1 The Spectral Theorem

Theorem 1.1 (Thm 7.9)**.** *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{C}$ and $\langle\,\cdot\,,\cdot\,\rangle$ be a positive definite inner product on $\mathcal{V}$. Let $T \in \mathcal{L}(\mathcal{V})$. Then $\mathcal{V}$ has an orthonormal basis of eigenvectors if and only if T is normal.*

Proof. ($\Leftarrow$) Assume T is normal. Set $n = \dim(\mathcal{V})$. Then there exists an orthonormal basis $\mathcal{B} = \{u_1, \ldots, u_n\}$ of $\mathcal{V}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(T)$ is upper-triangular. Thus,

$$M_{\mathcal{B}}^{\mathcal{B}}(T) = \begin{bmatrix} \langle Tu_1, u_1\rangle & \langle Tu_2, u_1\rangle & \cdots & \langle Tu_n, u_1\rangle \\ 0 & \langle Tu_2, u_2\rangle & \cdots & \langle Tu_n, u_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle Tu_n, u_n\rangle \end{bmatrix}$$

Let $v \in \mathcal{V}$. Then $v = \langle v, u_1\rangle u_1 + \ldots + \langle vu_n\rangle u_n$. Since $T\mathcal{U}_j \subseteq \mathcal{U}_j$, we have $Tu_j \in \operatorname{span}\{u_1, \ldots, u_j\}$, $\forall\, j \in \{1, \ldots, n\}$. It follows that $Tu_j = \langle Tu_j u_1\rangle u_1 + \ldots + \langle Tu_j u_j\rangle u_j$.

Now $M_{\mathcal{B}}^{\mathcal{B}}(T^*) = \left(M_{\mathcal{B}}^{\mathcal{B}}(T)\right)^* = \begin{bmatrix} \mathcal{C}_{\mathcal{B}}(T^*u_1) & \cdots & \mathcal{C}_{\mathcal{B}}(T^*u_n)\end{bmatrix}$ and $\|Tu_1\|^2 = \|T^*u_1\|^2$. It follows that $\|Tu_1\|^2 = |\langle Tu_1, u_1\rangle|^2$ and $\|T^*u_1\|^2 = \sum_{j=1}^n |\langle Tu_j, u_1\rangle|^2$. Thus, we have $\langle Tu_j, u_1\rangle = 0$, for $j = 2, \ldots, n$. A similar argument for $\|Tu_j\|^2$, $\forall\, j \in \{2, \ldots, n\}$, shows that all nondiagonal entries are zero.

($\Rightarrow$) Now assume $\{u_1, \ldots, u_n\}$ is an orthonormal basis of $\mathcal{V}$ such that $Tu_j = \lambda_j u_j$, $\forall\, j \in \{1, \ldots, n\}$. Then $M_{\mathcal{B}}^{\mathcal{B}}(T) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ and $M_{\mathcal{B}}^{\mathcal{B}}(T^*) = \begin{bmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{bmatrix}$.

Since $M_{\mathcal{B}}^{\mathcal{B}}(TT^*) = M_{\mathcal{B}}^{\mathcal{B}}(T)M_{\mathcal{B}}^{\mathcal{B}}(T^*) = \begin{bmatrix} \lambda_1\overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n\overline{\lambda_n} \end{bmatrix} = M_{\mathcal{B}}^{\mathcal{B}}(T^*T)$, we have $TT^* = T^*T$. Hence, T is normal. ☐

2 Invariance under a linear operator

Theorem 2.1. *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{C}$. Let $\langle\cdot,\cdot\rangle$ be a positive definite inner product on $\mathcal{V}$. Let $T \in \mathcal{L}(\mathcal{V})$ be normal. Lastly, let $\mathcal{U}$ be a subspace of $\mathcal{V}$. Then*

$$T\mathcal{U} \subseteq \mathcal{U} \Leftrightarrow T\mathcal{U}^\perp \subseteq \mathcal{U}^\perp$$

(Recall that we have previously proved that for any $T \in \mathcal{L}(\mathcal{V})$, $T\mathcal{U} \subseteq \mathcal{U} \Leftrightarrow T^*\mathcal{U}^\perp \subseteq \mathcal{U}^\perp$. Hence if T is normal, showing that any one of $\mathcal{U}$ or $\mathcal{U}^\perp$ is invariant under either T or T^* implies that the rest are, also.)

