

# Notes on Singular Value Decomposition

Z. Zhang \*

February 20, 2020 (revised on February 28, 2021)

## Abstract

We collect a few elementary facts about the singular value decomposition (SVD) of matrices. In particular, we present three approaches used by different authors in the history to establish the existence of SVD.

**Notation.** Throughout the document,  $\|\cdot\|$  stands for the 2-norm for vectors and matrices. In inline equations, we use the MATLAB-style notation  $[a; b]$  to denote a vertical array with  $a$  and  $b$  being its entries.

## 1 Eigenvalue decomposition

**Theorem 1.** *Given any Hermitian matrix  $A \in \mathbb{C}^{n \times n}$ , there exists a unitary matrix  $U \in \mathbb{C}^{n \times n}$  and a diagonal matrix  $\Lambda \in \mathbb{R}^{n \times n}$  such that*

$$A = U\Lambda U^H.$$

*If  $A$  is real, then we can require that  $U$  is real. Indeed,  $\Lambda_{1,1}, \dots, \Lambda_{n,n}$  are the eigenvalues of  $A$ , multiplicity included, and the  $j$ -th column of  $U$  is an eigenvector of  $A$  associated with  $\Lambda_{j,j}$ .*

**Definition 1.** Let  $A \in \mathbb{C}^{n \times n}$  be an Hermitian matrix.

1.  $U\Lambda U^H$  is called an eigenvalue decomposition of  $A$ , provided that  $A = U\Lambda U^H$ ,  $U \in \mathbb{C}^{n \times n}$  is a unitary matrix, and  $\Lambda \in \mathbb{R}^{n \times n}$  is a diagonal matrix.
2.  $U\Lambda U^H$  is called a compact eigenvalue decomposition of  $A$ , provided that  $A = U\Lambda U^H$ ,  $U \in \mathbb{C}^{n \times r}$  is a matrix with  $U^H U = I_r$ , and  $\Lambda \in \mathbb{R}^{r \times r}$  is a diagonal matrix whose diagonal entries are nonzero.

## 2 Singular value decomposition (SVD)

**Definition 2.** Let  $A \in \mathbb{C}^{m \times n}$  be a matrix with  $\text{rank}(A) = r$ .

1.  $U\Sigma V^H$  is called an singular value decomposition of  $A$ , provided that  $A = U\Sigma V^H$ ,  $U \in \mathbb{C}^{m \times m}$  and  $V \in \mathbb{C}^{n \times n}$  are unitary matrices, and  $\Sigma \in \mathbb{R}^{m \times n}$  is a matrix whose first  $r$  diagonal entries (if  $r \geq 1$ ) are positive while all the other entries are zero.

---

\*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China (zaikun.zhang@polyu.edu.hk).

2. When  $r \geq 1$ ,  $U\Sigma V^H$  is called a compact (or *reduced*) singular value decomposition of  $A$ , provided that  $A = U\Sigma V^H$ ,  $U \in \mathbb{C}^{m \times r}$  and  $V \in \mathbb{C}^{n \times r}$  are matrices with  $U^H U = V^H V = I_r$ , and  $\Sigma \in \mathbb{R}^{r \times r}$  is a diagonal matrix whose diagonal entries are positive.

**Remark 1.** Let  $U\Sigma V^H$  be a singular value decomposition of  $A$  and  $\sigma_i = \Sigma_{i,i}$  ( $1 \leq i \leq \min\{m, n\}$ ). Then  $\sigma_1, \dots, \sigma_r$  are called the (nonzero) singular values of  $A$ . It is often convenient to regard  $\sigma_{r+1} = \dots = \sigma_{\min\{m, n\}} = 0$  also as singular values of  $A$ .

**Remark 2.** If  $U\Sigma V^H$  is a (compact) singular value decomposition of  $A$ , then  $AV = U\Sigma$  and  $A^H U = V\Sigma$ . Let  $\sigma_i = \Sigma_{i,i}$ ,  $u_i$  be the  $i$ -th column of  $U$ , and  $v_i$  be the  $i$ -th column of  $V$ . Then  $Av_i = \sigma_i u_i$  and  $A^H u_i = \sigma_i v_i$ ;  $u_i$  and  $v_i$  are called a pair of left and right singular vectors of  $A$  associated with the singular value  $\sigma_i$ .

**Remark 3.** If  $U\Sigma V^H$  is a compact singular value decomposition of  $A \in \mathbb{C}^{m \times n}$ , then we can extend it to a singular value decomposition

$$(U \tilde{U}) \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V^H \\ \tilde{V}^H \end{pmatrix},$$

where  $\tilde{U} \in \mathbb{C}^{m \times (m-r)}$  is any matrix such that  $(U \tilde{U})$  is unitary, and  $\tilde{V} \in \mathbb{C}^{n \times (n-r)}$  is any matrix such that  $(V \tilde{V})$  is unitary. Conversely, if  $U\Sigma V^H$  is a singular value decomposition of  $A \neq 0$ , we can obtain a compact singular value decomposition by dropping the zero diagonal entries of  $\Sigma$  and the corresponding columns of  $U$  and  $V$ .

## 2.1 Uniqueness of (compact) SVD

**Theorem 2.** Suppose that  $U\Sigma V^H$  be a compact singular value decomposition of a matrix  $A$ .

1.  $U\Sigma^2 U^H$  is a compact eigenvalue decomposition of  $AA^H$ , and hence the diagonal entries of  $\Sigma^2$  are the positive eigenvalues of  $AA^H$ , multiplicity included.
2.  $V\Sigma^2 V^H$  is a compact eigenvalue decomposition of  $A^H A$ , and hence the diagonal entries of  $\Sigma^2$  are the positive eigenvalues of  $A^H A$ , multiplicity included.

**Lemma 1.** Consider matrices  $A \in \mathbb{C}^{n \times n}$  and  $\Lambda \in \mathbb{C}^{n \times n}$  with  $\Lambda$  being diagonal.

1.  $\Lambda A = A\Lambda$  if and only if  $A_{i,j} = 0$  for any  $i$  and  $j$  such that  $\Lambda_{i,i} \neq \Lambda_{j,j}$ .
2. If  $\Lambda$  is nonnegative and  $\Lambda A = A\Lambda$ , then  $\Lambda^p A = A\Lambda^p$  for any  $p \geq 0$ .

**Proof.** This is because  $\Lambda A = A\Lambda$  if and only if  $\Lambda_{i,i} A_{i,j} = A_{i,j} \Lambda_{j,j}$  for any  $i, j \in \{1, \dots, n\}$ .  $\square$

**Lemma 2.** Let  $U_1 \in \mathbb{C}^{n \times r}$  and  $U_2 \in \mathbb{C}^{n \times r}$  satisfy  $U_1^H U_1 = U_2^H U_2 = I_r$  and  $\text{range}(U_1) = \text{range}(U_2)$ .

