Polar decomposition

November 25, 2019

Contents

1 Introduction

The polar decomposition proven in the book (Theorem 7.45 on page 233) concerns a linear map $T \in \mathcal{L}(V)$ from a single inner product space to itself. Exactly the same ideas treat that the case of a map

$$T \in \mathcal{L}(V, W) \qquad (V, W \ \text{ fin-diml inner product spaces}). \tag{1.1a}$$

I stated the result in class; the point of these notes is to write down some details of the proof, as well as to talk a bit more about why the result is useful. To get to the statement, we need some notation. I'll think of the linear map T as an arrow going from V to W:

$$V \xrightarrow{T} W; \tag{1.1b}$$

this is just another notation for saying that T is a function that takes anything in $v \in V$ and gives you something $T(v) \in W$. Because these are inner product spaces, we get also the *adjoint of* T

$$W \xrightarrow{T^*} V, \qquad \langle Tv, w \rangle = \langle v, T^* w \rangle \quad (v \in V,\ w \in W). \tag{1.1c}$$

The property written on the right *defines* the linear map T^*. Using these two maps, we immediately get two subspaces of each of V and W:

$$\mathrm{Null}(T),\ \mathrm{Range}(T^*) \subset V, \qquad \mathrm{Null}(T^*),\ \mathrm{Range}(T) \subset W. \tag{1.1d}$$

The first basic fact is that these spaces provide orthogonal direct sum decompositions of V and W.

Proposition 1.2. *In the setting of* (1.1),

1. *The null space of T is the orthogonal complement of the range of T^*:* $\mathrm{Null}(T) = \mathrm{Range}(T^*)^{\perp}$.

2. *The null space of T^* is the orthogonal complement of the range of T:* $\mathrm{Null}(T^*) = \mathrm{Range}(T)^{\perp}$.

3. *The space V is the orthogonal direct sum of* $\mathrm{Null}(T)$ *and* $\mathrm{Range}(T^*)$*:* $V = \mathrm{Null}(T) \oplus \mathrm{Range}(T^*)$.

4. *The space W is the orthogonal direct sum of* $\mathrm{Null}(T^*)$ *and* $\mathrm{Range}(T)$*:* $W = \mathrm{Null}(T^*) \oplus \mathrm{Range}(T)$.

5. *The natural quotient map π from V to $V/\mathrm{Null}(T)$ (text, 3.88) restricts to an isomorphism of vector spaces*
$$\mathrm{Range}(T^*) \xrightarrow{\simeq} V/\mathrm{Null}(T).$$

6. *The natural isomorphism $V/\mathrm{Null}(T) \xrightarrow{\widetilde{T}} \mathrm{Range}(T)$ (text, 3.91) restricts to an isomorphism of vector spaces*
$$\mathrm{Range}(T^*) \xrightarrow{T} \mathrm{Range}(T).$$

7. *The natural quotient map π from W to $W/\mathrm{Null}(T^*)$ (see text, 3.88) restricts to an isomorphism of vector spaces*
$$\mathrm{Range}(T) \xrightarrow{\simeq} W/\mathrm{Null}(T^*).$$

8. *The natural isomorphism $W/\mathrm{Null}(T^*) \xrightarrow{\widetilde{T^*}} \mathrm{Range}(T^*)$ (see text, 3.91) restricts to an isomorphism of vector spaces*
$$\mathrm{Range}(T) \xrightarrow{T^*} \mathrm{Range}(T^*).$$

I'll outline proofs in Section 2. (Even better, you should try to write down proofs yourself.)

Here is the polar decomposition I stated in class.

Theorem 1.3. *In the setting of* (1.1)*, there is a factorization*

$$T = SP \qquad (S \in \mathcal{L}(V,W),\ P \in \mathcal{L}(V))$$

characterized uniquely by the following properties:

1. *P is a positive self-adjoint operator;*

2. *S restricts to an isometry from* $\mathrm{Range}(T^*)$ *onto* $\mathrm{Range}(T)$*; and*

3. $P|_{\mathrm{Null}(T)} = S|_{\mathrm{Null}(T)} = 0$.

