# 8. Orthonormal Triangularization and Diagonalization

**8.1 Definition:** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For a linear map $L : U \to U$, where $U$ is a finite dimensional inner product space over $\mathbb{F}$, we say that $L$ is **orthonormally triangularizable** when there exists an othonormal basis $\mathcal{A}$ for $U$ such that $[L]_\mathcal{A}$ is upper triangular. For a matrix $A \in M_{n\times n}(\mathbb{F})$, we say that $A$ is **orthonormaly triangularizable** when there exists a matrix $P \in M_{n\times n}(\mathbb{F})$ with $P^*P = I$ such that $P^*AP$ is upper triangular. Most books do not use the term *orthonormally triangularizable* but, instead, in the case that $\mathbb{F} = \mathbb{R}$ they use the term **orthogonally triangularizable** and when $\mathbb{F} = \mathbb{C}$ they use the term **unitarily triangularizable**.

**8.2 Theorem:** *Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}$ be an orthonormal basis for $U$. Let $L : U \to U$ be a linear map and let $A = [L]_\mathcal{A}$. Then $L$ is orthonormally triangularlizable if and only if $A$ is orthonormally triangularizable.*

Proof: The proof is left as an exercise.

**8.3 Theorem:** (*Schur*) *Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be linear. Then $L$ is orthonormally triangularizable if and only if the characteristic polynomial $f_L(x)$ splits.*

Proof: Suppose that $L$ is orthonormally trianglarizable. Choose an orthonormal basis $\mathcal{A}$ for $U$ such that $[L]_\mathcal{A}$ is upper triangular. Let $T = [L]_\mathcal{A} \in M_n(\mathbb{F})$. Then $f_L(x) = f_T(x)$, and since $T$ is upper triangular we have $f_T(x) = (-1)^n \prod_{k=1}^{n} \left(x - T_{k,k}\right)$, which splits.

Conversely, suppose that $f_L(x)$ splits. Choose any orthonormal basis $\mathcal{A}$ for $U$ and let $A = [L]_\mathcal{A}$. Since $f_A(x) = f_L(x)$, we know that $f_A(x)$ splits. We shall show, by induction on $n$, that for any matrix $A \in M_n(\mathbb{F})$ for which $f_A(x)$ splits, there exists a matrix $P \in M_n(\mathbb{F})$ with $P^*P = I$ such that $P^*AP$ is upper triangular. When $n = 1$, the $1 \times 1$ matrix $A$ is already upper triangular and we can take $P$ to be the $1 \times 1$ identity matrix. Fix $n \geq 2$, let $A \in M_n$, suppose that $f_A(x)$ splits, and suppose, inductively that for every matrix $B \in M_{n-1}(\mathbb{F})$ for which $f_B(x)$ splits, we can find a matrix $Q \in M_{n-1}(\mathbb{F})$ with $Q^*Q = I$ such that $Q^*BQ$ is upper triangular. Since $f_A(x)$ splits, $A$ has an eigenvalue. Let $\lambda_1$ be an eigenvalue of $A$ and let $u_1 \in \mathbb{F}^n$ be a corresponding eigenvector with $\|u_1\| = 1$. Extend $\{u_1\}$ to an orthonormal basis $\{u_1, u_2, \cdots, u_n\}$ for $\mathbb{F}^n$ and let $R = (u_1, u_2, \cdots, u_n) \in M_n(\mathbb{F})$. Note that since $\{u_1, u_2, \cdots, u_n\}$ is orthonormal we have $R^*R = I$. The $k^{\text{th}}$ entry of the first column of the matrix $R^*AR$ is equal to

$$(R^*AR)_{k,1} = {e_k}^* R^*AR\, e_1 = {u_k}^* A u_1 = \langle Au_1, u_k\rangle = \langle \lambda_1 u_1, u_k\rangle = \lambda_1 \delta_{k,1}$$

so we have

$$R^*AR = \begin{pmatrix} \lambda_1 & x^T \\ 0 & B \end{pmatrix}$$

for some $x \in \mathbb{F}^{n-1}$ and some $B \in M_{n-1}(\mathbb{F})$. Since $f_A(x) = f_{R^*AR}(x) = -(x - \lambda_1) f_B(x)$, so we see that $f_B(x)$ splits. By the induction hypothesis, we can choose $Q \in M_{n-1}(\mathbb{F})$ with $Q^*Q = I$ such that $Q^*BQ$ is upper triangular. Letting $P = R\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$, we have

$$P^*AP = \begin{pmatrix} 1 & 0 \\ 0 & Q^* \end{pmatrix}\begin{pmatrix} \lambda_1 & x^T \\ 0 & B \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & x^T Q \\ 0 & Q^*BQ \end{pmatrix}$$

which is upper triangular, and it is easy to check that $P^*P = I$.