1.  $U_1 U_1^H = U_2 U_2^H$ , both being the orthogonal projection onto  $\text{range}(U_1) = \text{range}(U_2)$ .
2.  $W = U_1^H U_2$  is a unitary matrix and  $U_1 W = U_2$ .

**Proof.**  $U_1^H U_1 = I_r$  ensures that  $U_1 U_1^H$  is the orthogonal projection onto  $\text{range}(U_1) = \text{range}(U_2)$  (see, e.g., [3]). In addition,  $U_1 W = U_1 U_1^H U_2 = U_2$ , and  $W^H W = U_2^H U_1 W = U_2^H U_2 = I_r$ .  $\square$

**Theorem 3.** Let  $U_i \in \mathbb{C}^{m \times r}$  and  $V_i \in \mathbb{C}^{n \times r}$  satisfy  $U_i^H U_i = V_i^H V_i = I_r$  ( $i = 1, 2$ ), and  $\Sigma \in \mathbb{C}^{r \times r}$  be a diagonal matrix whose diagonal entries are positive. Then  $U_1 \Sigma V_1^H = U_2 \Sigma V_2^H$  if and only if there exists a unitary matrix  $W \in \mathbb{C}^{r \times r}$  such that  $U_2 = U_1 W$ ,  $V_2 = V_1 W$ , and  $\Sigma W = W \Sigma$ .

**Proof.** The “if” part is trivial, so we focus on the “only if” part. Assuming  $U_1 \Sigma V_1^H = U_2 \Sigma V_2^H$ , We will show that  $W = U_1^H U_2 \in \mathbb{C}^{r \times r}$  fulfills all the desired requirements. Observe that both  $\Sigma V_1^H$  and  $\Sigma V_2^H$  have full column rank. Hence

$$\text{range}(U_1) = \text{range}(U_1 \Sigma V_1^H) = \text{range}(U_2 \Sigma V_2^H) = \text{range}(U_2).$$

By Lemma 2,  $W$  is a unitary matrix and  $U_1 W = U_2$ . It remains to show that  $\Sigma W = W \Sigma$  and  $V_2 = W V_1$ . Recalling that  $V_1^H V_1 = V_2^H V_2 = I_r$ , we have

$$U_1 \Sigma^2 U_1^H = (U_1 \Sigma V_1^H)(U_1 \Sigma V_1^H)^H = (U_2 \Sigma V_2^H)(U_2 \Sigma V_2^H)^H = U_2 \Sigma^2 U_2^H.$$

Hence

$$\Sigma^2 W = \Sigma^2 U_1^H U_2 = U_1^H (U_1 \Sigma^2 U_1^H) U_2 = U_1^H (U_2 \Sigma^2 U_2^H) U_2 = U_1^H U_2 \Sigma^2 = W \Sigma^2$$

Thus  $\Sigma W = W \Sigma$  by Lemma 1. Finally, since  $V_1 \Sigma U_1^H = (U_1 \Sigma V_1^H)^H = (U_2 \Sigma V_2^H)^H = V_2 \Sigma U_2^H$ ,

$$V_2 = (V_2 \Sigma U_2^H)(U_2 \Sigma^{-1}) = (V_1 \Sigma U_1^H)(U_2 \Sigma^{-1}) = V_1 \Sigma W \Sigma^{-1} = V_1 W \Sigma \Sigma^{-1} = V_1 W.$$

The proof is complete.  $\square$

## 2.2 Existence of SVD

### 2.2.1 Jordan’s deflation approach [4]

**Lemma 3.** Given a nonzero matrix  $A \in \mathbb{C}^{m \times n}$ , let  $(u, v) \in \mathbb{C}^m \times \mathbb{C}^n$  be a solution of

$$\max \{ \Re(x^H A y) : \|x\| = \|y\| = 1, x \in \mathbb{C}^m, y \in \mathbb{C}^n \},$$

and  $\sigma = \Re(u^H A v)$ . Then  $Av = \sigma u$ ,  $A^H u = \sigma v$ , and  $\sigma > 0$ .

**Proof.** Since  $A \neq 0$ , it is obvious that  $\sigma > 0$ . Hence  $Av \neq 0$ . According to the definition of  $u$ ,

$$\Re(u^H A v) \geq \Re((Av / \|Av\|)^H A v) = \|Av\| = \|u\| \|Av\|.$$

By the Cauchy-Schwarz inequality, there exists a scalar  $\lambda > 0$  such that  $\lambda u = Av$ . Hence

$$\sigma = \Re(u^H A v) = \Re(\lambda \|u\|^2) = \lambda.$$

Thus  $Av = \sigma u$ . Similarly, we can prove  $A^H u = \sigma v$  using the fact that

$$\Re(u^H A v) \geq \Re(u^H A (A^H u / \|A^H u\|)) = \|A^H u\| = \|A^H u\| \|v\|.$$

$\square$

**Remark 4.** When Lemma 3 is applied in the proof of Theorem 4 later, we only need the existence of unit vectors  $u \in \mathbb{C}^m$ ,  $v \in \mathbb{C}^n$ , and a scalar  $\sigma > 0$  such that  $Av = \sigma u$  and  $A^H u = \sigma v$ . The existence can be established in other ways.

1. Let  $\sigma = (\lambda_{\max}(AA^H))^{\frac{1}{2}} > 0$ ,  $u \in \mathbb{C}^{m \times m}$  be an eigenvector of  $AA^H$  associated with  $\lambda_{\max}(AA^H)$ , and  $v = A^H u / \sigma$ . Then  $Av = AA^H u / \sigma = \sigma^2 u / \sigma = \sigma u$ , and  $A^H u = \sigma v$ . This is the approach used in the proofs of [8, Theorem 4.1] and [5, Theorem 1].
2. Let  $\sigma = \lambda_{\max}(J) > 0$  with  $J$  being the Jordan-Wielandt form of  $A$  (see (6)),  $w \in \mathbb{C}^{m+n}$  be an eigenvector associated with  $\sigma$ ,  $x \in \mathbb{C}^m$  consist of the first  $m$  entries of  $w$ , and  $y \in \mathbb{C}^n$  consist of the last  $n$ . Then we can verify that  $Ay = \sigma x$  and  $A^H x = \sigma y$ . Meanwhile,  $A(-y) = -\sigma x$ , and  $A^H x = -\sigma(-y)$ , making  $[x; -y]$  an eigenvector of  $J$  associated with  $-\sigma \neq \sigma$ . Since eigenvectors for different eigenvalues are orthogonal, we have  $x^H x - y^H y = 0$ . Thus  $\|x\| = \|y\|$ , which are nonzero since  $w \neq 0$ . Finally, let  $u = x/\|x\|$  and  $v = y/\|y\|$ .