In addition, the decomposition has the following properties.

4. *The operator P is the (unique) positive square root of T^*T.*

Write λ_{max} and λ_{min} and for the largest and smallest eigenvalues of P (which are the square roots of the corresponding eigenvalues of T^*T). Then

5. $$\|Tv\|/\|v\| \le \lambda_{max} \qquad (0 \ne v \in V)$$

   with the maximum attained exactly on the λ_{max} eigenspace of P.

6. $$\|Tv\|/\|v\| \ge \lambda_{min} \qquad (0 \ne v \in V)$$

   with the minimum attained exactly on the λ_{min} eigenspace of P.

If $\dim(V) = \dim(W)$*, then there is another factorization (not unique)*

$$T = S'P$$

in which conditions (3) above is replaced by

3.′ *S' restricts to an isometry (which may be arbitrary) from* $\mathrm{Null}(T)$ *onto* $\mathrm{Null}(T^*)$.

It is equivalent to require that

3.″ *S' is an isometry from V onto W.*

It is the factorization $T = S'P$ that is established in the text (when $V = W$). Even in that case, I like $T = SP$ better because it's unique, and unique is (almost) always better. The two maps S and S' differ only on $\mathrm{Null}(T)$; so if T is invertible, $S = S'$.

Items (5) and (6) correspond to an application that I discussed in class: often one cares about controlling how a linear map can change the *sizes* of vectors. This result answers that question very precisely. I'll discuss this in Section 3.

The proof will be outlined in Section 4. This is harder than Proposition 1.2; but it's still very worthwhile first to try yourself to write proofs.

2 Orthogonal decompositions

This section is devoted to the proof of Proposition 1.2.

Proof of Proposition 1.2. The idea that's used again and again is that in an inner product space V

$$v = 0 \iff \langle v, v' \rangle = 0 \qquad (\text{all } v' \in V). \tag{2.1}$$

We use this idea to prove 1.

$$\begin{aligned}\mathrm{Null}(T) &= \{v \in V \mid Tv = 0\} \qquad (\text{definition})\\ &= \{v \in V \mid \langle Tv, w\rangle = 0 \quad (\text{all } w \in W)\} \qquad (\text{by (2.1)})\\ &= \{v \in V \mid \langle v, T^*w\rangle = 0 \quad (\text{all } w \in W)\} \qquad (\text{by (1.1c)})\\ &= \mathrm{Range}(T^*)^\perp \qquad (\text{definition of } U^\perp).\end{aligned}$$

This proves 1.

Part 2 is just 1 applied to T^*, together with the (very easy; you should write a proof!) $T^{**} = T$.

Part 3 follows from 1 and Theorem 6.47 in the text; and part 4 is 3 applied to T^*.

For part 5, suppose $V = V_1 \oplus V_2$ is any direct sum decomposition of any finite-dimensional V over any field F; the claim is that the natural quotient map

$$\pi \colon V \to V/V_1$$

restricts to an isomorphism

$$V_2 \xrightarrow{\simeq} V/V_1.$$

To see this, notice first that

$$\dim V_2 = \dim V - \dim V_1 = \dim(V/V_1);$$

(Theorem 3.78 on page 93 and 3.89 on page 97)so the two vector spaces V and V/V_1 have the same dimension. Second, $\text{Null}(\pi) = V_1$ (text, proof of Theorem 3.89), so

$$\text{Null}(V_2 \to V/V_1) = V_2 \cap V_1 = 0$$

(text, 1.45). So our map is an *injection* between vector spaces of the same dimension; so it is an isomorphism.

For 6, the map we are looking at is the composition of the isomorphism from 5 and the isomorphism from Proposition 3.91 in the text; so it is also an isomorphism.

