The Spectral Theorem for normal linear maps

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling
(March 14, 2007)

In this section we come back to the question when an operator on an inner product space V is diagonalizable. We first introduce the notion of the adjoint or hermitian conjugate of an operator and use this to define normal operators, which are those for which the operator and its adjoint commute with each other. The main result of this section is the Spectral Theorem which states that normal operators are diagonal with respect to an orthonormal basis. We use this to show that normal operators are "unitarily diagonalizable" and generalize this notion to find the singular-value decomposition of an operator.

1 Self-adjoint or hermitian operators

Let V be a finite-dimensional inner product space over $\mathbb{C}$ with inner product $\langle \cdot, \cdot \rangle$. A linear operator $T \in \mathcal{L}(V)$ is uniquely determined by the values of

$$\langle Tv, w \rangle \quad \text{for all } v, w \in V.$$

This means in particular that if $T, S \in \mathcal{L}(V)$ and

$$\langle Tv, w \rangle = \langle Sv, w \rangle \quad \text{for all } v, w \in V,$$

then $T = S$. To see this take w for example to be the elements of an orthonormal basis of V.

Definition 1. Given $T \in \mathcal{L}(V)$, the **adjoint** (sometimes hermitian conjugate) of T is the operator $T^* \in \mathcal{L}(V)$ such that

$$\langle Tv, w \rangle = \langle v, T^* w \rangle \quad \text{for all } v, w \in V$$

Moreover, we call T **self-adjoint** or **hermitian** if $T = T^*$.

The uniqueness of T^* is clear by the previous observation.

Example 1. Let $V = \mathbb{C}^3$ and let $T \in \mathcal{L}(\mathbb{C}^3)$ be defined as $T(z_1, z_2, z_3) = (2z_2 + iz_3, iz_1, z_2)$. Then

$$\begin{aligned}
\langle (y_1, y_2, y_3), T^*(z_1, z_2, z_3) \rangle &= \langle T(y_1, y_2, y_3), (z_1, z_2, z_3) \rangle \\
&= \langle (2y_2 + iy_3, iy_1, y_2), (z_1, z_2, z_3) \rangle \\
&= 2y_2\overline{z_1} + iy_3\overline{z_1} + iy_1\overline{z_2} + y_2\overline{z_3} \\
&= \langle (y_1, y_2, y_3), (-iz_2, 2z_1 + z_3, -iz_1) \rangle,
\end{aligned}$$

so that $T^*(z_1, z_2, z_3) = (-iz_2, 2z_1 + z_3, -iz_1)$. Writing the matrix for T in terms of the canonical basis we see that

$$M(T) = \begin{bmatrix} 0 & 2 & i \\ i & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad M(T^*) = \begin{bmatrix} 0 & -i & 0 \\ 2 & 0 & 1 \\ -i & 0 & 0 \end{bmatrix}.$$

Note that $M(T^*)$ can be obtained from $M(T)$ by taking the complex conjugate of each element and transposing.

Elementary properties that you should prove as exercises are that for all $S, T \in \mathcal{L}(V)$ and $a \in \mathbb{F}$ we have

$$\begin{aligned}
(S + T)^* &= S^* + T^* \\
(aT)^* &= \overline{a}T^* \\
(T^*)^* &= T \\
I^* &= I \\
(ST)^* &= T^*S^* \\
M(T^*) &= M(T)^*
\end{aligned}$$

where $A^* = (\overline{a_{ji}})_{i,j=1}^n$ if $A = (a_{ij})_{i,j=1}^n$. The matrix A^* is the **conjugate transpose** of A.

For $n = 1$ the conjugate transpose of the 1×1 matrix A is just the complex conjugate of its element. Hence requiring A to be self-adjoint ($A = A^*$) amounts to saying that the entry of A is real. Because of the transpose, reality is not the same as self-adjointness, but the analogy does carry over to the eigenvalues of self-adjoint operators as the next Proposition shows.

