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 the 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*

$$A = U\Lambda U^{\mathsf{H}}.$$

If A is real, then we can require that U is real. Indeed, $\Lambda_{1,1}, \ldots, \Lambda_{n,n}$ 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. $U\Lambda U^{\mathsf{H}}$ is called an eigenvalue decomposition of A, if $A = U\Lambda U^{\mathsf{H}}$, $U \in \mathbb{C}^{n\times n}$ is a unitary matrix, and $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix.

2. $U\Lambda U^{\mathsf{H}}$ is called a compact eigenvalue decomposition of A, if $A = U\Lambda U^{\mathsf{H}}$, $U \in \mathbb{C}^{n\times r}$ is a matrix with $U^{\mathsf{H}}U = I_r$, 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 $\operatorname{rank}(A) = r$.

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

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

Remark 2. *If $U\Sigma V^{\mathsf{H}}$ is a (compact) singular value decomposition of A, then $AV = U\Sigma$ and $A^{\mathsf{H}}U = V\Sigma$. 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 $A^{\mathsf{H}}u_i = \sigma_i v_i$; u_i and v_i are called a pair of left and right singular vectors of A associated with the singular value σ_i.*

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

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

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 $U\Sigma V^{\mathsf{H}}$ is a singular value decomposition of $A \neq 0$, we can obtain a compact singular value decomposition by dropping the zero diagonal entries of Σ and the corresponding columns of U and V.

2.1 Uniqueness of (compact) SVD

Theorem 2. *Let $U\Sigma V^{\mathsf{H}}$ be a compact singular value decomposition of a matrix A.*

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

Lemma 1. *Consider matrices $A \in \mathbb{C}^{n\times n}$ and $\Lambda \in \mathbb{C}^{n\times n}$ with Λ 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 Λ 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 $i, j \in \{1, \ldots, n\}$. □

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

1. *$U_1U_1^{\mathsf{H}} = U_2U_2^{\mathsf{H}}$, both being the orthogonal projection onto $\operatorname{range}(U_1) = \operatorname{range}(U_2)$.*
2. *$W = U_1^{\mathsf{H}}U_2$ is a unitary matrix and $U_1W = U_2$.*

Proof. $U_1^{\mathsf{H}}U_1 = I_r$ ensures that $U_1U_1^{\mathsf{H}}$ is the orthogonal projection onto $\operatorname{range}(U_1) = \operatorname{range}(U_2)$ (see, e.g., [4]). In addition, $U_1W = U_1U_1^{\mathsf{H}}U_2 = U_2$, and $W^{\mathsf{H}}W = U_2^{\mathsf{H}}U_1W = U_2^{\mathsf{H}}U_2 = I_r$. □

Theorem 3. *Let $U_i \in \mathbb{C}^{m\times r}$ and $V_i \in \mathbb{C}^{n\times r}$ satisfy $U_i^{\mathsf{H}} U_i = V_i^{\mathsf{H}} V_i = I_r$ $(i = 1, 2)$, and $\Sigma \in \mathbb{C}^{r\times r}$ be a diagonal matrix whose diagonal entries are positive. Then $U_1 \Sigma V_1^{\mathsf{H}} = U_2 \Sigma V_2^{\mathsf{H}}$ 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 $U_1 \Sigma V_1^{\mathsf{H}} = U_2 \Sigma V_2^{\mathsf{H}}$, We will show that $W = U_1^{\mathsf{H}} U_2 \in \mathbb{C}^{r\times r}$ fulfills all the desired requirements. Observe that both ΣV_1^{H} and ΣV_2^{H} have full column rank. Hence

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

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 $V_1^{\mathsf{H}} V_1 = V_2^{\mathsf{H}} V_2 = I_r$, we have

$$U_1 \Sigma^2 U_1^{\mathsf{H}} \;=\; (U_1 \Sigma V_1^{\mathsf{H}})(U_1 \Sigma V_1^{\mathsf{H}})^{\mathsf{H}} \;=\; (U_2 \Sigma V_2^{\mathsf{H}})(U_2 \Sigma V_2^{\mathsf{H}})^{\mathsf{H}} \;=\; U_2 \Sigma^2 U_2^{\mathsf{H}}.$$