Parts 7 and 8 are just 5 and 6 applied to T^*. □

3 Length of Tv

This section concerns the general question "how big is Tv?" whenever this question makes sense: for us, that's the setting (1.1a). I'll start with a slightly different result, proved in class.

Proposition 3.1. *Suppose $S \in \mathcal{L}(V)$ is a self-adjoint linear operator. Define*

$$\mu_{min} = \textit{smallest eigenvalue of } S$$

$$\mu_{max} = \textit{largest eigenvalue of } S$$

Then for all nonzero $v \in V$,

$$\mu_{min} \ \leq \ \langle Sv, v\rangle/\langle v, v\rangle \ \leq \ \mu_{max}.$$

All values in this range are attained. The first inequality is an equality if and only if v is an eigenvector for μ_{min}; that is, if and only if $v \in V_{\mu_{min}}$. The second inequality is an equality if and only if v is an eigenvector for μ_{max}; that is, if and only if $v \in V_{\mu_{max}}$.

Proof. According to the Spectral Theorem for self-adjoint operators (which is in the text, but not so easy to point to; there is a simple statement in the notes on the spectral theorem on the class web site) the eigenvalues of S are all real, so they can be arranged as

$$\mu_{\min} = \mu_1 < \mu_2 < \cdots < \mu_{r-1} < \mu_r = \mu_{\max}. \tag{3.2a}$$

Furthermore V is the orthogonal direct sum of the eigenspaces:

$$V = V_{\mu_1} \oplus \cdots \oplus V_{\mu_r}. \tag{3.2b}$$

If we choose an orthonormal basis $(e_i^1, \ldots, e_i^{n_i})$ of each eigenspace V_{μ_i} (with $n_i = \dim V_{\mu_i}$), then we get an orthonormal basis (e_j^i) of V, with i running from 1 to r and j running from 1 to n_i. (The total size of the orthonormal basis is therefore

$$n_1 + \cdots + n_r = n = \dim V.) \tag{3.2c}$$

Once we have chosen this basis, we can write any vector in V as

$$v = \sum_{i=1}^{r} \sum_{j=1}^{n_i} v_i^j e_i^j \qquad (v_i^j \in F). \tag{3.2d}$$

Because e_i^j belongs to the μ_i eigenspace V_{μ_i}, we find

$$Sv = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \mu_i v_i^j e_i^j, \tag{3.2e}$$

So far this calculation would have worked for any *diagonalizable* S. Now we use the fact that S is self-adjoint, and therefore that we could choose the e_i^j to be *orthonormal*. This allows us to calculate

$$\begin{aligned}
\langle \mu_i v_i^j e_i^j, v_{i'}^{j'} e_{i'}^{k'} \rangle &= \mu_i v_i^j \overline{v_{i'}^{j'}} \langle e_i^j, e_{i'}^{j'} \rangle \\
&= \begin{cases} (\mu_i v_i^j \overline{v_{i'}^{j'}}) \cdot 1 & (i = i',\ j = j') \\ (\mu_i v_i^j \overline{v_{i'}^{j'}}) \cdot 0 & i \neq i' \text{ or } j \neq j') \end{cases} \\
&= \begin{cases} \mu_i |v_i^j|^2 & (i = i',\ j = j') \\ 0 & i \neq i' \text{ or } j \neq j'). \end{cases}
\end{aligned}$$

Using this calculation to expand $\langle Sv, v \rangle$, we find

$$\langle Sv, v \rangle / \langle v, v \rangle = \left(\sum_{i=1}^{r} \mu_i (\sum_{j=1}^{n_r} |v_i^j|^2 \right) / \left(\sum_{j=1}^{n_r} \rangle |v_i^j|^2 \right). \tag{3.2f}$$