**8.4 Definition:** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For a linear map $L : U \to U$, where $U$ is a finite dimensional inner product space over $\mathbb{F}$, we say that $L$ is **orthonormally diagonalizable** when there exists an orthonormal basis $\mathcal{A}$ for $U$ such that $[L]_{\mathcal{A}}$ is diagonal. For a matrix $A \in M_n(\mathbb{F})$, we say that $A$ is **orthonormally diagonalizable** when there exists a matrix $P \in M_n(\mathbb{F})$ with $P^*P = I$ such that $P^*AP$ is diagonal. Most books do not use the term *orthonormally diagonalizable* but, instead, when $\mathbb{F} = \mathbb{R}$ they use the term **orthogonally diagonalizable** and when $\mathbb{F} = \mathbb{C}$ they use the term **unitarily diagonalizable**.

**8.5 Theorem:** *Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}$ be an orthonormal basis for $U$. Let $L : U \to U$ be a linear map and let $A = [L]_{\mathcal{A}}$. Then $L$ is orthonormally diagonalizable if and only if $A$ is orthonormally diagonalizable.*

Proof: The proof is left as an exercise.

**8.6 Definition:** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For a linear map $L : U \to U$, where $U$ is an inner product space over $\mathbb{F}$, we say that $L$ is **normal** when the adjoint $L^*$ exists and $L^*L = LL^*$. For a matrix $A \in M_n(\mathbb{F})$, we say that $A$ is **normal** when $A^*A = AA^*$. Note that when $U$ is finite dimensional and $\mathcal{A}$ is an orthonormal basis for $U$, the map $L$ is normal if and only if its matrix $[L]_{\mathcal{A}}$ is normal.

**8.7 Theorem:** (*Diagonalization of Normal Matrices) Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be linear. Then $L$ is orthonormally diagonalizable if and only if $L$ is normal and the characteristic polynomial $f_L(x)$ splits.*

Proof: Suppose first that $L$ is orthonormally diagonalizable. Choose an orthonormal basis $\mathcal{A}$ for $U$ so that $[L]_{\mathcal{A}}$ is diagonal, say $[L]_{\mathcal{A}} = D = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \in M_{n\times n}(\mathbb{F})$. Then $f_L(x)$ splits because $f_L(x) = f_D(x) = (-1)^n \prod\limits_{i=1}^{n} (x - \lambda_i)$, and $L$ is normal because $D$ is normal, indeed

$$\begin{aligned} D^*D &= \text{diag}\big(\overline{\lambda_1}, \cdots, \overline{\lambda_n}\big)\text{diag}\big(\lambda_1, \cdots, \lambda_n\big) = \text{diag}\big(|\lambda_1|^2, \cdots, |\lambda_n|^2\big) \\ &= \text{diag}\big(\lambda_1, \cdots, \lambda_n\big)\text{diag}\big(\overline{\lambda_1}, \cdots, \overline{\lambda_n}\big) = D\,D^*. \end{aligned}$$

Conversely, suppose that $L$ is normal and that $f_L(x)$ splits. Since $f_L(x)$ splits, by Schur's Theorem we can orthonormally triangularize $L$. Choose an orthonormal basis $\mathcal{A}$ for $U$ so that $[L]_{\mathcal{A}}$ is upper triangular. Let $T = [L]_{\mathcal{A}} \in M_n(\mathbb{F})$. Since $L$ is normal, it follows that $T$ is normal. Since $T$ is normal and upper triangular, it follows that $T$ is in fact diagonal: Indeed, the diagonal entries of $T^*T$ and of $TT^*$ are given by $(T^*T)_{k,k} = \sum\limits_{i=1}^{n} (T^*)_{k,i} T_{i,k} = \sum\limits_{i=1}^{n} \overline{T}_{i,k} T_{i,k} = \sum\limits_{i=1}^{n} |T_{i,k}|^2$ and $(TT^*)_{k,k} = \sum\limits_{i=1}^{n} T_{k,i}(T^*)_{i,k} = \sum\limits_{i=1}^{n} T_{k,i}\overline{T}_{k,i} = \sum\limits_{i=1}^{n} |T_{k,i}|^2$. Since $T$ is upper triangular, so that $T_{i,j} = 0$ whenever $i > j$, these expressions simplify to

