z) = (3z_1, 2z_2, 0, 3z_4)$$

$$\lambda(T^*) = \{-3, 0\}$$

$$\lambda(T^*T) = \{9, 4, 0\}$$

$$\lambda(\sqrt{T^*T}) = \{3, 2, 0\} \rightsquigarrow \text{the singular values}$$

## Singular Value Decomposition

Example ( $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$ )

(3) We need the eigenspaces of  $\sqrt{T^*T}$ :

$$E(0; \sqrt{T^*T}) = \text{span}((0, 0, 1, 0))$$

$$E(2; \sqrt{T^*T}) = \text{span}((0, 1, 0, 0))$$

$$E(3; \sqrt{T^*T}) = \text{span}((1, 0, 0, 0), (0, 0, 0, 1))$$

Thus, the singular values are  $\sigma(T) = \{3, 3, 2, 0\}$ .

Comment ( $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$ )

Note that  $\lambda(T) = \{0, -3\}$  did not "capture" the 2, but  $\sigma(T) = \{3, 3, 2, 0\}$  did.



## Singular Value Decomposition

Comment ( $T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4)$ )

$$\mathcal{M}(T, \{e_i\}) = \begin{bmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}, \quad \mathcal{M}(T^*, \{e_i\}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix},$$

$$\mathcal{M}(\sqrt{T^*T})^2 = \mathcal{M}(T^*)\mathcal{M}(T) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}^2$$

Using [EIGENVALUES AND DETERMINANTS: THE CHARACTERISTIC EQUATION (MATH-254, NOTES#7.2)], we can get

$$p_{\mathcal{M}(T)}(\lambda) = \lambda^3(\lambda + 3), \quad p_{\mathcal{M}(\sqrt{T^*T})}(\lambda) = \lambda(\lambda - 2)(\lambda - 3)^2$$

# Singular Value Decomposition

Each  $T \in \mathcal{L}(V)$  has  $\dim(V)$  singular values; this follows from [C/R SPECTRAL THEOREM (NOTES#7.1)], and [CONDITIONS EQUIVALENT TO DIAGONALIZABILITY (NOTES#5)] applied to the positive ( $\Rightarrow$  self-adjoint) operator  $\sqrt{T^*T}$ .

The next statement gives a characterization  $\forall T \in \mathcal{L}(V)$  in terms of the singular values, and two orthonormal bases of  $V$ .



# Singular Value Decomposition

## Theorem (Singular Value Decomposition)

Suppose  $T \in \mathcal{L}(V)$  has singular values  $\sigma_1, \dots, \sigma_n$ . Then there exists orthonormal bases  $v_1, \dots, v_n$ , and  $u_1, \dots, u_n$  of  $V$  such that

$$T(w) = \sigma_1 \langle w, v_1 \rangle u_1 + \dots + \sigma_n \langle w, v_n \rangle u_n$$

$\forall w \in V$ .

## Comment (The Fundamental Theorem of Data Science)

If you want to be Buzzword Compliant, you could call this the Fundamental Theorem of ~~Page-Rank~~ ~~Big/Data~~ Data Science



## Singular Value Decomposition

## Proof (Singular Value Decomposition)

By the [C/R SPECTRAL THEOREM (NOTES#7.1)], we can find an orthonormal basis  $v_1, \dots, v_n$  of  $V$  such that  $\sqrt{T^*T}(v_k) = \sigma_k v_k$ ,  $k = 1, \dots, n$ . Hence due to [WRITING A VECTOR AS A LINEAR COMBINATION OF ORTHONORMAL BASIS (NOTES#6)]  $\forall w \in V$

$$\begin{aligned}w &= \langle w, v_1 \rangle v_1 + \dots + \langle w, v_n \rangle v_n \\ \sqrt{T^*T}(w) &= \sqrt{T^*T}(\langle w, v_1 \rangle v_1 + \dots + \langle w, v_n \rangle v_n) \\ &= \sigma_1 \langle w, v_1 \rangle v_1 + \dots + \sigma_n \langle w, v_n \rangle v_n\end{aligned}$$