**Theorem 4.** Any  $A \in \mathbb{C}^{m \times n}$  has a singular value decomposition  $U\Sigma V^H$  as defined in Definition 2.

**Proof.** Assume without loss of generality that  $A \neq 0$ . We prove by an induction on  $\min\{m, n\}$ .

1. If  $\min\{m, n\} = 1$ , then  $A$  is either a row or a column. If  $A$  is a column, let  $U$  be a unitary matrix whose first column is  $A/\|A\|$ ,  $\Sigma = e_1$  (i.e., the first canonical coordinate vector), and  $V = \|A\|$ . Then  $U\Sigma V^H$  is a singular value decomposition of  $A$ . If  $A$  is a row, the decomposition can be found similarly.

2. Assume that the conclusion holds when  $\min\{m, n\} = k$ . Let us consider the scenario where  $\min\{m, n\} = k + 1$ . Let  $A$  be a matrix in  $\mathbb{C}^{m \times n}$ . By Lemma 3, there exist unit vectors  $u \in \mathbb{C}^m$ ,  $v \in \mathbb{C}^n$ , and a scalar  $\sigma > 0$  such that

$$Av = \sigma u, \quad A^H u = \sigma v. \quad (1)$$

Let  $U \in \mathbb{C}^{m \times m}$  be a unitary matrix whose first column is  $u$ , and  $V \in \mathbb{C}^{n \times n}$  be a unitary matrix whose first column is  $v$ . It is then straightforward to check that

$$U^H AV = \begin{pmatrix} \sigma & 0 \\ 0 & \hat{A} \end{pmatrix}, \quad (2)$$

where  $\hat{A}$  is a matrix in  $\mathbb{C}^{(m-1) \times (n-1)}$ . If  $\hat{A} = 0$ , then (2) provides a singular value decomposition for  $A$ . Otherwise, since  $\min\{m-1, n-1\} = \min\{m, n\} - 1$ , we know from the induction hypothesis that  $\hat{A}$  has a singular value decomposition  $\hat{U}\hat{\Sigma}\hat{V}^H$ . Consequently,

$$A = U(U^H AV)V^H = U \begin{pmatrix} \sigma & 0 \\ 0 & \hat{U}\hat{\Sigma}\hat{V}^H \end{pmatrix} V^H = \left[ U \begin{pmatrix} 1 & 0 \\ 0 & \hat{U} \end{pmatrix} \right] \begin{pmatrix} \sigma & 0 \\ 0 & \hat{\Sigma} \end{pmatrix} \left[ \begin{pmatrix} 1 & 0 \\ 0 & \hat{V} \end{pmatrix}^H V^H \right]. \quad (3)$$

It is easy to verify that the right-hand side of (3) provides a singular value decomposition for  $A$ . This completes the induction.  $\square$

**Remark 5.** We can also prove Theorem 4 by an induction on  $\text{rank}(A)$ . When  $\text{rank}(A) = 0$ , the desired conclusion is trivial. Assume that the conclusion holds when  $\text{rank}(A) \leq k$ . Let us consider the scenario with  $\text{rank}(A) = k + 1$ . By Lemma 3, there exists unit vectors  $u \in \mathbb{C}^m$ ,  $v \in \mathbb{C}^n$ , and a scalar  $\sigma > 0$  fulfilling (1). Define  $\hat{A} = A - \sigma uv^H$ . Then it is easy to check that  $\ker(A) \subset \ker(\hat{A})$  and  $v \in \ker(\hat{A})$ . Since  $v \in \text{range}(A^H) \perp \ker(A)$ , we know that  $\dim \ker(\hat{A}) \geq \dim \ker(A) + 1$ . Thus  $\text{rank}(\hat{A}) \leq \text{rank}(A) - 1$ . If  $\hat{A} = 0$ , we are done. Otherwise, by the induction hypothesis,  $\hat{A}$  has a compact singular value decomposition  $\hat{U}\hat{\Sigma}\hat{V}^H$ . Consequently,

$$A = \sigma uv^H + \hat{A} = \sigma uv^H + \hat{U}\hat{\Sigma}\hat{V}^H = (u \ \hat{U}) \begin{pmatrix} \sigma & 0 \\ 0 & \hat{\Sigma} \end{pmatrix} (v \ \hat{V})^H. \quad (4)$$

Noting that  $\hat{A}v = 0$ ,  $\hat{A}^H u = 0$ , and  $\hat{\Sigma}$  is nonsingular, we can see that  $\hat{V}^H v = 0$  and  $\hat{U}^H u = 0$ . Thus the columns of  $(u \ \hat{U})$  are orthonormal, and so are those of  $(v \ \hat{V})$ . Hence (4) provides a compact singular value decomposition for  $A$ , which can be extended to a singular value decomposition. The induction is complete.

### 2.2.2 The Eckart-Young approach [1]

**Lemma 4.** Suppose that  $A \in \mathbb{C}^{m \times n}$  and  $B \in \mathbb{C}^{n \times k}$ .

1.  $\|AB\|_F \leq \|A\| \|B\|_F$ , and the equality holds if and only if  $A^H AB = \|A\|^2 B$ .
2.  $\|AB\|_F \leq \|A\|_F \|B\|$ , and the equality holds if and only if  $ABB^H = \|B\|^2 A$ .

**Remark 6.** Recall that  $\|\cdot\|$  denotes the 2-norm for matrices.

**Proof.** Let  $C \in \mathbb{C}^{n \times n}$  be the square root of the positive semidefinite matrix  $\|A\|^2 I_n - A^H A$ . Then

$$\|A\|^2 \|B\|_F^2 - \|AB\|_F^2 = \text{tr}(\|A\|^2 B^H B) - \text{tr}(B^H A^H AB) = \text{tr}(B^H C^2 B) \geq 0.$$

Thus  $\|AB\|_F \leq \|A\| \|B\|_F$ , and

$$\|AB\|_F = \|A\| \|B\|_F \iff B^H C^2 B = 0 \iff C^2 B = 0 \iff A^H AB = \|A\|^2 B.$$

The proof concerning  $\|AB\|_F \leq \|A\|_F \|B\|$  is similar. □

**Lemma 5.** Let  $A$ ,  $B$ ,  $U$ , and  $V$  be complex matrices of proper sizes so that both  $U^H AV$  and  $UBV^H$  are well defined. Suppose that  $\|A\|_F = \|B\|_F$  and  $\|U\| = \|V\| = 1$ , then  $A = UBV^H$  if and only if  $B = U^H AV$ .