Proposition 1. *Every eigenvalue of a self-adjoint operator is real.*

Proof. Suppose $\lambda \in \mathbb{C}$ is an eigenvalue of T and $0 \neq v \in V$ the corresponding eigenvector such that $Tv = \lambda v$. Then

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

This implies that $\lambda = \overline{\lambda}$. $\square$

Example 2. The operator $T \in \mathcal{L}(V)$ be defined by $T(v) = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix} v$ is self-adjoint (or hermitian) and it can be checked that the eigenvalues are $\lambda = 1, 4$ by determining the zeroes of the polynomial $p(\lambda) = (2-\lambda)(3-\lambda) - (1+i)(1-i) = \lambda^2 - 5\lambda + 4$.

2 Normal operators

Normal operator are those which commute with their adjoint.

Definition 2. We call $T \in \mathcal{L}(V)$ **normal** if $TT^* = T^*T$.

In general $TT^* \neq T^*T$. Note that TT^* and T^*T are both self-adjoint for all $T \in \mathcal{L}(V)$. Also, any self-adjoint operator T is normal. We now give a different characterization for normal operators in terms of norms. In order to prove this results, we first discuss the next proposition.

Proposition 2. *Let V be a complex inner product space and $T \in \mathcal{L}(V)$ such that*

$$\langle Tv, v \rangle = 0 \quad \text{for all } v \in V.$$

Then $T = 0$.

Proof. Verify that

$$\begin{aligned} \langle Tu, w \rangle = \frac{1}{4} \{ &\langle T(u+w), u+w \rangle - \langle T(u-w), u-w \rangle \\ &+ i \langle T(u+iw), u+iw \rangle - i \langle T(u-iw), u-iw \rangle \}. \end{aligned}$$

Since each term on the right is of the form $\langle Tv, v \rangle$, we obtain 0 for all $u, w \in V$. Hence $T = 0$. $\square$

Proposition 3. *Let $T \in \mathcal{L}(V)$. Then T is normal if and only if*

$$\|Tv\| = \|T^*v\| \quad \text{for all } v \in V.$$

Proof. Note that

$$\begin{aligned} T \text{ is normal} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)v, v \rangle = 0 \quad \text{for all } v \in V \\ &\iff \langle TT^*v, v \rangle = \langle T^*Tv, v \rangle \quad \text{for all } v \in V \\ &\iff \|Tv\|^2 = \|T^*v\|^2 \quad \text{for all } v \in V. \end{aligned}$$

□

Corollary 4. *Let $T \in \mathcal{L}(V)$ be a normal operator. Then*

1. $\operatorname{null} T = \operatorname{null} T^*$.
2. *If $\lambda \in \mathbb{C}$ is an eigenvalue of T, then $\overline{\lambda}$ is an eigenvalue of T^* with the same eigenvector.*
3. *If $\lambda, \mu \in \mathbb{C}$ are distinct eigenvalues of T with associated eigenvectors $v, w \in V$ respectively, then $\langle v, w\rangle = 0$.*

Proof. Note that 1 follows from Proposition 3 and the positive definiteness of the norm.

To prove 2, first verify that if T is normal, then $T - \lambda I$ is also normal and $(T - \lambda I)^* = T^* - \overline{\lambda} I$. Therefore by Proposition 3 we have

$$0 = \|(T - \lambda I)v\| = \|(T - \lambda I)^* v\| = \|(T^* - \overline{\lambda} I)v\|,$$

so that v is an eigenvector of T^* with eigenvalue $\overline{\lambda}$.

Using part 2, note that

$$(\lambda - \mu)\langle v, w\rangle = \langle \lambda v, w\rangle - \langle v, \overline{\mu} w\rangle = \langle Tv, w\rangle - \langle v, T^* w\rangle = 0.$$

Since $\lambda - \mu \neq 0$ it follows that $\langle v, w\rangle = 0$, proving part 3. □

3 Normal operators and the spectral decomposition

Recall that an operator is diagonalizable if there exists a basis of V consisting of eigenvectors of V. The nicest operators on V are those that are diagonalizable with respect to an orthonormal basis of V. These are the operators such that there is an orthonormal basis consisting of eigenvectors of V. The spectral theorem for complex inner product spaces shows that these are precisely the normal operators.

