

10.3 Polar Decomposition

CAN WE GENERALIZE THE POLAR FORM FOR COMPLEX NUMBERS?

Since complex (1×1) -matrices are linear operators on the vector space \mathbb{C}^1 , we can ambitiously attempt to lift ideas from the complex numbers to linear maps. We have the following analogy in mind.

numbers	maps
complex number: $z \in \mathbb{C}$	linear map: $T: V \rightarrow W$
conjugate: $\bar{z} \in \mathbb{C}$	adjoint: $T^*: W \rightarrow V$
points on the unit circle: $\bar{z}z = 1$	isometries: $T^*T = \text{id}_V$
real numbers: $z = \bar{z}$	self-adjoint operators: $T = T^*$
nonnegative real numbers: $z \geq 0$	positive-semidefinite operators
polar form: $z = re^{i\theta}$??

Table 10.1: Analogy between complex numbers and linear operators

Given the success of the first parts in this analogy, one wonders if every linear operator can be expressed as a product of a positive-semidefinite operator and an isometry.

10.3.0 Lemma (Positive part). *Let V and W be finite-dimensional inner product spaces. For any linear map $T: V \rightarrow W$, let $\sqrt{T^*T}: V \rightarrow V$ denote the unique positive-semidefinite square root of the positive-semidefinite operator $T^*T: V \rightarrow V$.*

Theorem 10.2.2 and Corollary 10.2.3 already establish that T^*T has a unique positive-semidefinite square root.

- i. For any vector v in V , we have $\|T[v]\|_W = \|(\sqrt{T^*T})[v]\|_V$.
- ii. We have $\text{Ker}(T) = \text{Ker}(\sqrt{T^*T})$.

Proof.

- i. For any vector v in V , the definition [7.1.0] of the norm, properties [9.0.4] of adjoint maps, and the self-adjointness of the linear map $\sqrt{T^*T}$ give

$$\begin{aligned} \|T[v]\|_W^2 &= \langle T[v], T[v] \rangle_W = \langle (T^*T)[v], v \rangle_V = \langle (\sqrt{T^*T})^2[v], v \rangle_V = \langle (\sqrt{T^*T})[(\sqrt{T^*T})[v]], v \rangle_V \\ &= \langle (\sqrt{T^*T})[v], (\sqrt{T^*T})^*[v] \rangle_V = \langle (\sqrt{T^*T})[v], (\sqrt{T^*T})[v] \rangle_V = \|(\sqrt{T^*T})[v]\|_V^2. \end{aligned}$$

The nonnegative [7.0.0] of inner products shows that, by taking the square root, we have $\|T[v]\|_W = \|(\sqrt{T^*T})[v]\|_V$.

- ii. The definition of the kernel [3.1.0], the positivity [7.1.1] of norms, and part i yields the following equivalences:

$$v \in \text{Ker}(T) \Leftrightarrow T[v] = \mathbf{0}_W \Leftrightarrow \|T[v]\|_W = 0 \Leftrightarrow \|(\sqrt{T^*T})[v]\|_V = 0 \Leftrightarrow (\sqrt{T^*T})[v] = \mathbf{0}_V \Leftrightarrow v \in \text{Ker}(\sqrt{T^*T}),$$

which proves that $\text{Ker}(T) = \text{Ker}(\sqrt{T^*T})$. □

10.3.1 Theorem (Polar decomposition). *Let V and W be two finite-dimensional inner product spaces such that $\dim W \geq \dim V$. For any linear map $T: V \rightarrow W$, there exists an isometry $S: V \rightarrow W$ such that $T = S\sqrt{T^*T}$.*