Proof. Assume $T\mathcal{U} \subseteq \mathcal{U}$. We know $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$. Let $u_1, \ldots, u_m$ be an orthonormal basis of $\mathcal{U}$ and $u_{m+1}, \ldots, u_n$ be an orthonormal basis of $\mathcal{U}^\perp$. Then $u_1, \ldots, u_n$ is an orthonormal basis of $\mathcal{V}$. If $j \in \{1, \ldots, m\}$ then $u_j \in \mathcal{U}$, so $Tu_j \in \mathcal{U}$. Hence

$$Tu_j = \sum_{k=1}^m \langle Tu_j, u_k\rangle u_k.$$

Also, clearly,

$$T^*u_j = \sum_{k=1}^n \langle T^*u_j, u_k\rangle u_k.$$

On the most recent exam, we proved that $\|Tu_j\|^2 = \sum_{k=1}^m |\langle Tu_j, u_k\rangle|^2$. Further, by normality, $\|Tu_j\|^2 = \|T^*u_j\|^2$. Hence

$$\begin{aligned}
\sum_{j=1}^m \|Tu_j\|^2 &= \sum_{j=1}^m \|T^*u_j\|^2 \\
&= \sum_{j=1}^m \sum_{k=1}^n |\langle T^*u_j, u_k\rangle|^2 \\
&= \sum_{j=1}^m \sum_{k=1}^n |\langle u_j, Tu_k\rangle|^2 && \text{(by the definition of } T^*\text{)} \\
&= \sum_{j=1}^m \sum_{k=1}^n |\overline{\langle Tu_k, u_j\rangle}|^2 && \text{(by hermiticity)} \\
&= \sum_{j=1}^m \sum_{k=1}^n |\langle Tu_k, u_j\rangle|^2 && \text{(because moduli are real)} \\
&= \sum_{j=1}^m \left(\sum_{k=1}^m |\langle Tu_k, u_j\rangle|^2 + \sum_{k=m+1}^n |\langle Tu_k, u_j\rangle|^2 \right) \\
&= \sum_{j=1}^m \sum_{k=1}^m |\langle Tu_k, u_j\rangle|^2 + \sum_{j=1}^m \sum_{k=m+1}^n |\langle Tu_k, u_j\rangle|^2 \\
&= \sum_{k=1}^m \sum_{j=1}^m |\langle Tu_k, u_j\rangle|^2 + \sum_{j=1}^m \sum_{k=m+1}^n |\langle Tu_j, u_k\rangle|^2 && \text{(by exchanging the order of summation)} \\
&= \sum_{j=1}^m \sum_{k=1}^m |\langle Tu_j, u_k\rangle|^2 + \sum_{j=1}^m \sum_{k=m+1}^n |\langle Tu_j, u_k\rangle|^2 && \text{(by reindexing)} \\
&= \sum_{j=1}^m \|Tu_j\|^2 + \sum_{j=1}^m \sum_{k=m+1}^n |\langle Tu_j, u_k\rangle|^2,
\end{aligned}$$

implying $\sum_{j=1}^m \sum_{k=m+1}^n |\langle Tu_j, u_k\rangle|^2 = 0$. As each term is nonnegative, we conclude that $|\langle Tu_j, u_k\rangle|^2 = 0$ for all $j \in \{1, \ldots, n\}$ and all $k \in \{1, \ldots, n\}$. Thus $|\langle T^*u_j, u_k\rangle|^2 = 0, \forall\ 1 \le j \le m, m+1 \le k \le n$. Hence $\langle T^*u_j, u_k\rangle = 0, \forall\ 1 \le j \le m, m+1 \le k \le n$. Thus

$$T^*u_j = \sum_{k=1}^m \langle T^*u_j, u_k\rangle u_k.$$

Therefore $T^*\mathcal{U} \subseteq \mathcal{U}$. Then, because we know that $\mathcal{U}$ is invariant under T if and only if $\mathcal{U}^\perp$ is invariant under T^*, we conclude that $T\mathcal{U}^\perp \subseteq \mathcal{U}^\perp$. □

(Alternate proof)

Proof. Assume T is normal. Then there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ and $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{C}$ such that

$$Tu_j = \lambda_j u_j \iff T^*u_j = \overline{\lambda_j} u_j, j \in \{1, \ldots, n\}.$$

Let v be arbitrary in $\mathcal{V}$. We can write

$$Tv = \sum_{j=1}^n \lambda_j \langle v, u_j\rangle u_j$$

and

$$T^*v = \sum_{j=1}^{n} \overline{\lambda_j}\langle v, u_j\rangle u_j.$$

Set $p(z) = a_0 + a_1 z + \cdots + a_m z^m \in \mathbb{C}[z]$. Then $p(Tv) = \sum_{j=1}^{n} p(\lambda_j)\langle v, u_j\rangle u_j$. We need a $p \in \mathbb{C}[z]$ such that $p(\lambda_j) = \overline{\lambda_j},\ \forall j \in \{1, \ldots, n\}$. We proved in the homework (assignment 2, #3), that if $S : \mathbb{C}[z]_{<n} \to \mathbb{C}^n$ is defined by