Hence

$$\Sigma^2 W \;=\; \Sigma^2 U_1^{\mathsf{H}} U_2 \;=\; U_1^{\mathsf{H}} (U_1 \Sigma^2 U_1^{\mathsf{H}}) U_2 \;=\; U_1^{\mathsf{H}} (U_2 \Sigma^2 U_2^{\mathsf{H}}) U_2 \;=\; U_1^{\mathsf{H}} U_2 \Sigma^2 \;=\; W \Sigma^2$$

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

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

The proof is complete. □

2.2 Existence of SVD

2.2.1 Jordan's deflation approach [5]

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*

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

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

Proof. Since $A \neq 0$, it is obvious that $\sigma > 0$. Hence $Av \neq 0$. By assumption,

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

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

$$\sigma \;=\; \Re(u^{\mathsf{H}} A v) \;=\; \Re(\lambda \|x\|^2) \;=\; \lambda.$$

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

$$\Re(u^{\mathsf{H}} A v) \;\geq\; \Re(u^{\mathsf{H}} A\, (A^{\mathsf{H}} u / \|A^{\mathsf{H}} u\|)) \;=\; \|A^{\mathsf{H}} x\| = \|A^{\mathsf{H}} u\| \|v\|.$$

□

Remark 4. *Indeed, the σ in Lemma 3 is the largest singular value of A, because*

$$\max_{\|x\|=\|y\|=1} \Re(x^{\mathsf{H}}Ay) \;=\; \max_{\|y\|=1}\max_{\|x\|=1} \Re(x^{\mathsf{H}}Ay) \;=\; \max_{\|y\|=1} \|Ay\| \;=\; \|A\| \;=\; \sigma_{\max}(A).$$

Similarly, we can see that

$$\max_{\|x\|=\|y\|=1} |x^{\mathsf{H}}Ay| \;=\; \sigma_{\max}(A).$$

See [1] for more about variational representations for singular values of matrices.

Remark 5. *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 $A^{\mathsf{H}}u = \sigma v$. The existence can be established in other ways.*

1. *Let $\sigma = (\lambda_{\max}(AA^{\mathsf{H}}))^{\frac{1}{2}} > 0$, $u \in \mathbb{C}^{m\times m}$ be an eigenvector of AA^{H} associated with $\lambda_{\max}(AA^{\mathsf{H}})$, and $v = A^{\mathsf{H}}u/\sigma$. Then $Av = AA^{\mathsf{H}}u/\sigma = \sigma^2 u/\sigma = \sigma u$, and $A^{\mathsf{H}}u = \sigma v$. This is the approach used in the proofs of [9, Theorem 4.1] and [6, 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 σ, $x \in \mathbb{C}^m$ 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 $A^{\mathsf{H}}x = \sigma y$. Meanwhile, $A(-y) = -\sigma x$, and $A^{\mathsf{H}}x = -\sigma(-y)$, making $[x; -y]$ an eigenvector of J associated with $-\sigma \neq \sigma$. Since eigenvectors for different eigenvalues are orthogonal, we have $x^{\mathsf{H}}x - y^{\mathsf{H}}y = 0$. Thus $\|x\| = \|y\|$, which are nonzero since $w \neq 0$. Finally, let $u = x/\|x\|$ and $y = y/\|y\|$.*

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

$$Av \;=\; \sigma u, \quad A^{\mathsf{H}}u \;=\; \sigma v. \tag{1}$$

Let $U \in \mathbb{C}^{m\times m}$ be a unitary matrix whose fist 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

$$U^{\mathsf{H}}AV \;=\; \begin{pmatrix} \sigma & 0 \\ 0 & \hat{A} \end{pmatrix}, \tag{2}$$

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 $\hat{U}\hat{\Sigma}\hat{V}^{\mathsf{H}}$. Consequently,