$$(T^*T)_{k,k} = \sum_{i=1}^{k} |T_{i,k}|^2 \quad\text{and}\quad (TT^*)_{k,k} = \sum_{i=k}^{n} |T_{k,i}|^2.$$

Comparing these diagonal entries, we find that

$$\begin{aligned} (T^*T)_{1,1} = (TT^*)_{1,1} &\implies |T_{1,1}|^2 = |T_{1,1}|^2 + |T_{1,2}|^2 + \cdots + |T_{1,n}|^2 \implies T_{1,i} = 0 \text{ for } i > 1, \\ (T^*T)_{2,2} = (TT^*)_{2,2} &\implies |T_{2,2}|^2 = |T_{2,2}|^2 + |T_{2,3}|^2 + \cdots + |T_{2,n}|^2 \implies T_{2,i} = 0 \text{ for } i > 2, \end{aligned}$$

and so on, so that $T$ is diagonal.

**8.8 Definition:** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For a linear map $L : U \to U$, where $U$ is an inner product space over $\mathbb{F}$, we say that $L$ is **unitary** when the adjoint $L^*$ exists and we have $L^*L = I = LL^*$. For a matrix $A \in M_n(\mathbb{F})$, we say that $A$ is **unitary** when $A^*A = I$. Note that when $U$ is finite dimensional and $\mathcal{A}$ is an orthonormal basis for $U$, the map $L$ is unitary if and only if its matrix $[L]_\mathcal{A}$ is unitary. When $\mathbb{F} = \mathbb{R}$ the term *unitary* can be replaced by the term **orthogonal**.

**8.9 Theorem:** *Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be a linear map. Then the following are equivalent.*
*(1) $L$ is unitary,*
*(2) $L$ preserves inner product, that is $\big\langle L(x), L(y)\big\rangle = \langle x, y\rangle$ for all $x, y \in U$,*
*(3) $L$ preserves norm, that is $\big\|L(x)\big\| = \|x\|$ for all $x \in U$.*

Proof: First we show that (1) is equivalent to (2). Suppose that $L$ is unitary. Then for $x, y \in U$ we have $\big\langle Lx, Ly\big\rangle = \big\langle x, L^*Ly\big\rangle = \langle x, Iy\rangle = \langle x, y\rangle$, and so $L$ preserves inner product. Conversely, suppose that $L$ preserves inner product. Let $y \in U$. Then for all $x \in U$ we have $\big\langle x, L^*Ly\big\rangle = \big\langle Lx, Ly\big\rangle = \langle x, y\rangle$ Since $\big\langle x, L^*Ly\big\rangle = \big\langle x, y\big\rangle$ for all $x \in U$, it follows (from Theorem 5.8) that $L^*Ly = y$. Since $L^*Ly = y$ for all $y \in U$, it follows that $L^*L = I$, and so $L$ is unitary.

Next we shall show that (2) is equivalent to (3). Suppose that $L$ preserves inner product. Then for $x \in U$ we have

$$\big\|Lx\big\|^2 = \big\langle Lx, Lx\big\rangle = \langle x, x\rangle = \|x\|^2$$

so that $L$ preserves norm. Conversely, suppose that $L$ preserves norm. Then, using the Polarization Identity and the linearity of $L$, for $x, y \in U$ we have

$$\begin{aligned}\langle Lx, Ly\rangle &= \tfrac{1}{4}\Big(\big\|Lx+Ly\big\|^2 + i\big\|Lx+iLy\big\|^2 - \big\|Lx-Ly\big\|^2 - i\big\|Lx-iLy\big\|^2\Big)\\ &= \tfrac{1}{4}\Big(\big\|L(x+y)\big\|^2 + i\big\|L(x+iy)\big\|^2 - \big\|L(x-y)\big\|^2 - i\big\|L(x-iy)\big\|^2\Big)\\ &= \tfrac{1}{4}\Big(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\Big) = \langle x, y\rangle.\end{aligned}$$