**Proof.** Assume that  $A = UBV^H$ . Since  $\|U\| = \|V\| = 1$  and  $\|A\|_F = \|B\|_F$ , we have

$$\min\{\|UB\|_F, \|BV^H\|_F\} \geq \|UBV^H\|_F = \|A\|_F = \|B\|_F.$$

Hence Lemma 4 ensures

$$U^H UB = B, \quad BV^H V = B.$$

Therefore,

$$U^H AV = U^H UBV^H V = BV^H V = B.$$

In the same way,  $B = U^H AV$  implies  $A = UBV^H$ . □

**Theorem 5.** Let  $A \in \mathbb{C}^{m \times n}$  be a matrix.

1. If  $V\Lambda V^H$  is a compact eigenvalue decomposition of  $A^H A$  and  $U = AV\Lambda^{-\frac{1}{2}}$ , then  $U\Lambda^{\frac{1}{2}}V^H$  is a compact singular value decomposition of  $A$ .
2. If  $U\Lambda U^H$  is a compact eigenvalue decomposition of  $AA^H$  and  $V = A^H U\Lambda^{-\frac{1}{2}}$ , then  $U\Lambda^{\frac{1}{2}}V^H$  is a compact singular value decomposition of  $A$ .

**Proof.** We only prove 1. By assumption,  $V^H V = I$ ,  $A^H A = V \Lambda V^H$ , and  $U = AV \Lambda^{-\frac{1}{2}}$ . Hence

$$U^H U = (AV \Lambda^{-\frac{1}{2}})^H (AV \Lambda^{-\frac{1}{2}}) = \Lambda^{-\frac{1}{2}} (V^H A^H A V) \Lambda^{-\frac{1}{2}} = \Lambda^{-\frac{1}{2}} \Lambda \Lambda^{-\frac{1}{2}} = I. \quad (5)$$

Thus

$$U^H A V = U^H (U \Lambda^{\frac{1}{2}}) = \Lambda^{\frac{1}{2}}.$$

Meanwhile,  $\|U\| = 1$  by (5),  $\|V\| = 1$  because  $V^H V = I$ , and

$$\|A\|_F^2 = \text{tr}(A^H A) = \text{tr}(\Lambda) = \|\Lambda^{\frac{1}{2}}\|_F^2.$$

Therefore, Lemma 5 ensures

$$A = U \Lambda^{\frac{1}{2}} V^H.$$

Hence  $U \Lambda^{\frac{1}{2}} V^H$  is a compact singular value decomposition of  $A$ .  $\square$

**Remark 7.** The major point of the proof is to show that  $U \Lambda^{\frac{1}{2}} V^H = A$ . Here we use Lemma 5, but there are other ways to prove it.

**Theorem 6.** Let  $A \in \mathbb{C}^{m \times n}$  be a matrix.

1. If  $V \Lambda V^H$  is an eigenvalue decomposition of  $A^H A$  such that the diagonal entries of  $\Lambda$  are descending. Then there exist  $U \in \mathbb{C}^{m \times m}$  and  $\Sigma \in \mathbb{R}^{m \times n}$  such that  $U \Sigma V^H$  is a singular value decomposition of  $A$ .
2. If  $U \Lambda U^H$  is an eigenvalue decomposition of  $A A^H$  such that the diagonal entries of  $\Lambda$  are descending. Then there exist  $V \in \mathbb{C}^{n \times n}$  and  $\Sigma \in \mathbb{R}^{m \times n}$  such that  $U \Sigma V^H$  is a singular value decomposition of  $A$ .

**Proof.** We only prove 1. Suppose that  $\text{rank}(A) = r$ . Let  $\hat{\Lambda} = \text{diag}(\Lambda_{1,1}, \dots, \Lambda_{r,r})$  and  $\hat{V}$  be the first  $r$  columns of  $V$ . Then  $\hat{V} \hat{\Lambda} \hat{V}^H$  is a compact eigenvalue decomposition of  $A^H A$ . With  $\hat{U} = A \hat{V} \hat{\Lambda}^{-\frac{1}{2}}$ , we know that  $\hat{U} \hat{\Lambda}^{\frac{1}{2}} \hat{V}^H$  is a compact singular value decomposition of  $A$ . Let  $\tilde{U} \in \mathbb{C}^{m \times (m-r)}$  be any matrix such that  $(\hat{U} \tilde{U})$  is unitary. Then

$$(\hat{U} \tilde{U}) \begin{pmatrix} \hat{\Lambda}^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix} V^H$$

is a singular value decomposition of  $A$ .  $\square$

### 2.2.3 The Wielandt-Lanczos approach [6]

**Lemma 6.** Given a matrix  $A \in \mathbb{C}^{m \times n}$ , define its Jordan-Wielandt form [7] to be

$$J = \begin{pmatrix} 0 & A \\ A^H & 0 \end{pmatrix}. \quad (6)$$

Then the characteristic polynomial of  $J$  is

$$p(\sigma) = \sigma^{m-n} \det(\sigma^2 I_n - A^H A) = \sigma^{n-m} \det(\sigma^2 I_m - A A^H). \quad (7)$$

If the nonzero eigenvalues of  $A A^H$  (i.e., those of  $A^H A$ ) are  $\lambda_1, \dots, \lambda_r$ , multiplicity included, then the nonzero eigenvalues of  $J$  are  $\sqrt{\lambda_1}, -\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}, -\sqrt{\lambda_r}$ , multiplicity included.

**Proof.** We only prove the first equality in (7). For any  $\sigma \neq 0$ ,

$$\begin{pmatrix} I_m & 0 \\ \sigma^{-1}A^H & I_n \end{pmatrix} \begin{pmatrix} \sigma I_m & -A \\ -A^H & \sigma I_n \end{pmatrix} = \begin{pmatrix} \sigma I_m & -A \\ 0 & \sigma I_n - \sigma^{-1}A^H A \end{pmatrix}.$$

Taking determinants, we have

$$\det(\sigma I - J) = \det(\sigma I_m) \det(\sigma I_n - \sigma^{-1}A^H A) = \sigma^{m-n} \det(\sigma^2 I_n - A^H A). \quad (8)$$

In (8), two rational functions are equal for all  $\sigma \neq 0$ . Hence they are indeed identical.  $\square$

**Theorem 7.** Consider matrices  $A \in \mathbb{C}^{m \times n}$ ,  $\Sigma \in \mathbb{R}^{r \times r}$ ,  $U_i \in \mathbb{C}^{m \times r}$ , and  $V_i \in \mathbb{C}^{n \times r}$  ( $i = 1, 2$ ). Suppose that  $\Sigma$  is a diagonal matrix whose diagonal entries are positive. Then

$$\begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix} \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^H \quad (9)$$

is a compact eigenvalue decomposition of the Jordan-Wielandt matrix  $J$  in (6) if and only if both  $(\sqrt{2}U_1)\Sigma(\sqrt{2}V_1)^H$  and  $(-\sqrt{2}U_2)\Sigma(\sqrt{2}V_2)^H$  are compact singular value decompositions of  $A$ .