Proof. Since the linear operator $\sqrt{T^*T}$ is self-adjoint [10.2.3], the self-adjoint spectral theorem [10.1.4] implies that there exists an orthonormal basis of V consisting of eigenvectors for $\sqrt{T^*T}$. The basis vectors lying in the 0-eigenspace span $\text{Ker}(\sqrt{T^*T})$ and the basis vectors with nonzero eigenvalues span $\text{Im}(\sqrt{T^*T})$. Hence, for all vectors v in V , there exists unique vectors v' in $\text{Ker}(\sqrt{T^*T})$ and v'' in $\text{Im}(\sqrt{T^*T})$ such that $v = v' + v''$. The properties [10.1.3] of self-adjointness prove that the linear subspaces $\text{Ker}(\sqrt{T^*T})$ and $\text{Im}(\sqrt{T^*T})$ are orthogonal. To exhibit the isometry S , we construct linear maps on $\text{Ker}(\sqrt{T^*T})$ and $\text{Im}(\sqrt{T^*T})$ separately.

Set $n := \dim V$ and $r := \dim \text{Im}(T)$. The dimension formula [3.1.6] shows that $\dim \text{Ker}(T) = n - r$ and part *ii* of the positive part lemma shows that $\text{Ker}(T) = \text{Ker}(\sqrt{T^*T})$. Choose an orthonormal basis u_1, u_2, \dots, u_{n-r} for the linear subspace $\text{Ker}(\sqrt{T^*T}) \subseteq V$. Similarly, set $m := \dim W$, choose an orthonormal basis w_1, w_2, \dots, w_r for the linear subspace $\text{Im}(T) \subseteq W$, and extend it to an orthonormal basis w_1, w_2, \dots, w_m of W . Let $W' := \text{Span}(w_{r+1}, w_{r+2}, \dots, w_m)$. By construction, the linear subspaces W' and $\text{Im}(T)$ are orthogonal and, by hypothesis, we have $\dim W' = m - r \geq n - r = \dim \text{Ker}(\sqrt{T^*T})$. The linear map $S_1: \text{Ker}(\sqrt{T^*T}) \rightarrow W'$ is defined, for all $1 \leq j \leq n - r$, by $S_1[u_j] = w_{r+j}$. Using the Parseval identity [7.1.5] twice gives

$$\begin{aligned} \|S_1[c_1 u_1 + c_2 u_2 + \dots + c_{n-r} u_{n-r}]\|_W^2 &= \|c_1 w_{r+1} + c_2 w_{r+2} + \dots + c_{n-r} w_n\|_W^2 \\ &= |c_1|^2 + |c_2|^2 + \dots + |c_{n-r}|^2 = \|c_1 u_1 + c_2 u_2 + \dots + c_{n-r} u_{n-r}\|_V^2. \end{aligned}$$

The nonnegativity [7.0.0] shows that, by taking the square root, we obtain $\|S_1[u]\|_W = \|u\|_V$ for all vectors u in $\text{Ker}(\sqrt{T^*T})$.

We next focus on $\text{Im}(\sqrt{T^*T})$. Consider vectors v_1 and v_2 in V such that $(\sqrt{T^*T})[v_1] = (\sqrt{T^*T})[v_2]$. Part *i* of the positive part lemma and the linearity of the maps give

$$\|T[v_1] - T[v_2]\|_W = \|T[v_1 - v_2]\|_W = \|(\sqrt{T^*T})[v_1 - v_2]\|_V = \|(\sqrt{T^*T})[v_1] - (\sqrt{T^*T})[v_2]\|_V = 0,$$

so the properties [7.1.1] of norms show that $T[v_1] = T[v_2]$. Hence, the linear map $S_2: \text{Im}(\sqrt{T^*T}) \rightarrow \text{Im}(T)$ defined, for all vectors v in V , by $S_2[(\sqrt{T^*T})[v]] = T[v]$ is well-defined. Part *i* of the lemma also implies that, for all vectors v in $\text{Im}(\sqrt{T^*T})$, we have $\|S_2[v]\|_W = \|v\|_V$.