$$Sp = [p(z_1) \cdots p^{(l_1-1)}(z_1) \cdots p(z_m) \cdots p^{(l_m-1)(z_m)}]^{\intercal},$$

then S is an isomorphism. Hence by the surjectivity of S, we can find $p \in \mathbb{C}[z]$ such that $p(\lambda_j) = \overline{\lambda_j},\ \forall j \in \{1, \ldots, n\}$, Thus $p(Tv) = T^*v$. Now assume $T\mathcal{U} \subseteq \mathcal{U}$. It follows that $T^k\mathcal{U} \subseteq \mathcal{U}$ for all $k \in \mathbb{N}$ and also that $\alpha T\mathcal{U} \subseteq \mathcal{U}$ for all $\alpha \in \mathbb{C}$. Hence $p(T)\mathcal{U} = T^*\mathcal{U} \subseteq \mathcal{U}$. □

(Thm 7.18 Axler)

Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{C}$ with a positive definite inner product. Let $T \in \mathcal{L}(\mathcal{V})$ be normal. Let $\mathcal{U}$ be a subspace of $\mathcal{V}$. Then $TU \subseteq U \iff T(U^\perp) \subseteq U^\perp$.

Proof. Assume $T\mathcal{U} \subseteq \mathcal{U}$. Let $u \in \mathcal{U}$. Then $Tu \in \mathcal{U}$. Let $w \in \mathcal{U}^\perp$. Then $0 = \langle Tu, w\rangle = \langle u, T^*w\rangle$, which implies $T^*w \in \mathcal{U}^\perp$. Hence, $T^*(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$.

Now $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$. Let $n = \dim(\mathcal{V})$. Let $\{u_1, \ldots, u_m\}$ be an orthonormal basis of $\mathcal{U}$ and $\{u_{m+1}, \ldots, u_n\}$ be an orthonormal basis of $\mathcal{U}^\perp$. Then $\mathcal{B} = \{u_1, \ldots, u_n\}$ is an orthonormal basis of $\mathcal{V}$ such that

$$M_{\mathcal{B}}^{\mathcal{B}}(T) = \begin{array}{c} \\ u_1 \\ \vdots \\ u_m \\ u_{m+1} \\ \vdots \\ u_n \end{array} \begin{array}{ccc|ccc} u_1 & \cdots & u_m & u_{m+1} & \cdots & u_n \\ \langle Tu_1, u_m\rangle & \cdots & \langle Tu_m, u_1\rangle & & & \\ \vdots & \ddots & \vdots & & B & \\ \langle Tu_1, u_1\rangle & \cdots & \langle Tu_m, u_m\rangle & & & \\ \hline & & & & & \\ & 0 & & & C & \\ & & & & & \end{array}$$

Take $j \in \{1, \ldots, m\}$. Then $Tu_j = \sum_{k=1}^{m}\langle Tu_j, u_k\rangle u_k$. Calculate $\|Tu_j\|^2 = \sum_{k=1}^{m} |\langle Tu_j, u_k\rangle|^2$ and $\|T^*u_j\|^2 = \sum_{k=1}^{n} |\langle T^*u_j, u_k\rangle|^2$. Since T is normal, $\sum_{j=1}^{m} \|Tu_j\|^2 = \sum_{j=1}^{m} \|T^*u_j\|^2$. Now we have

$$\begin{aligned}\sum_{j=1}^{m}\sum_{k=1}^{m} |\langle Tu_j, u_k\rangle|^2 &= \sum_{j=1}^{m}\sum_{k=1}^{m} |\langle T^*u_j, u_k\rangle|^2 + \sum_{k=m+1}^{n} |\langle T^*u_j, u_k\rangle|^2 \\ &= \sum_{j=1}^{m}\sum_{k=1}^{m} |\langle T^*u_j, u_k\rangle|^2 + \sum_{j=1}^{m}\sum_{k=m+1}^{n} |\langle T^*u_j, u_k\rangle|^2.\end{aligned}$$

Since $|\langle T^*u_j, u_k\rangle|^2 = |\langle u_j, Tu_k\rangle|^2 = |\langle Tu_k, u_j\rangle|^2$, it follows that $\langle T^*u_j, u_k\rangle = 0,\ \forall j \in \{1, \ldots, m\}$, $\forall k \in \{m+1, \ldots, n\}$. Thus, $B = 0$. Hence, $T^*u_j \in \mathcal{U},\ \forall j \in \{1, \ldots, m\}$, which implies $T^*\mathcal{U} \subseteq \mathcal{U}$.