By [POLAR DECOMPOSITION],  $\exists$  an isometry  $S \in \mathcal{L}(V)$  such that  $T = S\sqrt{T^*T}$ ; thus

$$T(w) = \sigma_1 \langle w, v_1 \rangle S(v_1) + \dots + \sigma_n \langle w, v_n \rangle S(v_n)$$

Let  $u_k = S(v_k)$ ,  $k = 1, \dots, n$ , then  $u_1, \dots, u_n$  is an orthonormal basis [CHARACTERIZATION OF ISOMETRIES]; and we have

$$T(w) = \sigma_1 \langle w, v_1 \rangle u_1 + \dots + \sigma_n \langle w, v_n \rangle u_n$$

$\forall w \in V$ .  $\checkmark$

# Singular Value Decomposition

## Comment (Singular Value Decomposition and Polar Decomposition)

When considering linear maps  $T \in \mathcal{L}(V, W)$ , we considered

$$\mathcal{M}(T; \mathfrak{B}(V); \mathfrak{B}(W));$$

in the operator setting ( $W = V$ )  $T \in \mathcal{L}(V)$  we usually consider

$$\mathcal{M}(T; \mathfrak{B}(V)),$$

making the basis  $\mathfrak{B}(V)$  play both the input/domain and output/range roles.

In the Polar Decomposition setting, where  $T = S\sqrt{T^*T}$ , we may consider two bases for  $V$ ,  $\mathfrak{B}_1(V)$ , and  $\mathfrak{B}_2(V)$ , so that

$$\mathcal{M}(S; \mathfrak{B}_1(V)), \text{ and } \mathcal{M}(\sqrt{T^*T}; \mathfrak{B}_2(V))$$

both are diagonal matrices.

## Singular Value Decomposition

## Comment (Singular Value Decomposition)

Now, in the Singular Value Decomposition we use one basis  $\mathfrak{B}_1(V)$  for the input/domain side, and another  $\mathfrak{B}_2(V)$  for the output/range side, so that

$$\mathcal{M}(T; \mathfrak{B}_1(V), \mathfrak{B}_2(V)) = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n \end{bmatrix} = \text{diag}(\sigma_1, \dots, \sigma_n)$$

**Every**  $T \in \mathcal{L}(V)$  has orthonormal bases  $\mathfrak{B}_1(V) = (v_1, \dots, v_n)$  and  $\mathfrak{B}_2(V) = (u_1, \dots, u_n)$  so that

$$\mathcal{M}(T; \mathfrak{B}_1(V), \mathfrak{B}_2(V)) = \text{diag}(\sigma_1, \dots, \sigma_n)$$

## Singular Value Decomposition

The following result is useful when developing strategies for finding singular values:

### Theorem (Singular Values Without Taking Square Root of an Operator)

*Suppose  $T \in \mathcal{L}(V)$ . Then the singular values of  $T$  are the nonnegative square roots of the eigenvalues of  $T^*T$ , with each eigenvalue  $\sigma$  repeated  $\dim(E(\sigma, T^*T))$  times.*

### Proof (Singular Values Without Taking Square Root of an Operator)

The [C/R SPECTRAL THEOREM (NOTES#7.1)] implies that there is an orthonormal basis  $v_1, \dots, v_n$  and nonnegative numbers  $\sigma_1, \dots, \sigma_n$  such that  $T^*T(v_j) = \sigma_j v_j$ ,  $j = 1, \dots, n$ . As we have done previously, defining  $\sqrt{T^*T}(v_j) = \sqrt{\sigma_j} v_j$  gives the desired result.



⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 7D-**{4, 6, 7, 10}**



**7D-6:** Find the singular values of the differentiation operator  $D \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$  defined by  $Dp = p'$ , where the inner product on  $\mathcal{P}_2(\mathbb{R})$  is the “Legendre Inner Product”,  $\langle p, q \rangle = \int_{-1}^1 p(x)q(x) dx$ .