Combining S_1 and S_2 gives the linear map $S: V \rightarrow W$ defined by $S[v] = S_1[v'] + S_2[v'']$ where $v = v' + v''$, $v' \in \text{Ker}(\sqrt{T^*T})$, and $v'' \in \text{Im}(\sqrt{T^*T})$. For all vectors v in V , we have

$$(S\sqrt{T^*T})[v] = S[(\sqrt{T^*T})[v]] = S_2[(\sqrt{T^*T})[v]] = T[v],$$

so $T = S\sqrt{T^*T}$. Moreover, the Pythagorean theorem [7.1.2] gives

$$\|S[v]\|_W^2 = \|S_1[v'] + S_2[v'']\|_W^2 = \|S_1[v']\|_W^2 + \|S_2[v'']\|_W^2 = \|v'\|_V^2 + \|v''\|_V^2 = \|v\|_V^2,$$

which proves that S is an isometry. \square

10.3.2 Problem. Find the polar decomposition of $\mathbf{A} := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & 1 & 0 \end{bmatrix}$.

Solution. Since $\mathbf{A}^* \mathbf{A} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, it

follows that $\sqrt{\mathbf{A}^* \mathbf{A}} = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$ and

$$\begin{aligned} \mathbf{S} &= \mathbf{A}(\sqrt{\mathbf{A}^* \mathbf{A}})^{-1} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \end{bmatrix}. \end{aligned} \quad \square$$

10.4 Singular-Value Decomposition

CAN WE EXTEND THE SPECTRAL THEOREMS TO ALL LINEAR MAPS?

To associate a diagonal matrix to every linear map, we need a pair of ordered bases: one for the source and another for the target.

10.4.0 Definition. Let V and W be finite-dimensional inner product spaces. The *singular values* of the linear map $T: V \rightarrow W$ are the eigenvalues of the linear operator $\sqrt{T^* T}: V \rightarrow V$. Since $\sqrt{T^* T}$ is the unique positive-semidefinite square root of $T^* T: V \rightarrow V$, the singular values of T are nonnegative real numbers and they are typically listed in increasing order.

10.4.1 Theorem (Singular-value decomposition). *Let V and W be finite-dimensional inner product spaces such that $m := \dim W \geq \dim V =: n$. For any linear map $T: V \rightarrow W$ with singular values $\sigma_1, \sigma_2, \dots, \sigma_n$, there exists an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of V and an orthonormal basis $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ of W such that, for all vectors \mathbf{v} in V , we have*

$$T[\mathbf{v}] = \sigma_1 \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{w}_1 + \sigma_2 \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{w}_2 + \cdots + \sigma_n \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{w}_n.$$

Proof. The self-adjoint spectral theorem [10.1.4] establishes that there exists an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of V consisting of eigenvectors for the self-adjoint linear operator $\sqrt{T^* T}$. The polar decomposition [10.3.1] shows that there exists an isometry $S: V \rightarrow W$ such that $T = S \sqrt{T^* T}$. Expressing the vector \mathbf{v} in V in terms of its orthonormal coordinate [7.1.7] and applying linear operator

$T = S \sqrt{T^* T}$, we obtain

$$\begin{aligned} T[v] &= (S \sqrt{T^* T})[v] \\ &= S[(\sqrt{T^* T})[\langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_2 + \cdots + \langle v, u_n \rangle u_n]] \\ &= S[\langle v, u_1 \rangle (\sqrt{T^* T})[u_1] + \langle v, u_2 \rangle (\sqrt{T^* T})[u_2] + \cdots + \langle v, u_n \rangle (\sqrt{T^* T})[u_n]] \\ &= S[\sigma_1 \langle v, u_1 \rangle u_1 + \sigma_2 \langle v, u_2 \rangle u_2 + \cdots + \sigma_n \langle v, u_n \rangle u_n] \\ &= \sigma_1 \langle v, u_1 \rangle S[u_1] + \sigma_2 \langle v, u_2 \rangle S[u_2] + \cdots + \sigma_n \langle v, u_n \rangle S[u_n]. \end{aligned}$$

The characterizations [9.1.3] of surjective isometries demonstrate that the vectors $S[u_1], S[u_2], \dots, S[u_n]$ form an orthonormal list for W . For all $1 \leq j \leq n$, set $w_j := S[u_j]$. Extending the orthonormal list w_1, w_2, \dots, w_n to an orthonormal basis of W completes the proof. \square

10.4.2 Corollary. Let m and n be positive integers such that $m \geq n$. For any complex $(m \times n)$ -matrix \mathbf{A} , there is a factorization $\mathbf{A} = \mathbf{P} \mathbf{\Sigma} \mathbf{Q}^*$ where \mathbf{P} is a unitary $(m \times m)$ -matrix, \mathbf{Q} is a unitary $(n \times n)$ -matrix, and $\mathbf{\Sigma}$ is a diagonal $(m \times n)$ -matrix whose diagonal entries are the singular values of \mathbf{A} .