Considering $M_{\mathcal{B}}^{\mathcal{B}}(T)$ for $j \in \{m+1, \ldots, n\}$, we have $Tu_j \in \operatorname{span}\{u_{m+1}, \ldots, u_n\}$. Thus, $Tu_j \in \mathcal{U}^\perp$, which implies $T(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$. Finally, letting $\mathcal{U} = \mathcal{U}^\perp$, a similar argument shows that $T\mathcal{U} \subseteq \mathcal{U}$. □

3 Polar Decomposition

Consider an analogy between $\mathcal{L}(\mathcal{V})$ and $\mathbb{C}$. The adjoint of T, T^*, is analogous to $\overline{z}$, the conjugate of z, although $T^*T = TT^*$ only when T is normal, whereas $\overline{z}z = z\overline{z},\ \forall z \in \mathbb{C}$. Self-adjoint maps in $\mathcal{L}(\mathcal{V})$ correspond to $\mathbb{R} \subset \mathbb{C}$. The set of unitary operators, i.e. all $T \in \mathcal{L}(\mathcal{V})$ such that $T^*T = I$, correspond to $\Pi = \{z \in \mathbb{C} : |z| = 1\}$. Whence given that all $z \in \mathbb{C}$ have a polar decomposition, i.e. for all z there exists an $r \geq 0$ and a $u \in \mathbb{C}$ such that $|u| = 1$, such that $z = ru$, there exists an equivalent concept in $\mathcal{L}(\mathcal{V})$.

Definition 3.1. An operator $P \in \mathcal{L}(\mathcal{V})$ is nonnegative if $\langle Pv, v\rangle \geq 0,\ \forall v \in \mathcal{V}$. Please note, Axler uses the term "positive" to describe such an operator. Also note, if an operator is nonnegative, that implies that it is self-adjoint, and hence normal.

Definition 3.2. An operator $U \in \mathcal{L}(\mathcal{V})$ is unitary if $U^*U = I$. An operator is unitary if and only if it is angle preserving:

$$\begin{aligned} \langle u, v\rangle &= \langle Iu, v\rangle \qquad \text{for any } u, v \in \mathcal{V} \\ &= \langle U^*Uu, v\rangle \\ &= \langle Uu, Uv\rangle. \end{aligned}$$

Theorem 3.3. *For all nonnegative $P \in \mathcal{L}(\mathcal{V})$ there exists a unique nonnegative $Q \in \mathcal{L}(\mathcal{V})$ such that $P = Q^2$. We will use $\sqrt{P}$ to denote this Q.*

Proof. (This is a proof for existence only.) By the spectral theorem, we know there exists an orthonormal basis $u_1, \ldots, u_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n \geq 0$ such that $Pv = \sum_{j=1}^{n} \lambda_j \langle v, u_j\rangle u_j$. Set

$$Qv = \sum_{j=1}^{n} \sqrt{\lambda_j}\langle v, u_j\rangle u_j.$$

□

Notice that $\operatorname{nul} P = \operatorname{nul} Q$. Also, the eigenvalues of Q are in the form $\sqrt{\lambda_j}$.

Theorem 3.4. *(Polar Decomposition in $\mathcal{L}(\mathcal{V})$) For all $T \in \mathcal{L}(\mathcal{V})$ there exists a unitary operator U in $\mathcal{L}(\mathcal{V})$ and a nonnegative $P \in \mathcal{L}(\mathcal{V})$ such that $T = UP$.*

Proof. First, notice that T^*T is nonnegative: $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = ||Tv||^2 \geq 0$. Set $P = \sqrt{T^*T}$. Then $\operatorname{nul} P = \operatorname{nul}(T^*T) \supseteq \operatorname{nul}(T)$. Let $v \in \operatorname{nul}(T^*T)$. Then $T^*Tv = 0$. Thus $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = 0$. Hence $||Tv|| = 0$, implying $Tv = 0$. Therefore v is in $\operatorname{nul} T$, for all $v \in \operatorname{nul}(T^*T)$. Thus by symmetric containment, $\operatorname{nul} P = \operatorname{nul}(T^*T) = \operatorname{nul} T$. Then, by the rank-nullity theorem, $\dim \operatorname{ran}(P) = \dim \operatorname{ran}(T)$. Consider $\psi : \operatorname{ran}(P) \to \operatorname{ran}(T)$ such that $Pv \mapsto Tv$. Suppose $Pv_1 = Pv_2$. Then $v_1 - v_2 \in \operatorname{nul} P = \operatorname{nul} T$. Thus $v_1 - v_2 \in \operatorname{nul} T$. Thus $Tv_1 = Tv_2$. Hence ψ is injective. Thus by injectivity and the dimension argument, ψ is a bijection. Let $v, w \in \mathcal{V}$. Consider

$$\begin{aligned} \langle \psi Pv, \psi Pw\rangle &= \langle Tv, Tw\rangle \\ &= \langle T^*Tv, w\rangle \\ &= \langle P^2v, w\rangle \\ &= \langle P^*Pv, w\rangle \quad \text{(because } P \text{ is self-adjoint)} \\ &= \langle Pv, Pw\rangle \end{aligned}$$