\*

## Reference Orthonormal Basis

\*

In [NOTES#6] we derived an orthonormal basis for  $\mathcal{P}_2(\mathbb{R})$  with this particular inner product:

$$u_0 = \sqrt{\frac{1}{2}}, \quad u_1 = \sqrt{\frac{3}{2}}x, \quad u_2 = \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right)$$

\*

Matrix wrt. the Standard Basis  $\{1, x, x^2\}$ 

\*

The matrix of  $D$  with respect to the Standard Basis of  $\mathcal{P}_2(\mathbb{R})$  is

$$\mathcal{M}(D, \{1, x, x^2\}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

which shows that the **only Eigenvalue of  $D$  is 0** by [DETERMINATION OF EIGENVALUES FROM UPPER-TRIANGULAR MATRIX (NOTES#5)].

\*

Toward Singular Values...

\*

However,  $\mathcal{M}(D, \{1, x, x^2\})$  cannot be used to compute the singular values since  $\{1, x, x^2\}$  is not an orthonormal basis.

\*

$$\mathcal{M}(T, \{u_0(x), u_1(x), u_2(x)\})$$

\*

Getting the coefficients for the matrix with respect to the reference orthonormal basis is a little messy, but not too bad:

$$\begin{aligned} D\left(\sqrt{\frac{1}{2}}\right) &= 0 \\ D\left(\sqrt{\frac{3}{2}}x\right) &= \sqrt{\frac{3}{2}} = \sqrt{3} \cdot \sqrt{\frac{1}{2}} \\ D\left(\sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)\right) &= \sqrt{\frac{45}{2}}x = \sqrt{15} \cdot \sqrt{\frac{3}{2}}x \end{aligned}$$

$$\mathcal{M}(T, \{u_0(x), u_1(x), u_2(x)\}) = \begin{bmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{M}(T^*, \{u_0(x), u_1(x), u_2(x)\}) = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & \sqrt{15} & 0 \end{bmatrix}$$

$$\mathcal{M}(T^*T) = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & \sqrt{15} & 0 \end{bmatrix} \begin{bmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{bmatrix}$$

Hence, by [SINGULAR VALUES WITHOUT TAKING SQUARE ROOT OF AN OPERATOR], we have

$$\sigma(T) = \{ \sqrt{15}, \sqrt{3}, 0 \}$$

## Suggested Problems

**7.C—1, 2, 4, 6, 7**

**7.D—1, 2, 4, 5, 6, 7, 10**

## Assigned Homework

## HW#7.2, Due Date in Canvas/Gradescope

7.C—2, 4, 7

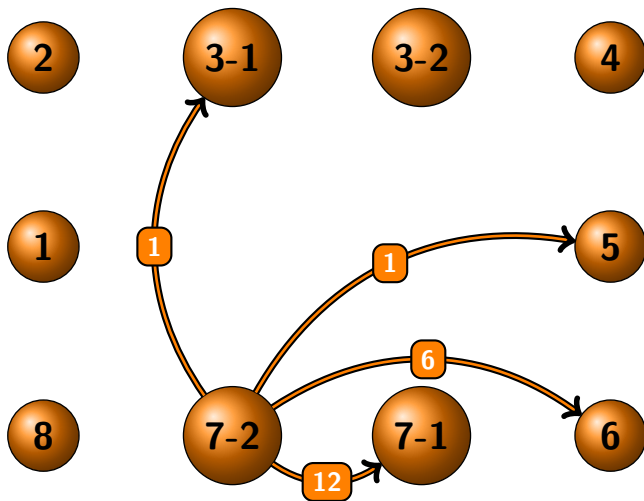
7.D—1, 2, 5

**Note:** Assignment problems are not official and subject to change until the first lecture on the chapter has been delivered (or virtually “scheduled.”)

Upload homework to [www.Gradescope.com](http://www.Gradescope.com)



## Explicit References to Previous Theorems or Definitions (with count)



## Explicit References to Previous Theorems or Definitions