Theorem 5 (Spectral Theorem)**.** *Let V be a finite-dimensional inner product space over $\mathbb{C}$ and $T \in \mathcal{L}(V)$. Then T is normal if and only if there exists an orthonormal basis for V consisting of eigenvectors for T.*

Proof.
"$\Longrightarrow$" Suppose that T is normal. We proved before that for any operator T on a complex inner product space V of dimension n, there exists an orthonormal basis $e = (e_1, \ldots, e_n)$ for which the matrix $M(T)$ is upper-triangular

$$M(T) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{bmatrix}.$$

We will show that $M(T)$ is in fact diagonal, which implies that $e_1, \ldots, e_n$ are eigenvectors of T. Since $M(T) = (a_{ij})_{i,j=1}^n$ with $a_{ij} = 0$ for $i > j$, we have $Te_1 = a_{11}e_1$ and $T^*e_1 = \sum_{k=1}^n \overline{a}_{1k}e_k$. Thus, by the Pythagorean Theorem and Proposition 3

$$|a_{11}|^2 = \|a_{11}e_1\|^2 = \|Te_1\|^2 = \|T^*e_1\|^2 = \|\sum_{k=1}^n \overline{a}_{1k}e_k\|^2 = \sum_{k=1}^n |a_{1k}|^2$$

from which follows that $|a_{12}| = \cdots = |a_{1n}| = 0$. One can repeat this argument, calculating $\|Te_j\|^2 = |a_{jj}|^2$ and $\|T^*e_j\|^2 = \sum_{k=j}^n |a_{jk}|^2$ to find $a_{ij} = 0$ for all $2 \leq i < j \leq n$. Hence T is diagonal with respect to the basis e and $e_1, \ldots, e_n$ are eigenvectors of T.
"$\Longleftarrow$" Suppose there exists an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors for T. Then the matrix $M(T)$ with respect to this basis is diagonal. Moreover, $M(T^*) = M(T)^*$ with respect to this basis must also be a diagonal matrix. Any two diagonal matrices commute. It follows that $TT^* = T^*T$ since their corresponding matrices commute

$$M(TT^*) = M(T)M(T^*) = M(T^*)M(T) = M(T^*T).$$

□

The next corollary is the best possible decomposition of the complex vector space V into subspaces invariant under a normal operator T. On each subspace null $(T - \lambda_i I)$ the operator T just acts by multiplication by λ_i.

Corollary 6. *Let $T \in \mathcal{L}(V)$ be a normal operator. Then*

1. *Denoting $\lambda_1, \ldots, \lambda_m$ the distinct eigenvalues for T,*

$$V = \operatorname{null}(T - \lambda_1 I) \oplus \cdots \oplus \operatorname{null}(T - \lambda_m I).$$

2. *If $i \neq j$, then* $\operatorname{null}(T - \lambda_i I) \perp \operatorname{null}(T - \lambda_j I)$.

As we will see next, the canonical matrix for T admits a "unitary diagonalization".

4 Applications of the spectral theorem: Diagonalization

We already discussed that if $e = (e_1, \ldots, e_n)$ is a basis of a vector space V of dimension n and $T \in \mathcal{L}(V)$, then we can associate a matrix $M(T)$ to T. To remember the dependency on the basis e let us now denote this matrix by $[T]_e$. That is

$$[Tv]_e = [T]_e[v]_e \qquad \text{for all } v \in V$$

where

$$[v]_e = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

is the coordinate vector for $v = v_1 e_1 + \cdots + v_n e_n$ with $v_i \in \mathbb{F}$.

The operator T is diagonalizable if there exists a basis e such that $[T]_e$ is diagonal, that is, there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that

$$[T]_e = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

The scalars $\lambda_1, \ldots, \lambda_n$ are necessarily eigenvalues of T and $e_1, \ldots, e_n$ are the corresponding eigenvectors. Therefore:

Proposition 7. *$T \in \mathcal{L}(V)$ is diagonalizable if and only if there exists a basis $(e_1, \ldots, e_n)$ consisting entirely of eigenvectors of T.*

We can reformulate this proposition as follows using the change of basis transformations. Suppose that e and f are bases of V such that $[T]_e$ is diagonal and let S be the change of basis transformation such that $[v]_e = S[v]_f$. Then $S[T]_f S^{-1} = [T]_e$ is diagonal.

Proposition 8. *$T \in \mathcal{L}(V)$ is diagonalizable if and only if there exists a invertible matrix $S \in \mathbb{F}^{n\times n}$ such that*

$$S[T]_f S^{-1} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

where $[T]_f$ is the matrix of T with respect to a given arbitrary basis $f = (f_1, \ldots, f_n)$.