Here's how you take a *weighted average* of the real numbers $\mu_1, \ldots \mu_r$. Fix r real numbers $w_1, \ldots, w_r$ such that

$$0 \leq w_i \leq 1, \qquad w_1 + \cdots + w_r = 1 \tag{3.2g}$$

These numbers are the *weights*. Simplest weights are the *uniform weight* $(1/r, \cdots, 1/r)$. The *weighted average* is

$$\overline{\mu} = w_1\mu_1 + \cdots + w_r\mu_r. \tag{3.2h}$$

If the weight is uniform, then the weighted average is

$$\overline{\mu} = (\mu_1 + \cdots + \mu_r)/r,$$

the ordinary average. The opposite extreme is the *teacher's pet weight*

$$w_i = \begin{cases} 1 & (i = p) \\ 0 & (i \neq p) \end{cases}$$

where all the weight is on a single value μ_p. The weighted average for the teacher's pet weight is

$$\overline{\mu} = \mu_p.$$

No matter what weights you use, it is always true that

$$\mu_{\min} \leq \overline{\mu} \leq \mu_{\max}. \tag{3.2i}$$

The first inequality is an equality if and only if the weights are concentrated on the minimum values:

$$w_i \neq 0 \iff \mu_i = \mu_{\min},$$

and similarly for the second inequality. I won't write out a proof of (3.2i), but here's how to start. For each i, the definitions of min and max say that

$$\mu_{\min} \leq \mu_i \leq \mu_{\max}.$$

We can multiply inequalities by a nonnegative real number like w_i, getting r inequalities

$$w_i\mu_{\min} \leq w_i\mu_i \leq w_i\mu_{\max}.$$

Now add up these r inequalities to get (3.2i).

Each nonzero vector v defines a set of nonzero weights

$$w_{i_0} = \left(\sum_{j=1}^{n_r} |v_{i_0}^j|^2\right) \Big/ \left(\sum_{i=1}^{r}\sum_{j=1}^{n_r} |v_i^j|^2\right). \tag{3.2j}$$

You should convince yourself that this really is a set of weights (that they are non-negative real numbers adding up to 1). Now (3.2f) says that

$$\langle Sv, v\rangle / \langle v, v\rangle = \sum_{i=1}^{r} w_i \mu_i = \text{ weighted average of eigenvalues of } S. \quad (3.2\text{k})$$

Now (3.2i) gives the inequalities in the proposition. You should also convince yourself that the conditions for equality given for weighted averages after (3.2i) lead exactly to the conditions for equality stated in the proposition. □

Here's an exercise in thinking about careful mathematical formulations. Suppose V has dimension zero. In this case there are *no* eigenvalues, so the largest and smallest eigenvalues are not defined. Is the formulation of Proposition 3.1 wrong in this case?

Now we can talk about the length of Tv.

Proposition 3.3. *Suppose we are in the setting* (1.1). *Write* λ_{max}^2 *and* λ_{min}^2 *and for the largest and smallest eigenvalues of* T^*T*, with* $0 \leq \lambda_{min} \leq \lambda_{max}$*. Then*

1. $$\|Tv\|/\|v\| \leq \lambda_{max} \qquad (0 \neq v \in V)$$

   with the maximum attained exactly on the λ_{max}^2 *eigenspace of* T^*T*.*

2. $$\|Tv\|/\|v\| \geq \lambda_{min} \qquad (0 \neq v \in V)$$

   with the minimum attained exactly on the λ_{min}^2 *eigenspace of* T^*T*.*

3. *Suppose* $P \in \mathcal{L}(V, V)$ *is the nonnegative self-adjoint square root of* T^*T*. Then*

   $$\|Tv\| = \|Pv\| \qquad (v \in V).$$

4. *Suppose* $R \in \mathcal{L}(V, U)$ *is any linear map such that* $R^*R = T^*T$*. Then*

   $$\|Tv\| = \|Rv\| \qquad (v \in V).$$

Proof. Because the inequalities in the proposition involve non-negative real numbers, they are equivalent to the squared versions:

$$\lambda_{\text{min}}^2 \leq \langle Tv, Tv\rangle / \langle v, v\rangle \leq \lambda_{\text{max}}^2.$$

Using the definition of adjoint, these can be written

$$\lambda_{\min}^2 \le \langle T^*Tv, v\rangle / \langle v, v\rangle \le \lambda_{\max}^2.$$

In this form it is precisely Proposition 3.1 applied to the self-adjoint operator T^*T.