Thus ψ is angle-preserving on $\operatorname{ran}(P)$. Let us consider $(\operatorname{ran}(P))^\perp$. Let $v_1, \ldots, v_m$ be an orthonormal basis on $(\operatorname{ran}(P))^\perp$ and let $u_1, \ldots, u_m$ be an orthonormal basis on $(\operatorname{ran}(T))^\perp$. Define $U_1 : (\operatorname{ran}(P))^\perp \to (\operatorname{ran}(T))^\perp$ by

$$\begin{aligned} U_1 v &= U_1\left(\sum_{j=1}^{m}\langle v, v_j\rangle v_j\right) \\ &= \sum_{j=1}^{m}\langle v, v_j\rangle u_j. \end{aligned}$$

Thus

$$\langle U_1 v, U_1 w\rangle = \sum_{j=1}^{m} \langle v, v_j\rangle \overline{\langle w, v_j\rangle}$$
$$= \langle v, w\rangle$$

Hence U_1 is unitary on $(\mathrm{ran}(P))^{\perp}$. Define $U : \mathcal{V} \to \mathcal{V}$ by

$$Uv = \psi P v + U_1(I-P)v.$$

Notice that $Pv \in (\mathrm{ran}(P))$ and $(I-P)v \in (\mathrm{ran}(P))^{\perp}$. We claim that U is unitary:

$$\begin{aligned}\langle Uv, Uw\rangle &= \langle \psi Pv + U_1(I-P)v, \psi Pw + U_1(I-P)w\rangle \\ &= \langle \psi Pv, \psi Pw\rangle + \langle U_1(I-P)v, U_1(I-P)w\rangle \\ &= \langle Tv, Tw\rangle + \langle (I-T)v, (I-T)w\rangle \\ &= \langle v, w\rangle\end{aligned}$$

Hence U is unitary. Thus we can write $T = U \circ \sqrt{T^*T}$, where U is unitary and $\sqrt{T^*T}$ is nonnegative. □

(Thm 7.41 Axler)

If $T \in \mathcal{L}(\mathcal{V})$, then there exists an isometry $S \in \mathcal{L}(\mathcal{V})$ such that $T = S\sqrt{T^*T}$.

Proof. Suppose $T \in \mathcal{L}(\mathcal{V})$. Let $v \in \mathcal{V}$. Then

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

Thus, $\|Tv\| = \|\sqrt{T^*T}v\|, \ \forall\, v \in \mathcal{V}$.

Define $S_1 : \mathrm{ran}(\sqrt{T^*T}) \to \mathrm{ran}(T)$ by $S_1(\sqrt{T^*T}v) = Tv$. We need to check that S_1 is well-defined. Let $v_1, v_2 \in \mathcal{V}$ such that $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$. Then $\|Tv_1 - Tv_2\| = \|T(v_1 - v_2)\| = \|\sqrt{T^*T}(v_1 - v_2)\| = \|\sqrt{T^*T}v_1 - \sqrt{T^*T}v_2\| = 0$. Thus, $Tv_1 = Tv_2$, and S_1 is well-defined.

Since S_1 maps $\mathrm{ran}(\sqrt{T^*T})$ onto $\mathrm{ran}(T)$, for every $u \in \mathrm{ran}(\sqrt{T^*T})$, we have $\|S_1 u\| = \|u\|$.

Now we need to show $\mathrm{nul}(T^*T) = \mathrm{nul}(T)$. First of all, we have $\mathrm{nul}(T) \subseteq \mathrm{nul}(T^*T)$. For the other direction, let $v \in \mathrm{nul}(T^*T)$. Then $T^*Tv = 0 \implies \langle T^*Tvv = 0 \implies \langle TvTv = 0 \implies Tv = 0 \implies v \in \mathrm{nul}(T)$. Thus, $\mathrm{nul}(T^*T) \subseteq \mathrm{nul}(T)$, so that $\mathrm{nul}(T^*T) = \mathrm{nul}(T)$.

Since $\mathrm{nul}(\sqrt{T^*T}) = \mathrm{nul}(T^*T)$, we have $\mathrm{nul}(\sqrt{T^*T}) = \mathrm{nul}(T)$. By the Rank-Nullity theorem, it follows that $\dim(\mathrm{ran}(\sqrt{T^*T})) = \dim(\mathrm{ran}(T))$. Hence, $\dim(\mathrm{ran}(\sqrt{T^*T}))^{\perp} = \dim(\mathrm{ran}(T))^{\perp}$.