**Proof.** 1. Assume that (9) is a compact eigenvalue decomposition of  $J$ . We will prove that  $(\sqrt{2}U_1)\Sigma(\sqrt{2}V_1)^H$  is a compact singular value decompositions of  $A$ , and the other one can be discussed similarly. It suffices to show that

$$U_1^H U_1 = V_1^H V_1 = \frac{I_r}{2}, \quad U_1 \Sigma V_1^H = \frac{A}{2}. \quad (10)$$

Due to the compact eigenvalue decomposition (9) of  $J$ , the columns of  $[U_1; V_1]$  are eigenvectors of  $J$  associated with all its  $r$  positive eigenvalues,<sup>1</sup> and

$$J \begin{pmatrix} U_1 \\ V_1 \end{pmatrix} = \begin{pmatrix} U_1 \\ V_1 \end{pmatrix} \Sigma, \quad (11)$$

This implies  $AV_1 = U_1 \Sigma$  and  $A^H U_1 = \Sigma V_1$ , which can be reformulated as

$$J \begin{pmatrix} U_1 \\ -V_1 \end{pmatrix} = \begin{pmatrix} U_1 \\ -V_1 \end{pmatrix} (-\Sigma), \quad (12)$$

i.e., the columns of  $[U_1; -V_1]$  are eigenvectors of  $J$  associated with the negative eigenvalues. Hence the columns of  $[U_1; V_1]$  and those of  $[U_1; -V_1]$  are orthogonal. Thus

$$U_1^H U_1 - V_1^H V_1 = (U_1^H \ V_1^H) \begin{pmatrix} U_1 \\ -V_1 \end{pmatrix} = 0.$$

With (9) being a compact eigenvalue decomposition, we also have  $U_1^H U_1 + V_1^H V_1 = I_r$ . Hence  $U_1^H U_1 = V_1^H V_1 = I_r/2$ , which is the first equality in (10). To establish the second one, define

$$\bar{U} = \begin{pmatrix} U_1 & U_1 \\ V_1 & -V_1 \end{pmatrix}, \quad \bar{\Sigma} = \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}.$$

---

<sup>1</sup>Recall that the MATLAB-style notation  $[a; b]$  denotes a vertical array with  $a$  and  $b$  being its entries.

Then

$$\bar{U}^H \bar{U} = \begin{pmatrix} U_1^H U_1 + V_1^H V_1 & U_1^H U_1 - V_1^H V_1 \\ U_1^H U_1 - V_1^H V_1 & U_1^H U_1 + V_1^H V_1 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & I_r \end{pmatrix}. \quad (13)$$

Meanwhile, we can reformulate (11)–(12) as  $J\bar{U} = \bar{U}\bar{\Sigma}$ . Therefore,

$$\bar{U}^H J \bar{U} = \bar{\Sigma}.$$

Note that  $\|J\|_F = \|\bar{\Sigma}\|_F$  according to the compact eigenvalue decomposition (9), and  $\|\bar{U}\| = 1$  due to (13). Thus Lemma 5 renders

$$J = \bar{U} \bar{\Sigma} \bar{U}^H, \quad (14)$$

from which we can obtain  $A = 2U_1 \Sigma V_1^H$  by straightforward calculations.

2. Assume that both  $(\sqrt{2}U_1)\Sigma(\sqrt{2}V_1)^H$  and  $(-\sqrt{2}U_2)\Sigma(\sqrt{2}V_2)^H$  are compact singular value decompositions of  $A$ . Then we have (10) and

$$U_2^H U_2 = V_2^H V_2 = \frac{I_r}{2}, \quad -U_2 \Sigma V_2^H = \frac{A}{2}. \quad (15)$$

To prove that (9) is a compact singular value decomposition for  $J$ , it suffices to show

$$\begin{pmatrix} 0 & A \\ A^H & 0 \end{pmatrix} = \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix} \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^H, \quad \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^H \begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & I_r \end{pmatrix},$$

which resolve to

$$\begin{cases} U_1 \Sigma U_1^H - U_2 \Sigma U_2^H = 0, & V_1 \Sigma V_1^H - V_2 \Sigma V_2^H = 0, & U_1 \Sigma V_1^H - U_2 \Sigma V_2^H = A, \\ U_1^H U_1 + V_1^H V_1 = I_r, & U_2^H U_2 + V_2^H V_2 = I_r, & U_1^H U_2 + V_1^H V_2 = 0 \end{cases} \quad (16)$$

Since  $U_1 \Sigma V_1^H = -U_2 \Sigma V_2^H = A/2$ , Theorem 3 implies the existence of a unitary matrix  $W \in \mathbb{C}^{r \times r}$  such that

$$U_2 = -U_1 W, \quad V_2 = V_1 W, \quad \Sigma W = W \Sigma.$$

Hence

$$U_2 \Sigma U_2^H = (-U_1 W) \Sigma (-U_1 W)^H = U_1 W \Sigma W^H U_1^H = U_1 \Sigma W W^H U_1^H = U_1 \Sigma U_1^H,$$

which implies  $U_1 \Sigma U_1^H - U_2 \Sigma U_2^H = 0$ . Similarly,  $V_1 \Sigma V_1^H - V_2 \Sigma V_2^H = 0$ . In addition,

$$U_1^H U_2 + V_1^H V_2 = -U_1^H U_1 W + V_1^H V_1 W = 0,$$

where we use the fact that  $U_1^H U_1 = V_1^H V_1 = I_r/2$  from (10). By (10) and (15), we also have

$$U_1^H U_1 + V_1^H V_1 = U_2^H U_2 + V_2^H V_2 = I_r, \quad U_1 \Sigma V_1^H - U_2 \Sigma V_2^H = A.$$

All the equations in (16)–(17) have been justified. The proof is complete.  $\square$

**Remark 8.** Due to (13) and (14),  $\bar{U} \bar{\Sigma} \bar{U}^H$  is indeed a compact eigenvalue decomposition of  $J$ .



### 3 Decompose a matrix into partial isometries [5]

**Definition 3.** A matrix  $A \in \mathbb{C}^{m \times n}$  is said to be a partial isometry if  $\|Ax\| = \|x\|$  for each  $x \in \text{range}(A^H)$  (i.e.,  $x \in (\ker A)^\perp$ ).

The following proposition collects various characterizations of partial isometries.

**Proposition 1.** For any  $A \in \mathbb{C}^{m \times n}$ , the following statements are equivalent.

1.  $A$  is a partial isometry.
2.  $A^H$  is a partial isometry.
3.  $A^H A$  is an orthogonal projection.
4.  $A A^H$  is an orthogonal projection.
5.  $A^H A A^H = A^H$ .
6.  $A A^H A = A$ .
7. All the nonzero singular values of  $A$  are 1.
8. The linear operator  $T : x \mapsto Ax$  is an isometric isomorphism from  $\text{range}(A^H)$  to  $\text{range}(A)$ .