On the other hand, the spectral theorem tells us that T is diagonalizable with respect to an orthonormal basis if and only if T is normal. Recall that

$$[T^*]_f = [T]_f^*$$

for any orthonormal basis f of V. Here

$$A^* = (\overline{a}_{ji})_{ij=1}^n \qquad \text{for } A = (a_{ij})_{i,j=1}^n$$

is the complex conjugate transpose of the matrix A. When $\mathbb{F} = \mathbb{R}$ then $A^* = A^t$ is just the transpose of the matrix, where $A^t = (a_{ji})_{i,j=1}^n$.

The change of basis transformation between two orthonormal bases is called **unitary** in the complex case or **orthogonal** in the real case. Let $e = (e_1, \ldots, e_n)$ and $f = (f_1, \ldots, f_n)$

be two orthonormal bases of V and U the change of basis matrix $[v]_f = U[v]_e$ for all $v \in V$. Then

$$\langle e_i, e_j\rangle = \delta_{ij} = \langle f_i, f_j\rangle = \langle Ue_i, Ue_j\rangle.$$

Since this is true on the basis e we in fact have that U is unitary if and only if

$$\langle Uv, Uw\rangle = \langle v, w\rangle \quad \text{for all } v, w \in V. \tag{1}$$

This means that unitary matrices preserve the inner product. Operators which preserve the inner product are also called **isometries**. Similar conditions hold for orthogonal matrices.

Since by the definition of the adjoint $\langle Uv, Uw\rangle = \langle v, U^*Uw\rangle$, equation 1 also shows that unitary matrices are characterized by the property

$$\begin{aligned} U^*U &= I && \text{for the unitary case} \\ O^tO &= I && \text{for the orthogonal case.} \end{aligned}$$

The equation $U^*U = I$ implies that $U^{-1} = U^*$. For finite-dimensional inner product spaces V the left inverse of an operator is also the right inverse, so that

$$\begin{aligned} UU^* = I \quad &\text{if and only if} \quad U^*U = I \\ OO^t = I \quad &\text{if and only if} \quad O^tO = I. \end{aligned} \tag{2}$$

It is easy to see that the columns of a unitary matrix are the coefficients of the elements of an orthonormal basis with respect to another orthonormal basis. Therefore the columns are orthonormal vectors in $\mathbb{C}^n$ (or in $\mathbb{R}^n$ in the real case). By (2) this is also true for the rows of the matrix.

The spectral theorem shows that T is normal if and only if $[T]_e$ is diagonal with respect to an orthonormal basis e, that is, there exists a unitary matrix U such that

$$UTU^* = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Conversely, if a unitary matrix U exists such that $UTU^* = D$ is diagonal, then

$$TT^* - T^*T = U^*(D\overline{D} - \overline{D}D)U = 0$$

since diagonal matrices commute, and hence T is normal.

Let us summarize all definitions so far.

Definition 3.

A is **hermitian** if $A^* = A$.
A is **symmetric** if $A^t = A$.
U is **unitary** if $UU^* = I$.
O is **orthogonal** if $OO^t = I$.

Note that all cases of Definition 3 are examples of normal operators. An example of a normal operator N that is none of the above is

$$N = i \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}.$$

You can easily verify that $NN^* = N^*N$. Note that iN is symmetric.

Example 3. Take the matrix

$$A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$$

of Example 2. To unitarily diagonalize A, we need to find a unitary matrix U and a diagonal matrix D such that $A = UDU^{-1}$. To do this, we want to change basis to one composed of orthonormal eigenvectors for $T \in \mathcal{L}(\mathbb{C}^2)$ defined by $Tv = Av$ for all $v \in \mathbb{C}^2$.