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

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 6. *We can also prove Theorem 4 by an induction on* $\operatorname{rank}(A)$*. When* $\operatorname{rank}(A) = 0$*, the desired conclusion is trivial. Assume that the conclusion holds when* $\operatorname{rank}(A) \le k$*. Let us consider the scenario with* $\operatorname{rank}(A) = k+1$*. 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* $\hat{A} = A - \sigma u v^{\mathsf{H}}$*. Then it is easy to check that* $\ker(A) \subset \ker(\hat{A})$ *and* $v \in \ker(\hat{A})$*. Since* $v \in \operatorname{range}(A^{\mathsf{H}}) \perp \ker(A)$*, we know that* $\dim\ker(\hat{A}) \ge \dim\ker(A) + 1$*. Thus* $\operatorname{rank}(\hat{A}) \le \operatorname{rank}(A) - 1$*. If* $\hat{A} = 0$*, then we are done. Otherwise, by the induction hypothesis,* $\hat{A}$ *has a compact singular value decomposition* $\hat{U}\hat{\Sigma}\hat{V}^{\mathsf{H}}$*. Consequently,*

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

Noting that $\hat{A}v = 0$*,* $\hat{A}^{\mathsf{H}}u = 0$*, and* $\hat{\Sigma}$ *is nonsingular, we can see that* $\hat{V}^{\mathsf{H}}v = 0$ *and* $\hat{U}^{\mathsf{H}}u = 0$*. 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 [2]

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

1. $\|AB\|_{\mathrm{F}} \le \|A\|\|B\|_{\mathrm{F}}$*, and the equality holds if and only if* $A^{\mathsf{H}}AB = \|A\|^2 B$*.*
2. $\|AB\|_{\mathrm{F}} \le \|A\|_{\mathrm{F}}\|B\|$*, and the equality holds if and only if* $ABB^{\mathsf{H}} = \|B\|^2 A$*.*

Remark 7. *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 $\|A\|^2 I_n - A^{\mathsf{H}}A$. Then

$$\|A\|^2\|B\|_{\mathrm{F}}^2 - \|AB\|_{\mathrm{F}}^2 \;=\; \operatorname{tr}(\|A\|^2 B^{\mathsf{H}}B) - \operatorname{tr}(B^{\mathsf{H}}A^{\mathsf{H}}AB) \;=\; \operatorname{tr}(B^{\mathsf{H}}C^2B) \;\ge\; 0.$$

Thus $\|AB\|_{\mathrm{F}} \le \|A\|\|B\|_{\mathrm{F}}$, and

$$\|AB\|_{\mathrm{F}} = \|A\|\|B\|_{\mathrm{F}} \iff B^{\mathsf{H}}C^2B = 0 \iff C^2B = 0 \iff A^{\mathsf{H}}AB = \|A\|^2 B.$$

The proof concerning $\|AB\|_{\mathrm{F}} \le \|A\|_{\mathrm{F}}\|B\|$ is similar. □

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

Proof. Assume that $A = UBV^{\mathsf{H}}$. Since $\|U\| = \|V\| = 1$ and $\|A\|_{\mathrm{F}} = \|B\|_{\mathrm{F}}$, we have

$$\min\{\|UB\|_{\mathrm{F}},\, \|BV^{\mathsf{H}}\|_{\mathrm{F}}\} \;\ge\; \|UBV^{\mathsf{H}}\|_{\mathrm{F}} \;=\; \|A\|_{\mathrm{F}} = \|B\|_{\mathrm{F}}.$$

Hence Lemma 4 ensures

$$U^{\mathsf{H}}UB \;=\; B, \quad BV^{\mathsf{H}}V \;=\; B.$$

Therefore,

$$U^{\mathsf{H}}AV \;=\; U^{\mathsf{H}}UBV^{\mathsf{H}}V \;=\; BV^{\mathsf{H}}V \;=\; B.$$

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

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

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

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

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

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

Thus

$$U^{\mathsf{H}}AV \;=\; U^{\mathsf{H}}(U\Lambda^{\frac{1}{2}}) \;=\; \Lambda^{\frac{1}{2}}.$$

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

$$\|A\|_{\mathrm{F}}^2 \;=\; \mathrm{tr}(A^{\mathsf{H}}A) \;=\; \mathrm{tr}(\Lambda) \;=\; \|\Lambda^{\frac{1}{2}}\|_{\mathrm{F}}^2.$$

