Polar Decomposition and Singular Value Decomposition (极分解和奇异值分解)

Lecture 22

Dept. of Math., SUSTech

2023.04

Polar Decomposition and Singular Value Decomposition

1. Polar Decomposition
2. Singular Value Decomposition
3. Homework Assignment 22

$z\in\mathbb{C} \longleftrightarrow T\in\mathscr{L}(V)$ $\quad z\geq 0 \longleftrightarrow T$ positive

$\bar{z} \longleftrightarrow T^*$ $\quad |z|=1 \longleftrightarrow$ isometry

$z=\bar{z} \longleftrightarrow T=T^*$ $\quad z=\left(\frac{z}{|z|}\right)\sqrt{\bar{z}z} \longleftrightarrow T=S\sqrt{T^*T}$

Introduction

- Recall our analogy between $\mathbb{C}$ and $\mathscr{L}(V)$. Under this analogy, a complex number z corresponds to an operator T, and $\bar{z}$ corresponds to T^*. The real numbers $z=\bar{z}$ correspond to the self-adjoint operators ($T=T^*$), and the nonnegative numbers correspond to the (badly named) positive operators.
- Another distinguished subset of $\mathbb{C}$ is the unit circle, which consists of the complex numbers z such that $|z|=1$. The unit circle in $\mathbb{C}$ corresponds to the isometries.
- Continuing with our analogy, note that each complex number z except 0 can be written in the form

$$z=\left(\frac{z}{|z|}\right)|z|=\left(\frac{z}{|z|}\right)\sqrt{\bar{z}z}$$

where the first factor, namely, $z/|z|$, is an element of the unit circle. Our analogy leads us to guess that each operator $T\in\mathscr{L}(V)$ can be written as an isometry times $\sqrt{T^*T}$.

Polar Decomposition

Recall our analogy between $\mathbb{C}$ and $\mathscr{L}(V)$. Under this analogy, a complex number z corresponds to an operator T, and $\bar{z}$ corresponds to T^*.

7.44 Notation $\sqrt{T}$

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

Now we can state and prove the Polar Decomposition, which gives a beautiful description of an arbitrary operator on V.

$T \in \mathcal{L}(V)$

$S \in \mathcal{L}(V)$ isometry.

$\underbrace{S}_{\text{isometry}} \underbrace{\sqrt{T^*T}}_{\text{positive}}$, $T = S\sqrt{T^*T}$

7.45 Polar Decomposition

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

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

Proof

Proof. If $v \in V$, then

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

Thus

$$\|Tv\| = \|\sqrt{T^*T}v\| \text{ for all } v \in V \cdots\cdots (1).$$

Define a linear map $S_1 : \text{range } \sqrt{T^*T} \to \text{range } T$ by

$$S_1(\sqrt{T^*T}v) = Tv \cdots\cdots (2)$$

The idea of the proof is to extend S_1 to an isometry $S \in \mathscr{L}(V)$ such that $T = S\sqrt{T^*T}$. Now for the details.

Proof.

First we must check that S_1 is well defined. To do this, suppose $v_1, v_2 \in V$ are such that $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$. For the definition given by (2) to make sense, we must show that $Tv_1 = Tv_2$. Note that

$$
\begin{aligned}
||Tv_1 - Tv_2|| &= ||T(v_1 - v_2)|| \\
&= ||\sqrt{T^*T}(v_1 - v_2)|| \\
&= ||\sqrt{T^*T}v_1 - \sqrt{T^*T}v_2)|| = 0
\end{aligned}
$$

where the second equality holds by (1). The equation above shows that $Tv_1 = Tv_2$, so S_1 is indeed well defined. You should verify that S_1 is a linear map. We see from (2) that S_1 maps range $\sqrt{T^*T}$ onto range T. Clearly (1) and (2) imply that

$$||S_1 u|| = ||u||$$

for all $u \in$ range $\sqrt{T^*T}$.

Proof.

In particular, S_1 is injective. Thus from the Fundamental Theorem of Linear Maps (3.22), applied to S_1, we have

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

This implies that $\dim(\text{range } \sqrt{T^*T})^\perp = \dim(\text{range } T)^\perp$. Thus orthonormal bases $e_1, \cdots, e_m$ of $(\text{range } \sqrt{T^*T})^\perp$ and $f_1, \cdots, f_m$ of $(\text{range } T)^\perp$ can be chosen; the key point here is that these two orthonormal bases have the same length (denoted m). Now define a linear map
$S_2 : (\text{range } \sqrt{T^*T})^\perp \to (\text{range } T)^\perp$ by

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

For all $w \in (\text{range } \sqrt{T^*T})^\perp$, we have $\|S_2w\| = \|w\|$.

Proof.

Now let S be the operator on V that equals S_1 on $(\text{range}\ \sqrt{T^*T})$ and equals S_2 on $(\text{range}\ \sqrt{T^*T})^\perp$. More precisely, recall that each $v \in V$ can be written uniquely in the form

$$v = u + w,$$

where $u \in (\text{range}\ \sqrt{T^*T})$ and $w \in (\text{range}\ \sqrt{T^*T})^\perp$. For $v \in V$ with decomposition as above, define Sv by

$$Sv = S_1 u + S_2 w.$$

For each $v \in V$ we have

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

so $T = S\sqrt{T^*T}$, as desired.