Let $\{u_1, \ldots, u_m\}$ be an orthonormal basis of $(\mathrm{ran}(\sqrt{T^*T}))^{\perp}$ and $\{v_1, \ldots, v_n\}$ be an orthonormal basis of $(\mathrm{ran}(T))^{\perp}$. Define $S_2 : (\mathrm{ran}(\sqrt{T^*T}))^{\perp} \to (\mathrm{ran}(T))^{\perp}$ by $S_2\left(\sum_{j=1}^{m}\langle v, u_j\rangle u_j\right) = \sum_{j=1}^{m}\langle v, u_j\rangle v_j$. We have $\|S_2 w\| = \|w\|, \ \forall\, w \in (\mathrm{ran}(\sqrt{T^*T}))^{\perp}$, since $\|S_2 w\| = \sum_{j=1}^{m}|\langle v, u_j\rangle|^2 = \|w\|$.

Now let $S : \mathcal{V} \to \mathcal{V}$ be defined by $S(v) = S_1 u + S_2 w$ where $v = u + w$ with $u \in \mathrm{ran}(\sqrt{T^*T})$ and $w \in (\mathrm{ran}(\sqrt{T^*T}))^{\perp}$. For each $v \in \mathcal{V}$, we have $S(\sqrt{T^*T}v) = S_1(\sqrt{T^*T}v) = Tv$. Thus, $T = S\sqrt{T^*T}$.

To show that S is an isometry, let $v \in \mathcal{V}$ such that $v = u + w$ where $u \in \mathrm{ran}(\sqrt{T^*T})$ and $w \in (\mathrm{ran}(\sqrt{T^*T}))^{\perp}$. Then $\|Sv\|^2 = \|S_1 u + S_2 w\|^2 = \|S_1 u\|^2 + \|S_2 w\|^2$ (since $S_1 u \perp S_2 w$), $= \|u\|^2 + \|w\|^2 = \|v\|^2$. □

Thm 7.46 **Singular-Value Decomposition.**

Suppose $T \in \mathcal{L}(\mathcal{V})$ has singular values $s_1, \ldots, s_n$. Then there exist an orthonormal bases $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$ such that $Tv = s_1\langle v, u_1\rangle v_1 + \cdots + s_n\langle v, u_n\rangle v_n$.

Proof. By the spectral theorem applied to $\sqrt{T^*T}$, there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathcal{V}$ such that $\sqrt{T^*T}u_j = s_j u_j, \ \forall\, j \in \{1, \ldots, n\}$. Let $v \in \mathcal{V}$. Then $v = \langle v u_1 u_1 + \ldots + \langle v u_n u_n$. Applying $\sqrt{T^*T}$ to both sides, we get $\sqrt{T^*T}v = s_1\langle v u_1 u_1 + \ldots + s_n\langle v u_n u_n$.

By polar decomposition, there exists an isometry $S \in \mathcal{L}(\mathcal{V})$ such that $T = S\sqrt{T^*T}$. Applying S to both sides, we get $S\sqrt{T^*T}v = Tv = s_1\langle v, u_1\rangle Su_1 + \cdots + s_n\langle v, u_n\rangle Su_n$. Now let $v_j = Su_j$, $\forall\, j \in \{1, \ldots, n\}$. Since S is an isometry, $\{v_1, \ldots, v_n\}$ is an orthonormal of $\mathcal{V}$. Hence, $Tv = s_1\langle v, u_1\rangle v_1 + \cdots + s_n\langle v, u_n\rangle v_n$, $\forall\, v \in \mathcal{V}$. □

4 Cauchy-Bunyakovsky-Schwarz Inequality

Theorem 4.1. *(Cauchy-Bunyakovsky-Schwarz Inequality) If $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ is an inner product space, where $\langle\cdot,\cdot\rangle$ is a nonnegative inner product, then $\forall\, u, v \in \mathcal{V}$,*

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle,$$

or equivalently,

$$|\langle u, v\rangle| \leq ||u||\ ||v||,$$

with equality if and only if there exists α, β, not both zero, in $\mathbb{F}$ such that

$$\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0.$$

■ ....■ Branko Ćurgus' comment starts here. ■ ....■

I don't see that the proof below proves what the claim.

There are two claims.

Assume that $\mathcal{V}$ a vector space over $\mathbb{F}$ and $\langle\,\cdot\,,\cdot\,\rangle$ is a nonnegative inner product on $\mathcal{V}$.

The first claim is:

Let $u, v \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{F}$. Then

$$|\alpha|^2 + |\beta|^2 > 0 \quad \text{and} \quad \langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0 \qquad \Rightarrow \qquad |\langle u, v\rangle|^2 = \langle u, u\rangle\langle v, v\rangle.$$

This is the easier part of the proof. I do not see that it is proved below. I will prove it here.

