

UNIVERSITY OF CALIFORNIA, BERKELEY

LITERALLY EVERYTHING I KNOW ABOUT

Linear Algebra

Warren Alphonso

A very reductionist summary of LINEAR ALGEBRA AND ITS APPLICATIONS by Lay, Lay, and McDonald, LINEAR ALGEBRA DONE WRONG by Treil, ADVANCED LINEAR ALGEBRA by Surowski, and QUANTUM COMPUTATION AND QUANTUM INFORMATION by Nielsen and Chuang.

Last updated: August 28, 2019

Contents

1	Basic Notations	1
1.1	Vector Spaces	1
1.2	Linear Combinations	1
1.3	Linear Transformations	2
1.3.1	Linear Transformations as a Vector Space	4
1.4	Invertible Transformations and Isomorphisms	4
2	Systems of Linear Equations	7
2.1	Representations of Linear Systems	7
2.2	Solving Linear Systems	7
2.3	Fundamental Subspaces	9
2.4	Change of Basis	12
3	Determinants	13
3.1	Properties	13
3.2	Computing the Determinant	14
3.3	Cofactor Expansion	15
4	Spectral Theory	19
4.1	Definitions	19
4.2	Diagonalization	21
5	Inner Product Spaces	23
5.1	Inner Product	23
5.2	Orthogonality	26
5.3	Orthogonal Projection and Gram-Schmidt Orthogonalization	27
5.3.1	Orthogonal Complement	30
5.4	Least Square Solution	31
5.5	Adjoint of a Linear Transformation	32
5.6	Isometries and Unitary Operators	33
6	Structure of Operators	37
6.1	Schur Representation	37
6.2	Spectral Theorem for Self-Adjoint and Normal Operators	37
6.3	Singular Value and Polar Decompositions	40
6.3.1	Schmidt Decomposition	41
6.3.2	Singular Value Decomposition	43
6.4	Applications of Singular Value Decomposition	44
6.4.1	Operator Norm	44

6.4.2	Moore-Penrose Pseudoinverse	45
6.5	Structure of Orthogonal Matrices	46
7	Outer Product Representation	49
7.1	Dirac Notation	49
7.2	Adjoint and Hermitian Operators	50
7.3	Diagonalization	52
7.4	Tensor Products	53
7.5	Operator Functions	54
7.6	Commutator and Anti-commutator	54
7.7	Polar Decomposition	56
7.8	Singular Value Decomposition	56
8	Bilinear and Quadratic Forms	59
8.1	Main Definitions	59
8.2	Diagonalization of Quadratic Forms	60
8.2.1	Orthogonal Diagonalization	60
8.2.2	Non-orthogonal Diagonalization	61
8.3	Sylvester's Law of Inertia	62
8.4	Sylvester's Criterion of Positivity and Minimax Characterization of Eigenvalues	63
8.4.1	Sylvester's Criterion of Positivity	64
8.4.2	Minimax Characterization of Eigenvalues	65
9	Matrix Functions - WIP	67
9.1	Matrix Exponential	67

Chapter 7

Outer Product Representation

This chapter is largely adapted from Chapter 2 of Nielsen and Chuang. Much of it will be spent providing an additional perspective on the core ideas in the previous two chapters. New concepts, like tensor products, that are introduced in the second half of this chapter will be developed more formally later.

7.1 Dirac Notation

We can use inner products to derive a useful representations of linear operators. Before doing so, we'll define a convenient notation.

Definition 7.1.1. The **Dirac notation** represents vectors as

$$|\psi\rangle$$

which is also known as a **ket**. The dual of this same vector (dual spaces will be defined later) is represented as

$$\langle\psi|$$

which is also known as a **bra**. For now, it suffices to think of the dual of a vector as the complex conjugate transpose. This notations allows us to represent the inner product of two vectors as

$$\langle\phi|\psi\rangle$$

Using this new notation, we can define the outer product.

Definition 7.1.2. The **outer product representation** is a representation of linear operators which uses the inner product. Suppose $|v\rangle$ is a vector in inner product space V , and $|w\rangle$ is a vector in inner product space W . We define $|w\rangle\langle v|$ to be the *linear operator* from V to W defined by

$$\left(|w\rangle\langle v|\right)|v_1\rangle \equiv |w\rangle\langle v|v_1\rangle = \langle v|v_1\rangle |w\rangle$$

To make this representation feel less abstract, we consider an important result using the outer product.