Therefore, Lemma 5 ensures

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

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

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

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

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

Proof. We only prove 1. Suppose that $\mathrm{rank}(A) = r$. Let $\hat{\Lambda} = \mathrm{diag}(\Lambda_{1,1}, \ldots, \Lambda_{r,r})$ and $\hat{V}$ be the first r columns of V. Then $\hat{V}\hat{\Lambda}\hat{V}^{\mathsf{H}}$ is a compact eigenvalue decomposition of $A^{\mathsf{H}}A$. With $\hat{U} = A\hat{V}\hat{\Lambda}^{-\frac{1}{2}}$, we know that $\hat{U}\hat{\Lambda}^{\frac{1}{2}}\hat{V}^{\mathsf{H}}$ 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

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

is a singular value decomposition of A. □

2.2.3 The Wielandt-Lanczos approach [7]

Lemma 6. *Given a matrix $A \in \mathbb{C}^{m\times n}$, define its Jordan-Wielandt form [8]*

$$J = \begin{pmatrix} 0 & A \\ A^{\mathsf{H}} & 0 \end{pmatrix}. \tag{6}$$

Then the characteristic polynomial of J is

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

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

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

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

Taking the determinant, we have

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

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

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 Σ is a diagonal matrix whose diagonal entries are positive. Then*

$$\begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}\begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}\begin{pmatrix} U_1 & U_2 \\ V_1 & V_2 \end{pmatrix}^{\mathsf{H}} \tag{9}$$

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

Proof. 1. Assume that (9) is a compact eigenvalue decomposition of J. To prove that both $(\sqrt{2}U_1)\Sigma(\sqrt{2}V_1)^{\mathsf{H}}$ and $(-\sqrt{2}U_2)\Sigma(\sqrt{2}V_2)^{\mathsf{H}}$ are compact singular value decompositions of A, it suffices to show

$$U_1^{\mathsf{H}}U_1 \;=\; V_1^{\mathsf{H}}V_1 \;=\; \frac{I_r}{2}, \quad U_1\Sigma V_1^{\mathsf{H}} \;=\; \frac{A}{2}, \quad U_2^{\mathsf{H}}U_2 \;=\; V_2^{\mathsf{H}}V_2 \;=\; \frac{I_r}{2}, \quad -U_2\Sigma V_2^{\mathsf{H}} \;=\; \frac{A}{2}. \tag{10}$$

Since the decomposition (9) of J renders immediately

$$U_1^{\mathsf{H}}U_1 + V_1^{\mathsf{H}}V_1 \;=\; I_r, \quad U_2^{\mathsf{H}}U_2 + V_2^{\mathsf{H}}V_2 \;=\; I_r, \quad U_1\Sigma V_1^{\mathsf{H}} - U_2\Sigma V_2^{\mathsf{H}} \;=\; A,$$

we only need to prove

$$U_1^{\mathsf{H}}U_1 - V_1^{\mathsf{H}}V_1 \;=\; 0, \quad U_2^{\mathsf{H}}U_2 - V_2^{\mathsf{H}}V_2 \;=\; 0, \quad U_1\Sigma V_1^{\mathsf{H}} + U_2\Sigma V_2^{\mathsf{H}} \;=\; 0. \tag{11}$$

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,[1] and

$$J\begin{pmatrix} U_1 \\ V_1 \end{pmatrix} = \begin{pmatrix} U_1 \\ V_1 \end{pmatrix}\Sigma,$$

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

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

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

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, justifying the first equation in (11). The second one can be proved in the same way. To verify the third, note that

$$U_1^{\mathsf{H}} U_1 + (-V_1)^{\mathsf{H}}(-V_1) \ = \ U_1^{\mathsf{H}} U_1 + V_1^{\mathsf{H}} V_1 \ = \ I_r,$$

which ensures that the columns of $[U_1; -V_1]$ form an orthonormal basis for the space spanned by the eigenvectors of J associated with all its r negative eigenvalues. The columns of $[U_2; V_2]$ also form such a basis according to the eigenvalue decomposition (9). Thus Lemma 2 ensures

$$\begin{pmatrix} U_2 \\ V_2 \end{pmatrix} (U_2^{\mathsf{H}} \ \ V_2^{\mathsf{H}}) = \begin{pmatrix} U_1 \\ -V_1 \end{pmatrix} (U_1^{\mathsf{H}} \ \ -V_1^{\mathsf{H}}).$$

Hence $V_1 V_1^{\mathsf{H}} = V_2 V_2^{\mathsf{H}}$. Recalling that $AV_1 = U_1\Sigma$ and $AV_2 = -U_2\Sigma$ according to (9), we have

$$U_1 \Sigma V_1^{\mathsf{H}} + U_2 \Sigma V_2^{\mathsf{H}} \ = \ AV_1V_1^{\mathsf{H}} - AV_2V_2^{\mathsf{H}} \ = \ 0,$$

which is the third equation in (11).