To find such an orthonormal basis, we start by finding the eigenspaces of T. We already determined that the eigenvalues of T are $\lambda_1 = 1$ and $\lambda_2 = 4$, so that $D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Hence

$$\begin{aligned} \mathbb{C}^2 &= \operatorname{null}(T - I) \oplus \operatorname{null}(T - 4I) \\ &= \operatorname{span}((-1-i, 1)) \oplus \operatorname{span}((1+i, 2)). \end{aligned}$$

Now apply the Gram-Schmidt procedure to each eigenspace to obtain the columns of U. Here

$$\begin{aligned} A = UDU^{-1} &= \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{-1+i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}. \end{aligned}$$

Note that the diagonal decomposition allows us to compute powers and the exponential of matrices. Namely if $A = UDU^{-1}$ where D is diagonal, we have

$$A^n = (UDU^{-1})^n = UD^nU^{-1}$$
$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = U \left(\sum_{k=0}^{\infty} \frac{1}{k!} D^k \right) U^{-1} = U \exp(D) U^{-1}.$$

Example 4. Continuing the previous Example

$$A^2 = (UDU^{-1})^2 = UD^2U^{-1} = U \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix} U^* = \begin{bmatrix} 6 & 5+5i \\ 5-5i & 11 \end{bmatrix}$$

$$A^n = (UDU^{-1})^n = UD^nU^{-1} = U \begin{bmatrix} 1 & 0 \\ 0 & 2^{2n} \end{bmatrix} U^* = \begin{bmatrix} \frac{2}{3}(1+2^{n-1}) & \frac{1+i}{3}(-1+2^{2n}) \\ \frac{1-i}{3}(-1+2^{2n}) & \frac{1}{3}(1+2^{2n+1}) \end{bmatrix}$$

$$\exp(A) = U \exp(D) U^{-1} = U \begin{bmatrix} e & 0 \\ 0 & e^4 \end{bmatrix} U^{-1} = \frac{1}{3} \begin{bmatrix} 2e + e^4 & e^4 - e + i(e^4 - e) \\ e^4 - e + i(e - e^4) & e + 2e^4 \end{bmatrix}.$$

5 Positive operators

Recall that self-adjoint operators are the operator analogue of real numbers. Let us now define the operator analogue of positive (or more precisely nonnegative) real numbers.

Definition 4. An operator $T \in \mathcal{L}(V)$ is called **positive** (in symbols $T \geq 0$) if $T = T^*$ and $\langle Tv, v \rangle \geq 0$ for all $v \in V$.

(If V is a complex vector space the condition of self-adjointness follows from the condition $\langle Tv, v \rangle \geq 0$ and can hence be dropped).

Example 5. Note that for all $T \in \mathcal{L}(V)$ we have $T^*T \geq 0$ since T^*T is self-adjoint and $\langle T^*Tv, v \rangle = \langle Tv, Tv \rangle \geq 0$.

Example 6. Let $U \subset V$ be a subspace of V and P_U the orthogonal projection onto U. Then $P_U \geq 0$. To see this write $V = U \oplus U^\perp$ and $v = u_v + u_v^\perp$ for all $v \in V$ where $u_v \in U$ and $u_v^\perp \in U^\perp$. Then $\langle P_U v, w \rangle = \langle u_v, u_w + u_w^\perp \rangle = \langle u_v, u_w \rangle = \langle u_v + u_v^\perp, u_w \rangle = \langle v, P_U w \rangle$, so that $P_U^* = P_U$. Also, setting $v = w$ in the above string of equations we obtain $\langle P_U v, v \rangle = \langle u_v, u_v \rangle \geq 0$ for all $v \in V$. Hence $P_U \geq 0$.

If λ is an eigenvalue of a positive operator T and $v \in V$ is the associated eigenvector, then $\langle Tv, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle \geq 0$. Since $\langle v, v \rangle \geq 0$ for all vectors $v \in V$, it follows that $\lambda \geq 0$. This fact can be used to define $\sqrt{T}$ by setting

$$\sqrt{T} e_i = \sqrt{\lambda_i} e_i,$$

where λ_i are the eigenvalues of T with respect to the orthonormal basis $e = (e_1, \ldots, e_n)$. We know that these exist by the spectral theorem.

6 Polar decomposition

Continuing the analogy between $\mathbb{C}$ and $\mathcal{L}(V)$, recall the polar form of a complex number $z = |z|e^{i\theta}$, where $|z|$ is the absolute value or length of z and $e^{i\theta}$ is an element on the unit circle. In terms of an operator $T \in \mathcal{L}(V)$ where V is a complex inner product space, a unitary operator U takes the role of $e^{i\theta}$ and $|T|$ takes the role of the length. As we discussed above $T^*T \geq 0$ so that $|T| := \sqrt{T^*T}$ exists and $|T| \geq 0$ as well.