Lemma 7.1.1 (Completeness Relation). *For an orthonormal basis $|i\rangle$,*

$$\sum_i |i\rangle\langle i| = I$$

Proof. We know any $v \in V$ can be written as $v = \sum_i v_i |i\rangle$ and that $\langle i|v\rangle = v_i$. Then

$$\left(\sum_i |i\rangle \langle i| \right) |v\rangle = \sum_i |i\rangle \langle i|v\rangle = \sum_i v_i |i\rangle = |v\rangle$$

Because this is true for all $|v\rangle$, it must be the identity operator. ■

One application of the Completeness Relation is the representation of any operator in the outer product notation. Suppose $A : V \rightarrow W$ is a linear operator and $|v_i\rangle$ and $|w_j\rangle$ are orthonormal bases for V and W , respectively. By using the Completeness Relation twice, we get

$$A = I_W A I_V = \sum_{ij} |w_j\rangle \langle w_j| A |v_i\rangle \langle v_i| = \sum_{ij} \langle w_j| A |v_i\rangle |w_j\rangle \langle v_i|$$

which is the **outer product representation** of A . This equation also shows that A has matrix element $\langle w_j| A |v_i\rangle$ in the i th column and j th row, with respect to input basis $|v_i\rangle$ and output basis $|w_j\rangle$.

To gain more familiarity with the outer product representation, we will prove the Cauchy-Schwarz Inequality.

Theorem 7.1.2 (Cauchy-Schwarz Inequality). *For any two vectors $|v\rangle$ and $|w\rangle$,*

$$|\langle v|w\rangle|^2 \leq \langle v|v\rangle \langle w|w\rangle$$

Proof. We use Gram-Schmidt to obtain an orthonormal basis $|i\rangle$, such that the first member of the basis is $|w\rangle/\sqrt{\langle w|w\rangle}$. Then

$$\begin{aligned} \langle v|v\rangle \langle w|w\rangle &= \sum_i \langle v|i\rangle \langle i|v\rangle \langle w|w\rangle \\ &\geq \frac{\langle v|w\rangle \langle w|v\rangle}{\langle w|w\rangle} \langle w|w\rangle \\ &= \langle v|w\rangle \langle w|v\rangle = |\langle v|w\rangle|^2 \end{aligned}$$

where the inequality in the second step follows from the fact that we only use the first member of our basis and thus drop some non-negative terms. ■

7.2 Adjoints and Hermitian Operators

Recall, if A is a linear operator on V , then there exists a unique linear operator A^\dagger (synonymous with A^* in quantum mechanics notation which we are adopting for this chapter) on V such that for all vectors $|v\rangle, |w\rangle \in V$,

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle)$$

This linear operator is the **adjoint** or **Hermitian conjugate** of A . Recall earlier that we also defined $|v\rangle^\dagger \equiv \langle v|$.

Definition 7.2.1. A mapping $f : V \rightarrow W$ between complex vector spaces is **antilinear** if

$$f(ax + by) = \bar{a}f(x) + \bar{b}f(y)$$

From Wikipedia: “antilinear maps occur in quantum mechanics in the study of time reversal and in spinor calculus.”

Lemma 7.2.1 (Antilinearity of the adjoint).

$$\left(\sum_i a_i A_i \right)^\dagger = \sum_i a_i^* A_i^\dagger$$

Proof. On the LHS: we know $(a+b)^* = a^* + b^*$. Thus, $(\sum_i a_i A_i)^\dagger = \sum_i (a_i A_i)^\dagger$.

On the RHS: $\sum_i a_i^* A_i^\dagger = (a_1 A_1)^\dagger + \cdots + (a_n A_n)^\dagger = \sum_i (a_i A_i)^\dagger$. ■

Recall, an operator A whose adjoint is itself is known as a **Hermitian** or **self-adjoint** operator.

An interesting usage of Hermitian operators that hint at their importance is the following theorem.

Theorem 7.2.2. *An arbitrary operator A can be decomposed into the sum $B + iC$ where B and C are Hermitian.*

Proof. Define $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$.

To prove B is Hermitian: $B^\dagger = \frac{A^\dagger+A}{2} = B$. To prove C is Hermitian: $C^\dagger = \frac{A^\dagger-A}{-2i} = C$.

Finally,

$$B + iC = \frac{A+A^\dagger}{2} + i \frac{A-A^\dagger}{2i} = \frac{A+A}{2} = A$$

■

Recall, *projector* operators are Hermitian operators. We can demonstrate this using outer product representation.