2. Assume that both $(\sqrt{2}U_1)\Sigma(\sqrt{2}V_1)^{\mathsf{H}}$ and $(-\sqrt{2}U_2)\Sigma(\sqrt{2}V_2)^{\mathsf{H}}$ are compact singular value decompositions of A, and hence (10) holds. To prove (9) is a compact singular value decomposition for J, it suffices to show

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

which resolve to

$$\begin{cases} U_1\Sigma U_1^{\mathsf{H}} - U_2\Sigma U_2^{\mathsf{H}} \ = \ 0, \quad V_1\Sigma V_1^{\mathsf{H}} - V_2\Sigma V_2^{\mathsf{H}} \ = \ 0, \quad U_1\Sigma V_1^{\mathsf{H}} - U_2\Sigma V_2^{\mathsf{H}} \ = \ A, & (12) \\ U_1^{\mathsf{H}}U_1 + V_1^{\mathsf{H}}V_1 \ = \ I_r, \quad U_2^{\mathsf{H}}U_2 + V_2^{\mathsf{H}}V_2 \ = \ I_r, \quad U_1^{\mathsf{H}}U_2 + V_1^{\mathsf{H}}V_2 \ = \ 0 & (13) \end{cases}$$

By Theorem 3, there exists a unitary matrix $W \in \mathbb{C}^{r\times r}$ such that

$$U_2 \ = \ -U_1W, \quad V_2 \ = \ V_1W, \quad \Sigma W \ = \ W\Sigma.$$

Hence

$$U_2\Sigma U_2^{\mathsf{H}} \ = \ (-U_1W)\Sigma(-U_1W)^{\mathsf{H}} \ = \ U_1W\Sigma W^{\mathsf{H}}U_1^{\mathsf{H}} \ = \ U_1\Sigma WW^{\mathsf{H}}U_1^{\mathsf{H}} \ = \ U_1\Sigma U_1^{\mathsf{H}},$$

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

$$U_1^{\mathsf{H}}U_2 + V_1^{\mathsf{H}}V_2 \ = \ -U_1^{\mathsf{H}}U_1W + V_1^{\mathsf{H}}V_1W \ = \ 0,$$

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

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

All the equations in (12)–(13) have been justified. The proof is complete. □

3 Decompose a linear operator into partial isometries [6]

Definition 3. A matrix $A \in \mathbb{C}^{m\times n}$ is said to be a partial isometry if $\|Ax\| = \|x\|$ for each $x \in \operatorname{range}(A^{\mathsf{H}})$ (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. *$A^{\mathsf{H}}A$ is an orthogonal projection.*
4. *AA^{H} is an orthogonal projection.*
5. *$A^{\mathsf{H}}AA^{\mathsf{H}} = A^{\mathsf{H}}$.*
6. *$AA^{\mathsf{H}}A = A$.*
7. *All the nonzero singular values of A are 1.*
8. *The linear operator $T : x \mapsto Ax$ is an isometric isomorphism from $\operatorname{range}(A^{\mathsf{H}})$ to $\operatorname{range}(A)$.*

Proof. 1 $\Rightarrow$ 8. Obvious.

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

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

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

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

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

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

5 $\Rightarrow$ 4. AA^{H} is Hermitian, and $(AA^{\mathsf{H}})^2 = AA^{\mathsf{H}}AA^{\mathsf{H}} = AA^{\mathsf{H}}$.