Assume $|\alpha|^2 + |\beta|^2 > 0$ and $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$. We consider two cases $\alpha \neq 0$ and $\beta \neq 0$. Assume $\alpha \neq 0$. Set $w = \alpha u + \beta v$. Then $\langle w, w\rangle = 0$. Also $u = \gamma v + \delta w$ where $\gamma = -\beta/\alpha$ and $\delta = 1/\alpha$. Notice that the Cauchy-Bunyakovsky-Schwarz inequality and $\langle w, w\rangle = 0$ implies that $\langle w, x\rangle = 0$ for all $x \in \mathcal{V}$. Now we calculate

$$\begin{aligned}
|\langle u, v\rangle|^2 &= |\langle \gamma v + \delta w, v\rangle|^2 \\
&= |\gamma\langle v, v\rangle + \delta\langle w, v\rangle|^2 \\
&= |\gamma\langle v, v\rangle|^2 \\
&= |\gamma|^2\langle v, v\rangle\langle v, v\rangle \\
&= \langle \gamma v, \gamma v\rangle\langle v, v\rangle \\
&= \langle \gamma v + \delta w, \gamma v + \delta w\rangle\langle v, v\rangle \\
&= \langle u, u\rangle\langle v, v\rangle.
\end{aligned}$$

This completes the proof of the first claim.

The proof of the second claim is more complicated.

The second claim is:

Let $u, v \in \mathcal{V}$. Then

$$|\langle u, v\rangle|^2 = \langle u, u\rangle\langle v, v\rangle \qquad \Rightarrow \qquad \exists\ \alpha, \beta \in \mathbb{F} \text{ s.t. } |\alpha|^2 + |\beta|^2 > 0 \quad \text{and} \quad \langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0.$$

To create $\alpha, \beta \in \mathbb{F}$ one has to go back to the proof of the Cauchy-Bunyakovsky-Schwarz inequality and use the high school theorem to create α and β. A correct proof of this must offer a way of creating α and β.

■ ....■ Branko Ćurgus' comment ends here. ■ ....■

Proof. (Proof of equality condition only) We know that when $\langle \cdot, \cdot \rangle$ is positive definite, $|\langle u, v\rangle|^2 = \langle u, u\rangle\langle vv\rangle$ if and only if u and v are linearly independent (Axler). When $\langle \cdot, \cdot \rangle$ is nonnegative, but not positive definite, there exists a $u_0 \neq 0$ in $\mathcal{V}$ such that $\langle u_0, u_0\rangle = 0$. Hence $\langle u_0, u_0\rangle\langle v, v\rangle = 0$ for all $v \in \mathcal{V}$. From the inequality, we know $|\langle u_0, v\rangle|^2 \leq 0$, but by the non-negativity of $\langle \cdot, \cdot \rangle$ we also know that $|\langle u_0, v\rangle|^2 \geq 0$. Hence

$$|\langle u_0, v\rangle|^2 = 0 = \langle u_0, u_0\rangle\langle v, v\rangle, \forall v \in \mathcal{V}.$$

To say that u, v are linearly independent is equivalent to saying there exists α, β, not both zero, in $\mathbb{F}$, such that $\alpha u + \beta v = 0$ implies $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$. Thus if $\langle \cdot, \cdot \rangle$ is nonnegative, whenever $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$ for some α, β not both zero, we have equality. Suppose $|\langle u, v\rangle|^2 = \langle u, u\rangle\langle v, v\rangle$. Then either u or v is such that $\langle u, u\rangle = 0$ or $\langle v, v\rangle = 0$. If, without loss of generality, $\langle u, u\rangle = 0$, then for any nonzero α and for $\beta = 0$, $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$. If $u, v \neq 0$, and neither $\langle u, u\rangle = 0$ nor $\langle v, v\rangle = 0$, then it must be that u, v are linearly independent. Hence $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$. Thus in either case we have equality if and only if there exists α, β, not both zero, in $\mathbb{F}$ such that $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$. □

For an example, suppose $\mathcal{V} = \mathcal{C}[0, 1]$, the set of all continuous functions on the interval $[0, 1]$. The inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$, where $\int dx$ denotes the Riemann integral, is a positive definite inner product $\mathcal{V}$. However, with the corresponding norm the space $\mathcal{V}$ is not complete. Since the completeness is the founding principle of analysis one needs to complete this space. The completion leads to the concept of the Lebesgue integral. We consider the space of all measurable functions f on $[0, 1]$ such that the Lebesgue integral $\int_0^1 (f(x))^2 d\mu$ is finite. The corresponding inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\mu(dx)$, where μ denotes the Lebesgue measure, is not positive definite. It is a nonnegative inner product. Hence the Cauchy-Bunyakovsky-Schwarz Inequality holds for the inner product $\int_0^1 f(x)g(x)\mu(dx)$.