Theorem 9. *For all $T \in \mathcal{L}(V)$ there exists a unitary U such that*

$$T = U|T|.$$

*This is called the **polar decomposition** of T.*

Sketch of proof. We start by noting that

$$\|Tv\|^2 = \||T|v\|^2,$$

since $\langle Tv, Tv\rangle = \langle v, T^*Tv\rangle = \langle \sqrt{T^*T}v, \sqrt{T^*T}v\rangle$. This implies that $\operatorname{null} T = \operatorname{null}|T|$. Because of the dimension formula $\dim \operatorname{null} T + \dim \operatorname{range} T = \dim V$, this also means that $\dim \operatorname{range} T = \dim \operatorname{range}|T|$. Moreover, we can hence define an isometry $S : \operatorname{range}|T| \to \operatorname{range} T$ by setting

$$S(|T|v) = Tv.$$

The trick is now to define a unitary operator U on all of V such that the restriction of U onto the range of $|T|$ is S

$$U|_{\operatorname{range}|T|} = S.$$

Note that $\operatorname{null}|T| \perp \operatorname{range}|T|$ because for $v \in \operatorname{null}|T|$ and $w = |T|u \in \operatorname{range}|T|$

$$\langle w, v\rangle = \langle |T|u, v\rangle = \langle u, |T|v\rangle = \langle u, 0\rangle = 0$$

since $|T|$ is self-adjoint.

Pick an orthonormal basis $e = (e_1, \ldots, e_m)$ of $\operatorname{null}|T|$ and an orthonormal basis $f = (f_1, \ldots, f_m)$ of $(\operatorname{range} T)^\perp$. Set $\tilde{S}e_i = f_i$ and extend $\tilde{S}$ on $\operatorname{null}|T|$ by linearity. Any $v \in V$ can be uniquely written as $v = v_1 + v_2$ where $v_1 \in \operatorname{null}|T|$ and $v_2 \in \operatorname{range}|T|$ since $\operatorname{null}|T| \perp \operatorname{range}|T|$. Then define $U : V \to V$ by $Uv = \tilde{S}v_1 + Sv_2$. Now U is an isometry and hence unitary as shown by the following calculation using the Pythagorean theorem

$$\begin{aligned}\|Uv\|^2 &= \|\tilde{S}v_1 + Sv_2\|^2 = \|\tilde{S}v_1\|^2 + \|Sv_2\|^2 \\ &= \|v_1\|^2 + \|v_2\|^2 = \|v\|^2.\end{aligned}$$

Also by construction $U|T| = T$ since $U|_{\mathrm{null}\,|T|}$ does not matter. □

7 Singular-value decomposition

The singular-value decomposition generalizes the notion of diagonalization. To unitarily diagonalize $T \in \mathcal{L}(V)$ means to find an orthonormal basis e such that T is diagonal with respect to this basis

$$M(T; e, e) = [T]_e = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

where the notation $M(T; e, e)$ indicates that the basis e is used both for the domain and codomain of T. The spectral theorem tells us that unitary diagonalization can only be done for normal operators. In general, we can find two orthonormal bases e and f such that

$$M(T; e, f) = \begin{bmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{bmatrix}$$

which means that $Te_i = s_i f_i$. The scalars s_i are called **singular values** of T. If T is diagonalizable they are the absolute values of the eigenvalues.

Theorem 10. *All $T \in \mathcal{L}(V)$ have a singular-value decomposition. That is, there exist orthonormal bases $e = (e_1, \ldots, e_n)$ and $f = (f_1, \ldots, f_n)$ such that*

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

where s_i are the singular values of T.

Proof. Since $|T| \geq 0$ and hence also self-adjoint, there is an orthonormal basis $e = (e_1, \ldots, e_n)$ by the spectral theorem so that $|T|e_i = s_i e_i$. Let U be the unitary matrix in the polar decomposition, so that $T = U|T|$. Since e is orthonormal, we can write any vector $v \in V$ as

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

and hence

$$Tv = U|T|v = s_1 \langle v, e_1 \rangle Ue_1 + \cdots + s_n \langle v, e_n \rangle Ue_n.$$

Now set $f_i = Ue_i$ for all $1 \leq i \leq n$. Since U is unitary, $(f_1, \ldots, f_n)$ is also an orthonormal basis, proving the theorem. $\square$