**Proof.**  $1 \Rightarrow 8$ . Obvious.

$8 \Rightarrow 2$ . Take any  $x \in \text{range}(A)$ . There is a  $y \in \text{range}(A^H)$  such that  $x = Ay$ . By assumption,  $\|y\| = \|x\|$ . Hence

$$\|A^H x\| \geq (y/\|y\|)^H A^H x = \frac{1}{\|x\|} y^H A^H x = \frac{1}{x} x^H x = \|x\|.$$

For any  $z \in \mathbb{C}^n$ , let  $z'$  be its orthogonal projection to  $\text{range}(A^H)$ . Then  $\|Az\| = \|Az'\| = \|z'\| \leq \|z\|$ . Thus

$$\|A^H x\| = \max_{\|z\|=1} z^H A^H x \leq \max_{\|z\|=1} \|Az\| \|x\| \leq \|x\|.$$

$2 \Rightarrow 3$ . Since  $A^H A$  is Hermitian, it suffices to show that it is idempotent. We only need to prove that  $x^H (A^H A)^2 y = x^H A^H A y$  for any  $x$  and  $y \in \mathbb{R}^n$ , or equivalently,  $u^H A A^H v = u^H v$  for any  $u$  and  $v \in \text{range}(A)$ . By assumption, For any  $x, y \in \text{range}(A)$ , we have  $\|A^H u\| = \|u\|$ ,  $\|A^H v\| = \|v\|$ , and  $\|A^H(u+v)\| = \|u+v\|$ . Squaring the last equality, we obtain  $u^H A A^H v = u^H v$ .

$3 \Rightarrow 5$ . Since  $A^H A$  is a projection, we have  $A^H A x = x$  for any  $x \in \text{range}(A^H A) = \text{range}(A^H)$ . Hence  $A^H A A^H = A^H$ .

$5 \Rightarrow 4$ .  $A A^H$  is Hermitian, and  $(A A^H)^2 = A A^H A A^H = A A^H$ .

$4 \Rightarrow 6$ . Similar to  $3 \Rightarrow 5$ .

$6 \Rightarrow 7$ . Since  $A A^H$  is positive semidefinite and  $(A A^H)^2 = A A^H A A^H = A A^H$ , we know that all the nonzero eigenvalues of  $A A^H$  are 1.

$7 \Rightarrow 1$ . Let  $r = \text{rank}(A)$ . If  $r = 0$ , then the conclusion is trivially true. Otherwise,  $A$  has a compact singular value decomposition of the form  $UV^H$ , where  $U \in \mathbb{C}^{m \times r}$  and  $V \in \mathbb{C}^{n \times r}$  satisfy  $U^H U = V^H V = I_r$ . Note that the columns of  $V$  is an orthonormal basis of  $\text{range}(A^H)$ . Therefore, for any  $x \in \text{range}(A^H)$ ,  $\|Ax\| = \|UV^H x\| = \|V^H x\| = \|x\|$ .  $\square$

**Theorem 8.** For any Hermitian matrix  $A \in \mathbb{C}^{n \times n}$ , there exists a unique decomposition

$$A = \sum_{i=1}^k \lambda_i P_i \quad (18)$$

such that

1.  $\{\lambda_i\}$  are all real numbers and  $\lambda_1 > \dots > \lambda_k$ .
2.  $\{P_i\}$  are all orthogonal projections,  $P_i P_j = 0$  for any distinct  $i$  and  $j$ , and  $\sum_{i=1}^k P_i = I$ .

**Proof.** The existence is easy to establish by any eigenvalue decomposition of  $A$ . We only prove the uniqueness.

Consider any decomposition in the form of (18). For each  $i$ , let  $V_i$  be a matrix whose columns form an orthonormal basis of  $\text{range}(P_i)$ . Then  $V_i^H V_i$  is an identity matrix, and  $P_i = V_i V_i^H$ . For any distinct  $i$  and  $j$ ,

$$V_i^H V_j = (V_i^H V_i) V_i^H V_j (V_j^H V_j) = V_i^H P_i P_j V_j = 0.$$

Define

$$V = (V_1 \dots V_k).$$

Then the columns of  $V$  are orthonormal. In addition,

$$V V^H = \sum_{i=1}^k V_i V_i^H = \sum_{i=1}^k P_i = I.$$

Thus  $V$  is a unitary matrix. In addition,

$$A = \sum_{i=1}^k \lambda_i P_i = \sum_{i=1}^k \lambda_i V_i V_i^H = \sum_{i=1}^k V_i \Lambda_i V_i^H = V \Lambda V^H,$$

where  $\Lambda_i = \lambda_i V_i^H V_i$ , and  $\Lambda$  is the block diagonal matrix whose diagonal blocks are  $\Lambda_i$ . Note that  $\Lambda$  is indeed a diagonal matrix since each  $\Lambda_i$  is diagonal. Thus  $V \Lambda V^H$  is an eigenvalue decomposition of  $A$ , with  $\lambda_1, \dots, \lambda_k$  being all the distinct eigenvalues, ranked in the descending order. Moreover, for each  $i$ , the columns of  $V_i$  form an orthonormal basis of the eigenspace associated with  $\lambda_i$ , and hence  $P_i$  is the orthogonal projection onto this eigenspace. In this way,  $\{\lambda_i\}$  and  $\{P_i\}$  are uniquely determined by  $A$ .  $\square$

**Theorem 9.** For any nonzero matrix  $A \in \mathbb{C}^{m \times n}$ , there exists a unique decomposition

$$A = \sum_{i=1}^k \sigma_i A_i \quad (19)$$

such that

1.  $\sigma_1 > \dots > \sigma_k > 0$ ;
2.  $\{A_i\}$  are all partial isometries, with  $A_i A_j^H$  and  $A_i^H A_j$  both being zero for any distinct  $i$  and  $j$ .

**Proof.** The existence is easy to establish by any singular value decomposition of  $A$ . We only prove the uniqueness.

Consider any decomposition in the form of (19). Since  $A_i^H A_j = 0$  for any distinct  $i$  and  $j$ , we have

$$A^H A = \left( \sum_{i=1}^k \sigma_i A_i \right)^H \left( \sum_{i=1}^k \sigma_i A_i \right) = \sum_{i=1}^k \sigma_i^2 A_i^H A_i. \quad (20)$$

For each  $i$ ,  $A_i^H A_i$  is an orthogonal projection as  $A_i$  is a partial isometry (see 3 of Proposition 1). Hence (20) is a decomposition specified in Theorem 8. Due to the uniqueness part of Theorem 8,  $\sigma_1, \dots, \sigma_k$  and  $A_1^H A_1, \dots, A_k^H A_k$  are uniquely determined by  $A$ .