The proof of the inequalities was based on the formula

$$\langle Tv, Tv\rangle = \langle T^*Tv, v\rangle > .$$

A consequence is that the length of Tv depends only on T^*T. For 3, we have

$$P^*P = P^2 = T^*T$$

by the definition of P; this proves 3. The same argument proves 4. □

4 Proof of the polar decomposition

Proof of Theorem 1.3. We address uniqueness first. So suppose that we have two factorizations $T = SP$ and $T = \widetilde{S}\widetilde{P}$ satisfying the properties 1, 2, and 3; we must show that $S = \widetilde{S}$ and $P = \widetilde{P}$. Write S_1 and $\widetilde{S_1}$ for the restrictions to $\mathrm{Range}(T^*)$. Because of hypothesis Theorem 1.3(2),

$$S_1^* S_1 = I_{\mathrm{Range}(T^*)}, \qquad \widetilde{S_1^*}\widetilde{S_1} = I_{\mathrm{Range}(T^*)}. \tag{4.1a}$$

(Here I_U means the identity operator on the vector space U.) Consequently

$$S^*S|_{\mathrm{Range}(T^*)} = \widetilde{S}^*\widetilde{S}|_{\mathrm{Range}(T^*)} = I_{\mathrm{Range}(T^*)}. \tag{4.1b}$$

Meanwhile condition 3 of the theorem guarantees that

$$S^*S|_{\mathrm{Null}(T)} = \widetilde{S}^*\widetilde{S}|_{\mathrm{Null}(T)} = 0_{\mathrm{Null}(T)}. \tag{4.1c}$$

Because of the direct sum decomposition Proposition 1.2(1), it follows that

$$S^*S = \widetilde{S}^*\widetilde{S} = P_{\mathrm{Range}(T^*)}, \tag{4.1d}$$

the orthogonal projection on the range of T^*.

Now the factorization $T = SP$ and the assumption that P is self-adjoint means that $T^* = PS^*$. Therefore

$$T^*T = PS^*SP = P(P_{\mathrm{Range}(T^*)})P. \tag{4.1e}$$

To continue, we need to know that

$$P_{\mathrm{Range}(T^*)}P = PP_{\mathrm{Range}(T^*)}. \tag{4.1f}$$

We already know from Theorem 1.3(3) that

$$\mathrm{Null}(T) \subset \mathrm{Null}(P). \tag{4.1g}$$

Write $V_\mu(P)$ for the μ eigenspace of P on V. Because the eigenspaces of the self-adjoint P must be orthogonal, it follows that for any nonzero μ,

$$V_\mu(P) \subset \mathrm{Null}(P)^\perp \subset \mathrm{Null}(T)^\perp = \mathrm{Range}(T^*). \tag{4.1h}$$

Similarly,

$$\mathrm{Null}(P) = V_0(P) = \mathrm{Null}(T) \oplus (V_0(P) \cap \mathrm{Range}(T^*), \tag{4.1i}$$

an orthogonal direct sum decomposition. That is, each eigenspace of P is the orthogonal direct sum of its intersections with $\mathrm{Null}(T)$ and $\mathrm{Range}(T^*)$. From this fact (4.1f) follows. We also find