**8.10 Theorem:** (*Diagonalization of Unitary Maps) Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $L : U \to U$ be linear. Then $L$ is orthonormally diagonalizable and all its eigenvalues have norm 1 if and only if $L$ is unitary and $f_L(x)$ splits.*

Proof: Suppose that $L$ is orthonormally diagonalizable and that all of its eigenvalues have norm 1. Choose an orthonormal basis $\mathcal{A}$ for $U$ so that $[L]_\mathcal{A} = D = \text{diag}\big(\lambda_1, \cdots, \lambda_n\big)$ where $|\lambda_i| = 1$ for $1 \le i \le n$. Then $f_L(x)$ splits since $f_L(x) = f_D(x) = (-1)^n \prod_{i=1}^{n}(x - \lambda_i)$, and $L$ is unitary since $D$ is unitary, indeed

$$D^*D = \text{diag}\big(\overline{\lambda_1}, \cdots, \overline{\lambda_n}\big)\text{diag}\big(\lambda_1, \cdots, \lambda_n\big) = \text{diag}\big(|\lambda_1|^2, \cdots, |\lambda_n|^2\big) = I.$$

Conversely, suppose that $L$ is unitary and that $f_L(x)$ splits. Since $L$ is unitary, it is also normal because $L^*L = I = LL^*$. Since $L$ is normal and $f_L(x)$ splits, $L$ is orthonormally diagonalizable. Choose an orthonormal basis $\mathcal{A}$ for $U$ such that $[L]_\mathcal{A} = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. Since $L$ is unitary, so is $D$, and so we have $I = D^*D = \text{diag}\big(|\lambda_1|^2, \cdots, |\lambda_n|^2\big)$, and hence $|\lambda_i| = 1$ for all indices $i$.

**8.11 Definition:** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For a linear map $L : U \to U$, where $U$ is an inner product space over $\mathbb{F}$, we say that $L$ is **Hermitian** (or **self-adjoint**) when the adjoint $L^*$ exists and we have $L^* = L$. For a matrix $A \in M_n(\mathbb{F})$, we say that $A$ is **Hermitian** (or **self-adjoint**) when $A^* = A$. Note that when $U$ is finite dimensional and $\mathcal{A}$ is an orthonormal basis for $U$, the map $L$ is Hermitian if and only if its matrix $A$ is Hermitian. When $\mathbb{F} = \mathbb{R}$, the terms *Hermitian* and *self-adjoint* can be replaced by the term **symmetric**.

**8.12 Theorem:** (*Diagonalization of Hermitian Maps) Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $L : U \to U$ be linear. Then $L$ is orthonormally diagonalizable and all its eigenvalues are real if and only if $L$ is Hermitian.*

Proof: Suppose that $L$ is orthonormally diagonalizable and all of its eigenvalues are real. Choose an orthonormal basis $\mathcal{A}$ for $U$ so that $[L]_\mathcal{A} = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ with each $\lambda_i \in \mathbb{R}$. Then $L$ is Hermitian since $D$ is Hermitian, indeed

$$D^* = \text{diag}\big(\overline{\lambda_1}, \cdots, \overline{\lambda_n}\big) = \text{diag}\big(\lambda_1, \cdots, \lambda_n\big) = D.$$

Conversely, suppose that $L$ is Hermitian, that is $L^* = L$. Since $L$ is Hermitian, it is also normal, indeed we have $L^*L = L^2 = LL^*$. Choose any orthonormal basis $\mathcal{A}$ for $U$ and let $A = [L]_\mathcal{A} \in M_n(\mathbb{F})$. Since $L^* = L$, we also have $A^* = A$. Whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we can consider $A$ as a matrix with complex entries $A \in M_n(\mathbb{C})$. Let $\lambda \in \mathbb{C}$ be a complex eigenvalue of $A$ and let $u \in \mathbb{C}^n$ be an eigenvector, scaled so that $\|u\| = 1$. Using the fact that $A^* = A$, we have