Now consider any two decompositions in the form of (19). According to what is proved above, we can formulate the decompositions as

$$A = \sum_{i=1}^k \sigma_i A_i \quad \text{and} \quad A = \sum_{i=1}^k \sigma_i \tilde{A}_i, \quad (21)$$

and we have  $A_i^H A_i = \tilde{A}_i^H \tilde{A}_i$  for each  $i$ . For any distinct  $i$  and  $j$ ,

$$A_i \tilde{A}_j^H = (A_i A_i^H A_i) \tilde{A}_j^H = (A_i \tilde{A}_i^H \tilde{A}_i) \tilde{A}_j^H = 0,$$

where the first equality is because  $A_i$  is a partial isometry (see 6 of Proposition 1), and the second is because  $\tilde{A}_i \tilde{A}_j^H = 0$ . Similarly,  $\tilde{A}_i A_j^H = 0$ . Hence

$$(A_i - \tilde{A}_i)(A_j - \tilde{A}_j)^H = 0.$$

Thus

$$\left[ \sum_{i=1}^k \sigma_i (A_i - \tilde{A}_i) \right] \left[ \sum_{i=1}^k \sigma_i (A_i - \tilde{A}_i) \right]^H = \sum_{i=1}^k \sigma_i^2 (A_i - \tilde{A}_i)(A_i - \tilde{A}_i)^H. \quad (22)$$

According to (21), the left-hand side of (22) is zero. Hence  $A_i = \tilde{A}_i$  for each  $i$ . Therefore, the two decompositions in (21) are identical. The proof is complete.  $\square$

Theorem 9 is indeed the matrix version of the following theorem.

**Theorem 10** ([5, Theorem 1]). *Let  $X$  and  $Y$  be finite dimensional Hilbert spaces and  $T : X \rightarrow Y$  be a linear operator. Then there exist unique orthogonal decompositions*

$$\text{range}(T^*) = X_1 \oplus \dots \oplus X_k, \quad \text{range}(T) = Y_1 \oplus \dots \oplus Y_k,$$

*scalars  $\sigma_1 > \dots > \sigma_k$ , and isometries  $T_i : X_i \rightarrow Y_i$  ( $i = 1, \dots, k$ ) such that*

$$T|_{X_i} = \sigma_i T_i \quad \text{for each } i \in \{1, \dots, k\}.$$

## 4 Understanding SVD as a change of basis

**Theorem 11.** *Suppose that  $X$  is finite dimensional vector space on  $\mathbb{F}$  with  $\{x_1, \dots, x_n\}$  being its basis and  $C_X : X \rightarrow \mathbb{F}^n$  being the map from any point in  $X$  to its coordinate under this basis;  $Y$ ,  $\{y_1, \dots, y_m\}$ , and  $C_Y : Y \rightarrow \mathbb{F}^m$  are similar. Consider a linear operator  $T : X \rightarrow Y$ .*

1. There is a unique matrix  $A \in \mathbb{F}^{m \times n}$  that represents  $T$  under the aforementioned bases of  $X$  and  $Y$  in the sense that

$$AC_X(x) = C_Y(T(x)) \quad \text{for all } x \in X,$$

meaning that applying  $T$  to  $x$  is equivalent to multiplying its coordinate by  $A$ . Indeed, the  $i$ -th column of  $A$  is  $A_i = C_Y(T(x_i))$ , namely the coordinate of  $T(x_i)$ .

2.  $A$  has full row rank if and only if  $T$  is injective;  $A$  has full column rank if and only if  $T$  is surjective; when  $m = n$ ,  $T$  is invertible if and only if  $A$  is invertible, and  $A^{-1}$  represents  $T^{-1} : Y \rightarrow X$  under the aforementioned bases for  $X$  and  $Y$ .
3.  $A^*$  represents  $T^* : Y^* \rightarrow X^*$  under the bases for  $X^*$  and  $Y^*$  that are dual to  $\{x_1, \dots, x_n\}$  and  $\{y_1, \dots, y_m\}$  respectively.
4. Let  $\{x'_1, \dots, x'_n\}$  be a basis for  $X$ ,  $\{y'_1, \dots, y'_m\}$  be a basis for  $Y$ , and  $B \in \mathbb{F}^{m \times n}$  be the representation of  $T$  under such bases. Then  $A = Q^{-1}BP$ , where  $P \in \mathbb{F}^{n \times n}$  and  $Q \in \mathbb{F}^{m \times m}$  are the changing of basis matrices such that

$$(x_1 \cdots x_n) = (x'_1 \cdots x'_n)P, \quad (y_1 \cdots y_m) = (y'_1 \cdots y'_m)Q.$$

With the view point presented in Theorem 11, we can understand SVD as follows.

Let  $A \in \mathbb{C}^{m \times n}$  be a nonzero matrix, and  $U\Sigma V^H$  be its SVD. Consider the linear operator  $T : x \mapsto Ax$  from  $\mathbb{C}^n$  to  $\mathbb{C}^m$ . Then  $A$  represents  $T$  under the canonical bases. If we take the columns of  $V$  as the basis for  $\mathbb{C}^n$  and those of  $U$  as the basis for  $\mathbb{C}^m$ , then SVD provides a simple representation for  $T$ , which is  $\Sigma$ .

Recall the decompositions

$$\mathbb{C}^n = \ker(A) \oplus \text{range}(A^H), \quad \mathbb{C}^m = \ker(A^H) \oplus \text{range}(A).$$

When  $T$  acts on  $\mathbb{C}^n$ , it drops out the information in  $\ker(A)$ , and provides no information in  $\ker(A^H)$ . Consequently,  $T$  is not an isomorphism if either  $\ker(A)$  or  $\ker(A^H)$  is nonzero. The restriction  $\hat{T} : \text{range}(A^H) \rightarrow \text{range}(A)$  with  $\hat{T}(x) = Ax$  is however always an isomorphism. Suppose that  $\hat{U}\hat{\Sigma}\hat{V}^H$  is a compact SVD of  $A$ . Note that the columns of  $\hat{U}$  form an orthonormal basis for  $\text{range}(A)$ , and those of  $\hat{V}$  form an orthonormal basis for  $\text{range}(A^H)$ . Under these bases,  $\hat{T}$  is represented by  $\hat{\Sigma}$ . The representation for  $\hat{T}^{-1} : \text{range}(A) \rightarrow \text{range}(A^H)$  is  $\hat{\Sigma}^{-1}$ .

The operator  $T^+ : \mathbb{C}^m \rightarrow \mathbb{C}^n$  defined by

$$T^+|_{\text{range}(A)} = \hat{T}^{-1}, \quad T^+|_{\ker(A^H)} = 0$$

is called the Moore-Penrose pseudoinverse of  $T$ . The representation of  $T^+$  under the canonical bases is called the Moore-Penrose pseudoinverse of  $A$ , which turns out to be  $A^+ = \hat{U}\hat{\Sigma}^{-1}\hat{V}^H$ .

