Linear Algebra Done Right
Notes of Section 7.D

> Good mathematicians see analogies between theorems; great mathematicians see analogies between analogies.
>
> *Stefan Banach*

1 Motivation

Axler motivates the polar decomposition by continuing his analogy between the complex numbers, **C**, and the space of operators, $\mathcal{L}(V)$.

C concept	**notation**	$\mathcal{L}(V)$ **concept**	**notation**
Number	z	Operator	T
Complex conjugate	$\bar{z}$	Adjoint	T^*
Real numbers	$z = \bar{z}$	Self-adjoint	$T = T^*$
Nonnegative numbers[1]	$z \geq 0$	Positive operators[2]	$\langle Tv, v \rangle \geq 0$
Unit circle	$\bar{z}z = \lvert z\rvert^2 = 1$	Isometry[3]	$T^*T = I$

Table 1: An analogy between complex numbers and operators.

From this analogy, it follows that each complex number z (with the exception of 0) can be written as

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

where $\frac{z}{|z|}$ is an element of the unit circle and $\sqrt{\bar{z}z}$ is nonnegative. Following this decomposition of a complex number, we write each operator T as

$$T = S\sqrt{T^*T} \tag{1}$$

where S is an isometry and $\sqrt{T^*T}$ is a positive operator.

[1] Note that z is real
[2] Note that T is self-adjoint
[3] Note that the definition is $\|Sv\| = \|v\|$

2 Polar Decomposition

We call (1) the polar decomposition, and while Axler's analogy motivates it well, we still need to prove that it exists. Note in the decomposition of the complex number, our element of the unit circle was easy to find ($\frac{z}{|z|}$). For operators, our proof will be an existence proof that an isometry S exists.

2.1 Proof

Let's begin by deriving a useful relationship between the norm of T and the norm of $\sqrt{T^*T}$. Note that Exercise 7.C(4) allows us to state T^*T is a positive operator, and following Theorem 7.36 ('Each positive operator has only one positive square root'), $\sqrt{T^*T}$ is the unique positive operator of T^*T; therefore, $T^*T = \sqrt{T^*T}\sqrt{T^*T}$. Thus, if $v \in V$, then

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

Hence,

$$\|Tv\| = \|\sqrt{T^*T}v\| \tag{2}$$

for all $v \in V$.

Now we can start to construct our isometry S. Define a linear map $S_1 \in \mathcal{L}(\operatorname{range}\sqrt{T^*T}, \operatorname{range} T)$ by

$$S_1(\sqrt{T^*T}v) = Tv. \tag{3}$$

The linear map S_1 can be a bit tricky to dissect at first, so we can think about it as follows:

1. Write $S_1(u)$
2. This implies $u \in \operatorname{range}\sqrt{T^*T}$
3. Therefore, there exists v such that $v = \sqrt{T^*T}u$
4. Thus $S_1(u) = S_1(\sqrt{T^*T}u)$
5. Following from (3), $S_1(\sqrt{T^*T}u) = Tv$ and $S_1(u) = Tv$

We eventually want to show that S_1 is an isometry and can be extended to an isometry on V. But first, we must understand if S_1 is well-defined. By well-defined, we want $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$ to imply that $Tv_1 = Tv_2$. In other words, the same input should lead to the same output. We can use (2) to write

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

Hence, the same inputs imply the same outputs, so S_1 is well-defined. But even if S_1 is well-defined, is it a linear map? Let's test for additivity and homogeneity.

Additivity: Suppose $u, w \in \operatorname{range}\sqrt{T^*T}$. This implies there exist $u_0, w_o \in V$ such that $u = \sqrt{T^*T}u_0$ and $w = \sqrt{T^*T}w_0$. Hence we can write

$$\begin{aligned}S_1(u) + S_1(w) &= S_1(\sqrt{T^*T}u_0) + S_1(\sqrt{T^*T}w_0) \\ &= Tu_0 + Tw_0 \\ &= T(u_0 + w_0) \\ &= S_1((\sqrt{T^*T})(u_0 + w_0)) \\ &= S_1(\sqrt{T^*T}u_0 + \sqrt{T^*T}u_0 w_0) = S_1(u + w)\end{aligned}$$

Homogeneity: Suppose $v \in \operatorname{range}\sqrt{T^*T}$ and $\lambda \in \mathbf{F}$. This implies there exist $v_0 \in V$ such that $v = \sqrt{T^*T}v_0$. Hence we can write

$$\begin{aligned}S_1(\lambda v) &= S_1(\lambda(\sqrt{T^*T})(v_0)) = S_1((\sqrt{T^*T})(\lambda v_0)) \\ &= (T)(\lambda v_0) = \lambda(Tv_0) \\ &= \lambda(S_1(\sqrt{T^*T}v_0)) = \lambda(Sv)\end{aligned}$$

Cool. Now we know S_1 is a well-defined, linear map.

To show that S_1 is an isometry, note that (2) and (3) imply

$$\|S_1 u\| = \|S_1(\sqrt{T^*T}u_0)\| = \|Tu_0\| = \|\sqrt{T^*T}u_0\| = \|u\| \tag{4}$$

for all $u \in \operatorname{range}\sqrt{T^*T}$ where $u = \sqrt{T^*T}u_0$. Hence S_1 is an isometry.