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

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

7 $\Rightarrow$ 1. Let $r = \operatorname{rank}(A)$. 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 $U^{\mathsf{H}}U = V^{\mathsf{H}}V = I_r$. Note that the columns of V is an orthonormal basis of $\operatorname{range}(A^{\mathsf{H}})$. Therefore, for any $x \in \operatorname{range}(A^{\mathsf{H}})$, $\|Ax\| = \|UV^{\mathsf{H}}x\| = \|V^{\mathsf{H}}x\| = \|x\|$. $\square$

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

$$A = \sum_{i=1}^{k} \lambda_i P_i \tag{14}$$

such that

1. *$\{\lambda_i\}$ are all real numbers and $\lambda_1 > \cdots > \lambda_k$.*
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 (14). For each i, let V_i be a matrix whose columns form an orthonormal basis of range(P_i). Then $V_i^{\mathsf{H}} V_i$ is an identity matrix, and $P_i = V_i V_i^{\mathsf{H}}$. For any distinct i and j,

$$V_i^{\mathsf{H}} V_j = (V_i^{\mathsf{H}} V_i) V_i^{\mathsf{H}} V_j (V_j^{\mathsf{H}} V_j) = V_i^{\mathsf{H}} P_i P_j V_j = 0.$$

Define

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

Then the columns of V are orthonormal. In addition,

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

Thus V is a unitary matrix. In addition,

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

where $\Lambda_i = \lambda_i V_i^{\mathsf{H}} V_i$, and Λ is the block diagonal matrix whose diagonal blocks are Λ_i. Note that Λ indeed a diagonal matrix since each Λ_i is diagonal. Thus $V \Lambda V^{\mathsf{H}}$ is an eigenvalue decomposition of A, with $\lambda_1, \ldots, \lambda_k$ 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 λ_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. □

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

$$A = \sum_{i=1}^{k} \sigma_i A_i \tag{15}$$

such that

1. *$\sigma_1 > \cdots > \sigma_k > 0$;*
2. *$\{A_i\}$ are all partial isometries, with $A_i A_j^{\mathsf{H}}$ and $A_i^{\mathsf{H}} A_j$ 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 (15). Since $A_i^{\mathsf{H}} A_j = 0$ for any distinct i and j, we have

$$A^{\mathsf{H}} A \;=\; \left(\sum_{i=1}^{k} \sigma_i A_i\right)^{\mathsf{H}} \left(\sum_{i=1}^{k} \sigma_i A_i\right) \;=\; \sum_{i=1}^{k} \sigma_i^2 A_i^{\mathsf{H}} A_i. \tag{16}$$

For each i, $A_i^{\mathsf{H}} A_i$ is a projection, as A_i is a partial isometry. Hence (16) is a decomposition specified in Theorem 8. Due to the uniqueness part of Theorem 8, σ_1, ..., σ_k and $A_1^{\mathsf{H}} A_1$, ..., $A_k^{\mathsf{H}} A_k$ are uniquely determined by A.

Now consider any two decompositions in the form of (15). According to what is proved above, we can formulate the decompositions as

$$A \;=\; \sum_{i=1}^{k} \sigma_i A_i \quad \text{and} \quad A \;=\; \sum_{i=1}^{k} \sigma_i \tilde{A}_i, \tag{17}$$

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

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

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

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

Thus

$$\left[\sum_{i=1}^{k} \sigma_i (A_i - \tilde{A}_i)\right] \left[\sum_{i=1}^{k} \sigma_i (A_i - \tilde{A}_i)\right]^{\mathsf{H}} \;=\; \sum_{i=1}^{k} \sigma_i^2 (A_i - \tilde{A}_i)(A_i - \tilde{A}_i)^{\mathsf{H}}. \tag{18}$$

According to (17), the left-hand side of (18) is zero. Hence $A_i = \tilde{A}_i$ for each i. Therefore, the two decompositions in (17) are indeed identical. The proof is complete. $\square$

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

Theorem 10 ([6, Theorem 1])**.** *Let X and Y be finite dimensional Hilbert spaces and $T : X \to Y$ be a linear operator. Then there exist unique orthogonal decompositions*

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

scalars $\sigma_1 > \cdots > \sigma_k$, and isometries $U_i : X_i \to Y_i$ $(i = 1, \ldots, k)$ such that

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

4 Understanding SVD as a change of basis

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

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*

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

   Indeed, the i*-th column of* A *is* $A_i = C_Y T(x_i)$*, namely the coordinate of* $T(x_i)$*.*

2. $T = C_Y^{-1} A C_X$*, meaning that applying* T *to any vector in* X *is equivalent to multiplying its coordinate by* A *and then using the result as the coordinate to locate a vector in* Y*.*