Proof. Combining the singular-value decomposition theorem and the changes of basis theorem [4.0.2] proves the claim. \square

10.4.3 Remark. The singular-value decomposition of an $(m \times n)$ -matrix \mathbf{A} , where $m \geq n$, can be computed using the following steps.

- Compute a unitary diagonalization of the product $\mathbf{A}^* \mathbf{A} = \mathbf{Q}^* \mathbf{\Lambda} \mathbf{Q}$ where $\mathbf{Q}^* \mathbf{Q} = \mathbf{I}$, $\mathbf{\Lambda} := \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$, and $\lambda_j = 0$ for all $r + 1 \leq j \leq n$.
- Consider the invertible $(r \times r)$ -matrix $\mathbf{D} := \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r})$ and let $\mathbf{\Sigma} := \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ be a diagonal $(m \times n)$ -matrix.
- For all $1 \leq j \leq r$, set $w_j := \frac{1}{\sqrt{\lambda_j}} \mathbf{A} u_j$ where the vector u_j denotes the j -th column in the matrix \mathbf{Q} . Extend the list w_1, w_2, \dots, w_r to an orthonormal basis of \mathbb{K}^m . This orthonormal basis determines the columns of the $(m \times m)$ -matrix \mathbf{P} .

10.4.4 Problem. Find a singular-value decomposition of $\mathbf{A} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Solution. Since

$$\mathbf{A}^* \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

we have

$$\mathbf{\Sigma} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and $\mathbf{Q} = \mathbf{I}$. Because $\mathbf{w}_1 = \frac{1}{2}\mathbf{a}_1$ and $\mathbf{w}_2 = \frac{1}{2}\mathbf{a}_2$, we obtain an orthonormal basis for \mathbb{R}^4 by choosing $\mathbf{w}_3 := \frac{1}{2}[1 \ 1 \ -1 \ -1]^\top$ and $\mathbf{w}_4 := \frac{1}{2}[1 \ -1 \ 1 \ -1]^\top$. Thus, a singular-value decomposition is

$$\mathbf{A} = \left(\frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \right) \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad \square$$

10.4.5 Problem. Find a singular value decomposition of

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

Solution. Since

$$\mathbf{B}^* \mathbf{B} = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix},$$

we see the eigenvalues are 18 and 0 with unit eigenvectors given by the columns of the matrix $\mathbf{Q} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Hence, we have

$$\mathbf{\Sigma} := \begin{bmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^\top.$$

Since we have $\mathbf{w}_1 = \frac{1}{3\sqrt{2}} \mathbf{A} \mathbf{u}_1 = \frac{1}{3}[1 \ -2 \ 2]^\top$, we may choose $\mathbf{w}_2 := \frac{1}{\sqrt{5}}[2 \ 1 \ 0]^\top$ and $\mathbf{w}_3 := \frac{1}{\sqrt{45}}[-2 \ 4 \ 1]^\top$. Thus, a singular-value decomposition is

$$\mathbf{B} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \right)^*. \quad \square$$

Exercises

10.4.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

- i. The singular values of any linear operator on a finite-dimensional vector space are also eigenvalues of the operator.
- ii. The singular values of any matrix \mathbf{A} are the eigenvalues of $\mathbf{A}^* \mathbf{A}$.
- iii. The singular values of any linear operator are nonnegative.
- iv. Every eigenvalue of a self-adjoint matrix \mathbf{A} is a singular value of \mathbf{A} .