At this point, we know S_1 is a well-defined function, a linear map, and an isometry. Now let's show that S_1 is an isomorphism from $\operatorname{range}\sqrt{T^*T}$ onto $\operatorname{range} T$. Clearly (3) implies that $\operatorname{range} S_1 = \operatorname{range} T$ and S_1 is surjective.

To show that S_1 is injective, suppose there exists $u \in \operatorname{range}\sqrt{T^*T}$ such that $S_1(u) = 0$. This implies there exists $v \in V$ such that $\sqrt{T^*T}v = u$ and $S_1(\sqrt{T^*T}v) = 0$. Hence, it follows that $Tv = 0$. Via (2), we can write

$$0 = \|Tv\| = \|\sqrt{T^*T}v\| = \|u\|.$$

Via the **definiteness** property of inner products, $u = 0$. Therefore, it follows that $\operatorname{nul} S_1 = \{0\}$ and S_1 is injective.

Hence, $\operatorname{range}\sqrt{T^*T}$ and $\operatorname{range} T$ isomorphic, and via Theorem 3.59 ('Dimension shows whether vector spaces are isomorphic'), we can write

$$\dim \operatorname{range}\sqrt{T^*T} = \dim \operatorname{range} T. \tag{5}$$

Now let's think back to Theorem 6.50 ('Dimension of the orthogonal complement') that states $\dim U = \dim V - \dim U^\perp$. Thus, following (5), we can write

$$\begin{aligned}\dim V - \dim(\operatorname{range}\sqrt{T^*T})^\perp &= \dim V - \dim(\operatorname{range} T)^\perp \\ \dim(\operatorname{range}\sqrt{T^*T})^\perp &= \dim(\operatorname{range} T)^\perp\end{aligned}$$

Now we can start to construct an isometry S on V by extending S_1 to $(\operatorname{range}\sqrt{T^*T})^\perp$. Let's start by Theorem 6.34 ('Existence of orthonormal basis') to choose orthonormal bases $e_1 \ldots, e_m$ on $(\operatorname{range}\sqrt{T^*T})^\perp$ and $f_1, \ldots, f_m$ on $(\operatorname{range} T)^\perp$. Via our best friend and colleague Theorem 3.5 ('Linear maps and basis of domain'), we can define a linear map $S_2 \in \mathcal{L}((\operatorname{range}\sqrt{T^*T})^\perp, (\operatorname{range} T)^\perp)$ by

$$S_2(a_1e_1 + \cdots + a_me_m) = a_1f_1 + \cdots + a_mf_m. \tag{6}$$

Via Theorem 6.25 ('The norm of an orthonormal linear combination'), for $w \in ((\operatorname{range}\sqrt{T^*T})^\perp)$ we can write

$$\|w\|^2 = \|a_1e_1 + \cdots + a_me_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

and with (6), it follows that

$$\|S_2w\|^2 = \|a_1f_1 + \cdots + a_mf_m\|^2 = |a_1|^2 + \cdots + |a_m|^2 = \|w\|^2. \tag{7}$$

Hence S_2 is an isometry.

Are you ready? Let's now construct $S \in \mathcal{L}(V)$ such that S equals S_1 on $\operatorname{range}\sqrt{T^*T}$ and equals S_2 on $(\operatorname{range}\sqrt{T^*T})^\perp$. More precisely, via Theorem 6.47 ('Direct sum of a subspace and its orthogonal complement'), we know that $V = \operatorname{range}\sqrt{T^*T} \oplus (\operatorname{range}\sqrt{T^*T})^\perp$; thus, we can uniquely write every $v \in V$ as

$$v = u + w \tag{8}$$

where $u \in \operatorname{range}\sqrt{T^*T}$ and $w \in (\operatorname{range}\sqrt{T^*T})^\perp$. Following our orthogonal decomposition of $v \in V$, let's define Sv by

$$Sv = S_1u + S_2w \tag{9}$$

Alright. You can probably guess that S is an isometry from (4) and (7), but let's be explicit. Following (8), we know that u and w are orthogonal; hence we can use the Pythagorean Theorem to write $\|v\|^2 = \|u\|^2 + \|w\|^2$. With a similar logic, we know that $S_1u \in \operatorname{range} T$ and $S_2w \in (\operatorname{range} T)^\perp)$; hence it follows that $\|Sv\|^2 = \|S_1u\|^2 + \|S_2w\|^2$. Putting it all together, following from (4) and (7), we can write

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

and S is an isometry.

To end this proof, let's show that $T = S(\sqrt{T^*T})$. For every $v \in V$, it follows that

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

This can be initially confusing because we didn't include S_2. However, let's think about the vector $\sqrt{T^*T}v$. Clearly $\sqrt{T^*T}v \in \operatorname{range}\sqrt{T^*T}$; hence our decomposition (8) becomes

$$\sqrt{T^*T}v = \sqrt{T^*T}v + 0.$$

Thus, we can more explicitly use (9) to write the following

$$S(\sqrt{T^*T}v) = S_1(\sqrt{T^*T}v) + S_2(0) = Tv + 0 = Tv$$

for all $v \in V$, which was to be shown.