$$\lambda = \lambda\langle u, u\rangle = \langle \lambda u, u\rangle = \langle Au, u\rangle = \langle u, A^*u\rangle = \langle u, Au\rangle = \langle u, \lambda u\rangle = \overline{\lambda}\langle u, u\rangle = \overline{\lambda}.$$

so the eigenvalues of $A \in M_n(\mathbb{C})$ are all real. Thus the eigenvalues of $L$ are all real (even when $\mathbb{F} = \mathbb{C}$), and $f_L(x) = f_A(x)$ splits (even when $\mathbb{F} = \mathbb{R}$). Since $L$ is normal and $f_L(x)$ splits, $L$ is orthonormally diagonalizable.

**8.13 Example:** Let $U$ be a finite dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : U \to U$ be a linear map. Then $L$ is an orthogonal projection (onto some subspace $U_0 \subseteq U$) when there exists an orthonormal basis $\mathcal{A}$ for $U$ (obtained by extending an orthonormal basis $\mathcal{A}_0$ for $U_0$ to all of $U$) such that

$$[L]_\mathcal{A} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus we see that

$L$ is an orthogonal projection map

$\iff$ $L$ is orthonormally diagonalizable and all of its eigenvalues are 0 or 1

$\iff$ $L^* = L$ and $L^2 = L$

because when $[L]_\mathcal{A} = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ we have

$$L^2 = L \iff D^2 = D \iff {\lambda_i}^2 = \lambda_i \text{ for all } i \iff \lambda_i \in \{0, 1\} \text{ for all } i.$$

Similarly, $L$ is a reflection (onto some subspace $U_0 \subseteq U$) when there is an orthonormal basis $\mathcal{A}$ for $U$ such that

$$[L]_\mathcal{A} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

and so we see that

$L$ is a reflection map

$\iff$ $L$ is orthogonally diagonalizable and all of its eigenvalues are 1 or -1

$\iff$ $L^* = L$ and $L^2 = I \iff L^* = L$ and $L^*L = I$.

**8.14 Theorem:** (*Singular Value Decomposition*) *Let $U$ and $V$ be finite dimensional inner product spaces over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $L:U\to V$ be a linear map. Then there exist orthonormal bases $\mathcal{A}$ and $\mathcal{B}$ for $U$ and $V$ such that $[L]_{\mathcal{B}}^{\mathcal{A}}$ is of the block form*

$$[L]_{\mathcal{B}}^{\mathcal{A}}=\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix} \quad \text{with} \quad D=\operatorname{diag}\big(\sigma_1,\sigma_2,\cdots,\sigma_r\big)$$

*with $r=\operatorname{rank}(L)$ and $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0$. The positive real numbers $\sigma_i$ are unique.*

Proof: First, let us prove that the numbers $\sigma_1,\cdots,\sigma_r$ are uniquely determined from $L$. Suppose $\mathcal{A}=\{u_1,\cdots,u_k\}$ and $\mathcal{B}=\{v_1,\cdots,v_l\}$ are orthonormal bases for $U$ and $V$ such that $[L]_{\mathcal{B}}^{\mathcal{A}}=\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}\in M_{l\times k}(\mathbb{F})$ where $D=\operatorname{diag}(\sigma_1,\cdots,\sigma_r)$ with $\sigma_1\geq\cdots\geq\sigma_r>0$. Since $[L]_{\mathcal{B}}^{\mathcal{A}}=\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}\in M_{l\times k}(\mathbb{F})$, we must have $L(u_i)=\sigma_i v_i$ for $1\leq i\leq r$ and $L(u_i)=0$ for $r<i\leq k$. Since $[L^*]_{\mathcal{A}}^{\mathcal{B}}=\begin{pmatrix} D & 0\\ 0 & 0\end{pmatrix}\in M_{k\times l}(\mathbb{F})$, we must have $L^*(v_i)=\sigma_i u_i$ for $1\leq i\leq r$ and $L^*(v_i)=0$ for $r<i\leq l$. It follows that

$$L^*L(u_i)=L^*(\sigma_i v_i)=\sigma_i L^*(v_i)={\sigma_i}^2 u_i$$

for $1\leq i\leq r$ and $L^*L(u_i)=L^*(0)=0$ for $r<i\leq k$. Thus for $1\leq i\leq r$, the values $\lambda_i={\sigma_i}^2$ must be the non-zero eigenvalues of $L^*L$, and they must be positive and real, and the vectors $u_i$ must be corresponding eigenvectors.