Suppose W is a k -dimensional vector *subspace* of the d -dimensional vector space V . We use Gram-Schmidt to construct an orthonormal basis $|1\rangle, \dots, |d\rangle$ for V such that $|1\rangle, \dots, |k\rangle$ is an orthonormal basis for W . Then, we define

$$P = \sum_{i=1}^k |i\rangle \langle i|$$

to be the projector onto subspace W . Notice that the Completeness Relation states that if $k = d$, $P = I$, which is exactly what a projector from $V \rightarrow V$ would be. This outer product representation of P makes it clear that P is Hermitian: since $(|v\rangle \langle v|)^\dagger = (|v\rangle \langle v|)$, antilinearity means

$$P^\dagger = \left(\sum_i^k |i\rangle \langle i| \right)^\dagger = \sum_i^k (|i\rangle \langle i|)^\dagger = \sum_i^k |i\rangle \langle i| = P$$

The Completeness Relation also makes it clear that the orthogonal complement of P is defined as $Q = I - P$, and that Q is a projector onto the vector space spanned by $|k+1\rangle, \dots, |d\rangle$.

Recall, an operator A is **normal** if $AA^\dagger = A^\dagger A$. Clearly, a Hermitian operator is also normal.

Recall, an operator U is **unitary** if $U^\dagger U = I$, so unitary operators are also normal. In Chapter 5, we proved unitary operators preserve the norm, but we can generalize this to prove that they also preserve inner products.

Theorem 7.2.3. *Let U be unitary. Then*

$$(U|v\rangle, U|w\rangle) = (|v\rangle, |w\rangle)$$

Proof.

$$(U|v\rangle, U|w\rangle) = \langle v|U^\dagger U|w\rangle = \langle v|I|w\rangle = \langle v|w\rangle = (|v\rangle, |w\rangle)$$

■

This result allows an elegant outer product representation of U . Let $|v_i\rangle$ be an orthonormal basis. We define

$$|w_i\rangle = U|v_i\rangle$$

so that $|w_i\rangle$ is also an orthonormal basis because unitary operators preserve inner products. Right multiplying both sides by $\langle v_i|$ and using the Completeness Relation, we get

$$|w_i\rangle\langle v_i| = U|v_i\rangle\langle v_i| \rightarrow U = \sum_i |w_i\rangle\langle v_i|$$

Recall, an operator A such that $(|v\rangle, A|v\rangle)$ is real and non-negative is called a **positive operator** or positive semidefinite. If this same inner product is *strictly* positive for all $|v\rangle \neq 0$, then A is **positive definite**.

Theorem 7.2.4 (Hermiticity of positive operators). *A positive operator A is Hermitian.*

Proof. Since we can write $A = B + iC$ where B and C are Hermitian, we can use definition of a positive operator to write

$$\begin{aligned} (|v\rangle, A|v\rangle) &= (|v\rangle, B|v\rangle + iC|v\rangle) \\ &= (|v\rangle, B|v\rangle) + (|v\rangle, iC|v\rangle) \\ &= (B|v\rangle, |v\rangle) + (-iC|v\rangle, |v\rangle) \end{aligned}$$

where the third equality follows from B and C being Hermitian. This implies that $A = B - iC$, so

$$A^\dagger = (B - iC)^\dagger = B^\dagger + iC^\dagger = B + iC = A$$

■

7.3 Diagonalization

The spectral theorem for normal operators proved in the previous chapter ($N = UDU^\dagger$ where U is a unitary matrix of eigenvectors and D is a diagonal matrix of corresponding eigenvalues) states an operator is diagonalizable if and only if it is normal. We can also write the diagonal representation of a normal matrix in outer product representation.

Theorem 7.3.1. *A diagonal representation for a normal operator $N : V \rightarrow V$ is*

$$N = \sum_i \lambda_i |i\rangle\langle i|$$

where $|i\rangle$ are an orthonormal set of eigenvectors of V with corresponding eigenvalues λ_i .

Proof. Let $|e\rangle$ be an orthonormal basis of N . Since each basis vector $|e\rangle$ can be understood as providing a coordinate of $|v\rangle$ in that basis through $\langle e|v\rangle$, we know $|e\rangle\langle e|$ is the 0 matrix except with a 1 on the e th diagonal entry. The spectral theorem tells us that if U is unitary and D is diagonal, then

$$N = UDU^\dagger = U\left(\sum_{i=1}^n \lambda_i |e\rangle\langle e|\right)U^\dagger = \sum_{i=1}^n \lambda_i U|e\rangle\left(U|e\rangle\right)^\dagger$$

Now we define $U|i\rangle$ to be the normalized eigenvectors $|i\rangle$ to get

$$N = \sum_{i=1}^n \lambda_i |i\rangle \langle i|$$

■

As an example, we will introduce the Pauli $Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ matrix, which is significant in quantum computation, and write its diagonal representation.