$$\mathrm{Range}(P) \subset \mathrm{Range}(T^*) \tag{4.1j}$$

Now we can plug (4.1f) into (4.1e) to get

$$T^*T = P_{\mathrm{Range}(T^*)}P^2. \tag{4.1k}$$

Taking into account (4.1j), this becomes

$$T^*T = P^2. \tag{4.1l}$$

Now Theorem 1.3(4) follows immediately; and in particular this shows that

$$P = \widetilde{P}. \tag{4.1m}$$

We also know from Proposition 1.2(6) and (8) that T^*T is an *isomorphism* on $\mathrm{Range}(T^*)$. Therefore (using (4.1l) again) $\mathrm{Null}(P) \cap \mathrm{Range}(T^*) = 0$, so $P|_{\mathrm{Range}(T^*)}$ is invertible. Therefore the factorization gives

$$S_1 = T(P|_{\mathrm{Range}(T^*)})^{-1} = \widetilde{S_1}. \tag{4.1n}$$

Since S and $\widetilde{S}$ both vanish on $\mathrm{Null}(T)$, it follows that

$$S = \widetilde{S}, \tag{4.1o}$$

completing the uniqueness proof for the decomposition. (We also proved (4) in the process.

For the existence of the decomposition, we define $P = (T^*T)^{1/2}$, the unique positive square root, as we must. By Proposition 3.3(3)

$$\|Tv\| = \|Pv\| \qquad (v \in V). \tag{4.1p}$$

This implies immediately that $\mathrm{Null}(T) = \mathrm{Null}(P)$, so

$$\mathrm{Null}(P) \cap \mathrm{Range}(T^*) = \mathrm{Null}(P) \cap (\mathrm{Null}(T)^\perp) = 0. \tag{4.1q}$$

Therefore

$$P_1 =_{\mathrm{def}} P|_{\mathrm{Range}(T^*)} \in \mathcal{L}(\mathrm{Range}(T^*) \tag{4.1r}$$

is *invertible*; and P_1 inherits from P the property

$$\|P_1 v\| = \|Tv\| \qquad (v \in \mathrm{Range}(T^*) = \mathrm{Null}(T)^\perp). \tag{4.1s}$$

Now the invertibility of P_1 allows us to define

$$S_1 = TP_1^{-1} \in \mathcal{L}(\mathrm{Range}(T^*), W), \tag{4.1t}$$

and (4.1s) (applied to $P_1 v$) says that

$$\|S_1 v\| = \|v\| \qquad (v \in \mathrm{Range}(T^*); \tag{4.1u}$$

this is Theorem 1.3(2). Now the orthogonal decomposition

$$V = \mathrm{Null}(T) \oplus \mathrm{Range}(T^*)$$

of Proposition 1.2(3) allows us to define $S \in \mathcal{L}(V, W)$ by

$$S(n + v') = S_1(v') \qquad (n \in \mathrm{Null}(T),\ v' \in \mathrm{Range}(T^*)) \tag{4.1v}$$

Then Theorem 1.3(2) and (3) are true by this definition and (4.1u). Furthermore

$$\begin{aligned} T(n + v') &= (T|_{\mathrm{Range}(T^*)})(v') \\ &= (T|_{\mathrm{Range}(T^*)})P_1^{-1}P_1 v' \\ &= S_1 P_1 v' \\ &= SP(n + v'). \end{aligned} \tag{4.1w}$$

This proves the factorization $T = SP$.

Parts (5) and (6) of Theorem 1.3 are contained in Proposition 3.3.

For the last assertion (about an alternate factorization), it's clear from the preceding proof that we can achieve the factorization using any S' which agrees with S on $\mathrm{Range}(P) = \mathrm{Null}(T)^{\perp}$. To complete the definition of S', we just need to define it (as any linear map to W) on $\mathrm{Null}(T)$. The hypothesis $\dim V = \dim W$ and Proposition 1.2 guarantee that $\dim \mathrm{Null}(T) = \dim \mathrm{Null}(T^*)$; so we can choose S' on $\mathrm{Null}(T)$ to be an isometry to $\mathrm{Null}(T^*)$. I'll omit the rest of the details. $\square$