Next, let us prove that there do indeed exist orthonormal bases which put $L$ into the desired form. Note that $\operatorname{Null}(L^*L)=\operatorname{Null}(L)$, indeed for $x\in U$ we have

$$Lx=0\Longrightarrow L^*Lx=0\ \text{, and}$$
$$L^*Lx=0\Longrightarrow\big\langle x,L^*Lx\big\rangle=0\Longrightarrow\big\langle Lx,Lx\big\rangle=0\Longrightarrow Lx=0.$$

In particular, we have $\operatorname{rank}(L^*L)=\operatorname{rank}(L)=r$. Also note that $L^*L$ is Hermitian since $(L^*L)^*=L^*L^{**}=L^*L$, and so $L^*L$ is orthonormaly diagonalzable and its eigenvalues are all real. Furthermore, note that the eigenvalues of $L^*L$ are all non-negative because if $\lambda$ is an eigenvalue of $L^*L$ and $u$ is a corresponding unit eigenvector so that we have $Lu=\lambda u$ and $\|u\|=1$, then we have

$$\lambda=\lambda\|u\|^2=\lambda\langle u,u\rangle=\langle\lambda u,u\rangle=\big\langle L^*Lu,u\big\rangle=\overline{\big\langle u,L^*Lu\big\rangle}=\overline{\big\langle Lu,Lu\big\rangle}=\big\|Lu\big\|^2\geq 0.$$

Let $\lambda_1,\cdots,\lambda_k$ be the eigenvalues of $L^*L$ with $\lambda_1\geq\cdots\geq r>0$ and $\lambda_{r+1}=\cdots=\lambda_k=0$. let $\sigma_i=\sqrt{\lambda_i}$ for $1\leq i\leq k$ so that $\sigma_1\geq\cdots\geq\sigma_r>0$ and $\sigma_{r+1}=\cdots=\sigma_k=0$. Choose an orthonormal basis $\mathcal{A}$ for $U$ so that $[L^*L]_{\mathcal{A}}=\operatorname{diag}(\lambda_1,\cdots,\lambda_k)$. Note that $\{u_{r+1},\cdots,u_k\}$ is an orthonormal basis for $\operatorname{Null}(L^*L)=\operatorname{Null}(L)$ and $\{u_1,\cdots,u_r\}$ is an orthonormal basis for $\operatorname{Null}(L)^\perp$. For $1\leq i\leq r$, let $v_i=\frac{1}{\sigma_i}L(u_i)$. Note that $\{v_1,\cdots,v_r\}$ is orthonormal since

$$\begin{aligned}\langle v_i,v_j\rangle&=\big\langle \tfrac{1}{\sigma_i}L(u_i)\,,\,\tfrac{1}{\sigma_j}L(u_j)\big\rangle=\tfrac{1}{\sigma_i\overline{\sigma_j}}\big\langle L(u_i),L(u_j)\big\rangle=\tfrac{1}{\sigma_i\sigma_j}\big\langle u_i,L^*L(u_j)\big\rangle\\&=\tfrac{1}{\sigma_i\sigma_j}\big\langle u_i,\lambda_j u_j\big\rangle=\tfrac{\overline{\lambda_j}}{\sigma_i\sigma_j}\langle u_i,u_j\rangle=\tfrac{\lambda_j}{\sigma_i\sigma_j}\delta_{i,j}=\delta_{i,j}.\end{aligned}$$

since $\frac{\lambda_j}{\sigma_j\sigma_j}=1$. Extend $\{v_1,\cdots,v_r\}$ to an orthonormal basis for $V$, and note that $[L]_{\mathcal{B}}^{\mathcal{A}}$ is of the desired form.

**8.15 Definition:** The **singular values** of a linear map $L:U\to V$ are the square roots of the eigenvalues of the map $L^*L$. The **singular values** of a matrix $A$ are the square roots of the eigenvalues of the matrix $A^*A$.