Proof

All that remains is to show that S is an isometry. However, this follows easily from two uses of the Pythogorean Theorem: if $v \in V$ has decomposition $v = u + w$, then

$$||Sv||^2 = ||S_1u + S_2w||^2 = ||S_1u||^2 + ||S_2w||^2 = ||u||^2 + ||w||^2 = ||v||^2.$$

The Polar Decomposition states that each operator on V is the product of an isometry and a positive operator. Thus we can write each operator on V as the product of two operators, each of which comes from a class that we can completely describe and that we understand reasonably well.

Remarks

1. The isometries are described by 7.43 and 9.36; the positive operators are described by the Spectral Theorem (7.24 and 7.29).
2. Specifically, consider the case $\mathbb{F} = \mathbb{C}$, and suppose $T = S\sqrt{T^*T}$ is a Polar Decomposition of an operator $T \in \mathscr{L}(V)$, where S is an isometry. Then there is an orthonormal basis of V with respect to which S has a diagonal matrix, and there is an orthonormal basis of V with respect to which $\sqrt{T^*T}$ has a diagonal matrix.
3. Warning: there may not exist an orthonormal basis that simultaneously puts the matrices of both S and $\sqrt{T^*T}$ into these nice diagonal forms. In other words, S may require one orthonormal basis and $\sqrt{T^*T}$ may require a different orthonormal basis.

Singular Values

The eigenvalues of an operator tell us something about the behavior of the operator. Another collection of numbers, called the singular values, is also useful. Recall that eigenspaces and the notation E are defined in 5.36.

7.49 **Definition** ***singular values***

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

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

Example

7.50 **Example** Define $T \in \mathcal{L}(\mathbf{F}^4)$ by

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

Find the singular values of T.

Solution A calculation shows $T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4)$, as you should verify. Thus

$$\sqrt{T^*T}(z_1, z_2, z_3, z_4) = (3z_1, 2z_2, 0, 3z_4),$$

and we see that the eigenvalues of $\sqrt{T^*T}$ are $3, 2, 0$ and

$$\dim E(3, \sqrt{T^*T}) = 2, \ \dim E(2, \sqrt{T^*T}) = 1, \ \dim E(0, \sqrt{T^*T}) = 1.$$

Hence the singular values of T are $3, 3, 2, 0$.

Singular Value Decomposition

Each T has $\dim\ V$ singular values, as can be seen by applying the Spectral Theorem and 5.41 [see specially part (e)] to the positive (hence self-adjoint) operator $\sqrt{T^*T}$. For example, the operator T defined in Example 7.50 on the four-dimensional vector space $\mathbb{F}^4$ has four singular values, as we saw above.

Singular Value Decomposition

The next result shows that every operator on V has a clean description in terms of its singular values and two orthonormal bases of V.

$$T(e_1,\cdots,e_n)=(f_1,\cdots,f_n)\begin{pmatrix} s_1 & & 0 \\ & s_2 & \\ 0 & & \ddots s_n \end{pmatrix}$$

7.51 Singular Value Decomposition

Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \ldots, s_n$. Then there exist orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for every $v \in V$.

$$v=\langle v,e_1\rangle e_1+\cdots+\langle v,e_n\rangle e_n$$

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

$$\sqrt{T^*T}\,e_j=s_je_j$$

$$\begin{aligned}
Tv&=T(\langle v,e_1\rangle e_1+\cdots+\langle v,e_n\rangle e_n)\\
&=S\left(\sqrt{T^*T}(\langle v,e_1\rangle e_1+\cdots+\langle v,e_n\rangle e_n)\right)\\
&=S\langle v,e_1\rangle\sqrt{T^*T}e_1+\cdots+S\langle v,e_n\rangle\sqrt{T^*T}e_n\\
&=S(\langle v,e_1\rangle s_1e_1+\cdots+\langle v,e_n\rangle s_ne_n)\\
&=s_1\langle v,e_1\rangle\underbrace{\boxed{Se_1}}_{f_1}+\cdots+s_n\langle v,e_n\rangle\underbrace{\boxed{Se_n}}_{f_n}
\end{aligned}$$

Remarks

1. When we worked with linear maps from one vector space to a second vector space, we considered the matrix of a linear map with respect to a basis of the first vector space and a basis of the second vector space. When dealing with operators, which are linear maps from a vector space to itself, we almost always use only one basis, making it play both roles.
2. The Singular Value Decomposition allows us a rare opportunity to make good use of two different bases for the matrix of an operator. To do this,

Singular Value Decomposition

suppose $T \in \mathcal{L}(V)$. Let $s_1, \ldots, s_n$ denote the singular values of T, and let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be orthonormal bases of V such that the Singular Value Decomposition 7.51 holds. Because $Te_j = s_j f_j$ for each j, we have

$$\boxed{\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)\big) = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix}}.$$

Singular values without taking square root of an operator

The singular values of T can be approximated without computing the square root of T^*T. The next result helps justify working with T^*T instead of $\sqrt{T^*T}$:

7.52 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 λ repeated $\dim E(\lambda, T^*T)$ times.

Proof The Spectral Theorem implies that there are an orthonormal basis $e_1, \dots, e_n$ and nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $T^*Te_j = \lambda_j e_j$ for $j = 1, \dots, n$. It is easy to see that $\sqrt{T^*T}e_j = \sqrt{\lambda_j}e_j$ for $j = 1, \dots, n$, which implies the desired result. ■

Homework Assignment 22

7.D: 5, 7, 9, 10, 12, 13, 15, 17.