5 Jordan Normal Form

Let $\mathcal{V}$ be vector space over $\mathbb{C}$. Let $\dim \mathcal{V} = n$. Let $T \in \mathcal{L}(\mathcal{V})$. Consider the set of nilpotent operators in $\mathcal{L}(\mathcal{V})$: $\{N \in \mathcal{L}(\mathcal{V}) : \exists k \in \mathbb{N}$ such that $N^k = 0\}$. We define the degree of nilpotency of N to be q such that $N^q = 0$, but $N^{q-1} \neq 0$. For an example, suppose $n = 3$ and there exists a basis $\mathcal{B}$ of $\mathcal{V}$ such that

$$M_{\mathcal{B}}^{\mathcal{B}}(N) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Suppose $\mathcal{B} = \{v_1, v_2, v_3\}$. Notice that v_1 is an eigenvector of N, with eigenvalue 0. Also, $Nv_2 = v_1$ and $Nv_3 = v_2$. Because $M_{\mathcal{B}}^{\mathcal{B}}(N)$ is upper triangular, span$\{v_3, Nv_3, N^2v_3\}$ is invariant under N. The sequence v_3, Nv_3, N^2v_3 is an example of a *Jordan chain.* If $v \in \mathcal{V}$, $l \in \mathbb{N}$ and N is nilpotent, we define a Jordan chain to be $\{v, Nv, \ldots, N^{l-1}v\}$, where $N^{l-1}v$ is an eigenvector (and hence $\neq 0$) and $N^l v = 0$. The span of a Jordan chain is an invariant subspace for N.

Theorem 5.1. *Every Jordan chain is linearly independent.*

Proof. The proof will proceed by induction on l. When $l = 1$, the chain is v_1, which is clearly linearly independent since it is an eigenvector, which is by definition different from 0. Next, let $m \in \mathbb{N}$ be arbitrary and assume that each Jordan chain of length m is linearly independent. Consider a Jordan chain of length $m+1$: $\{w, Nw, \dots, N^m w\}$ is a Jordan chain of a nilpotent linear operator N. Suppose there exists $\alpha_0, \dots, \alpha_m$ such that

$$\alpha_0 w + \alpha_1 Nw + \dots + \alpha_m N^m w = 0.$$

Take N of both sides of the equation:

$$\begin{aligned} N(\alpha_0 w + \alpha_1 Nw + \dots + \alpha_m N^m w) &= N(0) \\ \alpha_0 N(w) + \alpha_1 N(Nw) + \dots + \alpha_{m-1} N(N^{m-1} w) + \alpha_m N(N^m w) &= 0 && \text{(by linearity)} \\ \alpha_0 Nw + \alpha_1 N^2 w + \dots + \alpha_{m-1} N^m w + \alpha_m N^{m+1} w &= 0 \\ \alpha_0 Nw + \alpha_1 N^2 w + \dots + \alpha_{m-1} N^m w + 0 &= 0 && \text{(because the chain is Jordan)} \end{aligned}$$

Notice that $\{Nw, N^2 w, \dots, N^m w\}$ is a Jordan chain of N of length m. Hence by the inductive hypothesis, $\{Nw, N^2 w, \dots, N^m w\}$ is linearly independent, so $\alpha_j = 0$ for all $j \in \{0, \dots, m-1\}$. Thus

$$\alpha_m N^m w = 0.$$

Thus as $N^m w \neq 0$, $\alpha_m = 0$. Thence $\alpha_j = 0$ for all $j \in \{0, \dots, m\}$. So $\{w, Nw, \dots, N^{m+1} w\}$ is linearly independent. □

Theorem 5.2. *Let N be nilpotent in $\mathcal{L}(\mathcal{V})$. Then there exists a basis $\mathcal{B}$ of $\mathcal{V}$ that consists of Jordan chains corresponding to N.*

Before we begin the proof, a point of clarification: if $\{w, Nw, \dots, N^{l-1} w\}$ is a Jordan chain, for any $k \in \{0, \dots, l-2\}$, $\{w, Nw, \dots, N^k w\}$ is not a Jordan chain, because $N^{k+1} w \neq 0$, but $\{N^k w, N^{k+1} w, \dots, N^{l-1} w\}$ is.

Proof. Let $\dim \mathcal{V}$. Let N be nilpotent in $\mathcal{L}(\mathcal{V})$. Let $\dim \mathcal{N}(N) = m$. Then there exists $\{v_1 \dots, v_m\} \in \mathcal{V}$ and $q_1, \dots, q_m \in \mathbb{N}$ such that $\{N^{q_1 - 1} v_1, N^{q_m - 1} v_m\}$ is a basis for $\mathcal{N}(N)$. We claim that $v_j, Nv_j \dots N^{q_j - 1} v_j\}, \forall j \in \{1 \dots m\}$ is a basis for $\mathcal{V}$. □