3. $C_X(\ker(T)) = \ker(A)$*, and* $C_Y(\mathrm{range}(T)) = \mathrm{range}(A)$*.*

4. 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* $T^{-1} : Y \to X$ *under the aforementioned bases for* X *and* Y*.*

5. A^* *represents* $T^* : Y^* \to X^*$ *under the bases for* X^* *and* Y^* *that are dual to* $\{x_1, \ldots, x_n\}$ *and* $\{y_1, \ldots, y_n\}$ *respectively.*

6. *Let* $\{x'_1, \ldots, x'_n\}$ *be a basis for* X*,* $\{y'_1, \ldots, y'_m\}$ *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*

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

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 $U\Sigma V^{\mathsf{H}}$ 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 Σ.

Recall the decompositions

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

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

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

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

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 $A^+ = \hat{U}\hat{\Sigma}^{-1}\hat{V}^{\mathsf{H}}$.

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 also leads 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 $U^{\mathsf{H}}U = I_n$; 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 $V^{\mathsf{H}}V = I_n$. In this case, $P = (A^{\mathsf{H}}A)^{\frac{1}{2}}$.*

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 $UU^{\mathsf{H}} = I_m$; 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 $VV^{\mathsf{H}} = I_m$. In this case, $Q = (AA^{\mathsf{H}})^{\frac{1}{2}}$.*

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

Proof. We only prove 1. Suppose that $W\Sigma Z^{\mathsf{H}}$ is an SVD of A.

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

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

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

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

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

If A is real, then W, Σ, and Z can all be real, ensuring P, U, Q, and V to be real. □

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

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 $x \in \mathbb{C}^n$,

$$x^{\mathsf{H}}Hx = \frac{1}{2}x^{\mathsf{H}}(A^{\mathsf{H}} + A)x = \Re(x^{\mathsf{H}}Ax) \leq |x^{\mathsf{H}}Ax| = |x^{\mathsf{H}}UPx| \leq \|Px\| = (x^{\mathsf{H}}P^2x)^{\frac{1}{2}} = (x^{\mathsf{H}}A^{\mathsf{H}}Ax)^{\frac{1}{2}}.$$

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

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

□

Proposition 4. *For matrices A_1 and $A_2 \in \mathbb{C}^{m\times n}$, $A_1^{\mathsf{H}}A_1 = A_2^{\mathsf{H}}A_2$ if and only if there exists a unitary matrix $U \in \mathbb{C}^{m\times m}$ such that $A_2 = UA_1$.*

Proof. The "if" part is trivial. We focus on the "only if" part. Let $V\Lambda V^{\mathsf{H}}$ be an eigenvalue decomposition of $A_1^{\mathsf{H}}A_1 = A_2^{\mathsf{H}}A_2$ such that the diagonal entries of Λ 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 $A_1 = W_1\Sigma V^{\mathsf{H}}$ and $A_2 = W_2\Sigma V^{\mathsf{H}}$. Set $U = W_2W_1^{\mathsf{H}}$. □

References

[1] Z.-H. Cao and L.-H. Feng. A note on variational representation for singular values of matrix. *Appl. Math. Comput.*, 143:559–563, 2003.

[2] C. Eckart and G. Young. A principal axis transformation for non-Hermitian matrices. *Bull. Amer. Math. Soc.*, 45:118–121, 1939.

[3] Ky Fan and A. J. Hoffman. Some metric inequalities in the space of matrices. *Proc. Amer. Math. Soc.*, 6:111–116, 1955.

[4] 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.

[5] C. Jordan. Mémoire sur les formes bilinéaires. *J. Math. Pures Appl.*, 19:35–54, 1874.

[6] A. Korányi. Around the finite-dimensional spectral theorem. *Amer. Math. Monthly*, 108:120–125, 2001.

[7] C. Lanczos. Linear systems in self-adjoint form. *Amer. Math. Monthly*, 65:665–679, 1958.

[8] R. Mathias. Singular values and singular value inequalities. In L. Hogben, editor, *Handbook of Linear Algebra.* CRC Press, Boca Raton, FL, 2013.

[9] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra.* SIAM, Philadelphia, 1997.