We can regard any nonzero linear operator as a bijection by restricting its domain and image space. SVD tells us that any nonzero linear operator between finite dimensional Hilbert spaces can be represented by a positive diagonal matrix under properly chosen orthonormal bases for the restricted domain and image space. This can also lead us to Theorem 10.

## 5 Examples of applications

**Proposition 2** (Polar decomposition). *Let  $A \in \mathbb{C}^{m \times n}$  be a matrix.*

1. *If  $m \geq n$ , there exists a positive semidefinite matrix  $P \in \mathbb{C}^{n \times n}$  and a matrix  $U \in \mathbb{C}^{m \times n}$  such that  $A = UP$  and  $U^H U = I_n$ ; there also exists a positive semidefinite matrix  $Q \in \mathbb{C}^{m \times m}$  and a matrix  $V \in \mathbb{C}^{m \times n}$  such that  $A = QV$  and  $V^H V = I_n$ . In this case,  $P = (A^H A)^{\frac{1}{2}}$ .*
2. *If  $n \geq m$ , there exists a positive semidefinite matrix  $P \in \mathbb{C}^{n \times n}$  and a matrix  $U \in \mathbb{C}^{m \times n}$  such that  $A = UP$  and  $U U^H = I_m$ ; there also exists a positive semidefinite matrix  $Q \in \mathbb{C}^{m \times m}$  and a matrix  $V \in \mathbb{C}^{m \times n}$  such that  $A = QV$  and  $V V^H = I_m$ . In this case,  $Q = (A A^H)^{\frac{1}{2}}$ .*

*If  $A$  is real, we can require  $P$ ,  $U$ ,  $Q$ , and  $V$  to be real.*

**Proof.** We only prove 1. Let  $W \Sigma Z^H$  be an SVD of  $A$ .

Note that the last  $m - n$  rows of  $\Sigma$  are zero. Let  $\hat{\Sigma}$  be the first  $n$  rows of  $\Sigma$ , and  $\hat{W}$  be the first  $n$  columns of  $W$ . Then  $A = \hat{W} \hat{\Sigma} Z^H$ . Define  $U = \hat{W} Z^H$  and  $P = Z \hat{\Sigma} Z^H$ . Then  $P$  is positive semidefinite, and

$$A = \hat{W} \hat{\Sigma} Z^H = UP, \quad U^H U = Z \hat{W}^H \hat{W} Z^H = Z Z^H = I_n.$$

Consequently,  $A^H A = P^H U^H U P = P^2$ , and hence  $P = (A^H A)^{\frac{1}{2}}$ .

Let  $\bar{\Sigma} = (\Sigma \ 0_{m \times (m-n)})$  and  $\bar{Z} = (Z \ 0_{n \times (m-n)})$ . Then  $A = W \bar{\Sigma} \bar{Z}^H$ . Define  $Q = W \bar{\Sigma} W^H$  and  $V = W \bar{Z}^H$ . Then  $Q$  is positive semidefinite, and

$$A = W \bar{\Sigma} \bar{Z}^H = QV, \quad V^H V = \bar{Z} W^H W \bar{Z}^H = \bar{Z} \bar{Z}^H = Z Z^H = I_n.$$

If  $A$  is real, then  $W$ ,  $\Sigma$ , and  $Z$  can all be real, ensuring  $P$ ,  $U$ ,  $Q$ , and  $V$  to be real.  $\square$

**Proposition 3** ([2]). *Let  $H$  be the Hermitian part of a matrix  $A \in \mathbb{C}^{n \times n}$ . Enumerating the eigenvalues of  $H$  as  $\lambda_1(H) \geq \dots \geq \lambda_n(H)$ , and the singular values of  $A$  as  $\sigma_1(A) \geq \dots \geq \sigma_n(A)$ , we have  $\sigma_i(A) \geq \lambda_i(H)$  for each  $i = 1, \dots, n$ .*

**Proof.** By Theorem 2, there exists a positive semidefinite matrix  $P \in \mathbb{C}^{n \times n}$  and a unitary matrix  $U \in \mathbb{C}^{n \times n}$  such that  $A = UP$ . For any unit vector  $x \in \mathbb{C}^n$ ,

$$x^H H x = \frac{1}{2} x^H (A^H + A) x = \Re(x^H A x) \leq |x^H A x| = |x^H U P x| \leq \|P x\| = (x^H P^2 x)^{\frac{1}{2}} = (x^H A^H A x)^{\frac{1}{2}}.$$

Therefore, by the Courant-Fischer-Weyl min-max principle, we know that

$$\lambda_i(H) \leq \lambda_i(A^H A)^{\frac{1}{2}} = \sigma_i(A).$$

$\square$

**Proposition 4.** *For matrices  $A_1$  and  $A_2 \in \mathbb{C}^{m \times n}$ ,  $A_1^H A_1 = A_2^H A_2$  if and only if there exists a unitary matrix  $U \in \mathbb{C}^{m \times m}$  such that  $A_2 = U A_1$ .*

**Proof.** The “if” part is trivial. We focus on the “only if” part. Let  $V \Lambda V^H$  be an eigenvalue decomposition of  $A_1^H A_1 = A_2^H A_2$  such that the diagonal entries of  $\Lambda$  is descending. By Theorem 6, there exists  $W_1, W_2 \in \mathbb{C}^{m \times m}$  and  $\Sigma \in \mathbb{R}^{m \times n}$  such that  $A_1 = W_1 \Sigma V^H$  and  $A_2 = W_2 \Sigma V^H$ . Set  $U = W_2 W_1^H$ .  $\square$

## References

- [1] C. Eckart and G. Young. A principal axis transformation for non-Hermitian matrices. *Bull. Amer. Math. Soc.*, 45:118–121, 1939.
- [2] Ky Fan and A. J. Hoffman. Some metric inequalities in the space of matrices. *Proc. Amer. Math. Soc.*, 6:111–116, 1955.
- [3] L. Han and M. Neumann. Inner product spaces, orthogonal projection, least squares, and singular value decomposition. In L. Hogben, editor, *Handbook of Linear Algebra*. CRC Press, Boca Raton, FL, 2013.
- [4] C. Jordan. Mémoire sur les formes bilinéaires. *J. Math. Pures Appl.*, 19:35–54, 1874.
- [5] A. Korányi. Around the finite-dimensional spectral theorem. *Amer. Math. Monthly*, 108:120–125, 2001.
- [6] C. Lanczos. Linear systems in self-adjoint form. *Amer. Math. Monthly*, 65:665–679, 1958.
- [7] R. Mathias. Singular values and singular value inequalities. In L. Hogben, editor, *Handbook of Linear Algebra*. CRC Press, Boca Raton, FL, 2013.
- [8] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.