The characteristic equation $\det(A - \lambda I)$ determines that the eigenvalues for Z are 1 and -1. The corresponding eigenvectors are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which we will denote $|0\rangle$ and $|1\rangle$, respectively. This means our diagonal representation is

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = |0\rangle \langle 0| - |1\rangle \langle 1|$$

7.4 Tensor Products

Tensor products are probably the most unintuitive Linear Algebra concept I've come across thus far. This section is a very top-level overview of the subject; later chapters will cover this topic in more detail.

Definition 7.4.1. Suppose V and W are m and n dimensional vector spaces, respectively. Then $V \otimes W$ is the **tensor product** of these two spaces and is an mn dimensional vector space. The elements of $V \otimes W$ are linear combinations of **tensor products** $|v\rangle \otimes |w\rangle$. If $|i\rangle$ and $|j\rangle$ are orthonormal bases for spaces V and W , then $|i\rangle \otimes |j\rangle$ is a basis for $V \otimes W$.

The tensor product satisfies the following properties:

1. For scalar z , $|v\rangle \in V$, and $|w\rangle \in W$,

$$z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle)$$

2. For $|v_1\rangle, |v_2\rangle \in V$ and $|w\rangle \in W$,

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle$$

3. For $|v\rangle \in V$ and $|w_1\rangle, |w_2\rangle \in W$,

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle$$

If A and B are linear operators on V and W , respectively, then we can define a linear operator $A \otimes B$ on $V \otimes W$,

$$(A \otimes B)(|v\rangle \otimes |w\rangle) \equiv A|v\rangle \otimes B|w\rangle$$

We can extend this definition of $A \otimes B$ to ensure linearity of $A \otimes B$ so

$$(A \otimes B) \left(\sum_i \alpha_i |v_i\rangle \otimes |w_i\rangle \right) \equiv \sum_i \alpha_i A|v_i\rangle \otimes B|w_i\rangle$$

An arbitrary linear operator C mapping $V \otimes W$ to $V' \otimes W'$ can be represented as a linear combination of tensor products of operators mapping $A : V \rightarrow V'$ and $B : W \rightarrow W'$,

$$C = \sum_i c_i A_i \otimes B_i$$

We can use the inner products on V and W to define an inner product on $V \otimes W$,

$$\left(\sum_i \alpha_i |v_i\rangle \otimes |w_i\rangle, \sum_j \beta_j |v_j\rangle \otimes |w_j\rangle \right) \equiv \sum_{ij} \alpha_i^* \beta_j \langle v_i | v_j \rangle \langle w_i | w_j \rangle$$

To make all this less abstract, we consider a matrix representation known as the *Kronecker product*. Suppose A is an $m \times n$ matrix and B is a $p \times q$ matrix. Then

$$A \otimes B \equiv \overbrace{\begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{2n}B \\ \vdots & \vdots & & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{bmatrix}}^{nq} \Bigg\}^{mp}$$

where $A_{11}B$ denotes the $p \times q$ submatrix B multiplied by the constant A_{11} .

For example,

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \times 2 \\ 1 \times 3 \\ 2 \times 2 \\ 2 \times 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 6 \end{bmatrix}$$

and the tensor product of Pauli X and Y is

$$X \otimes Y = \begin{bmatrix} 0 \cdot Y & 1 \cdot Y \\ 1 \cdot Y & 0 \cdot Y \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}$$

The notation $|\psi\rangle^{\otimes k}$ denotes $|\psi\rangle$ tensored with itself k times.

Do problems from Nielsen and Chuang.

7.5 Operator Functions

Given a function $f : C \rightarrow C$, we can define a corresponding matrix function. Let $A = \sum_a a |a\rangle \langle a|$, then

$$f(A) \equiv \sum_a f(a) |a\rangle \langle a|$$

We can use this construction to define the square root of a positive operator, the logarithm of a positive-definite operator, or the exponential of a normal operator.

One important matrix function is the trace. Since we know trace is *cyclic* (that is, $\text{tr}(AB) = \text{tr}(BA)$) then $\text{tr}(UAU^\dagger) = \text{tr}(U^\dagger UA) = \text{tr}(A)$.

Suppose $|\psi\rangle$ is a unit vector and A is an arbitrary operator. To evaluate $\text{tr}(A|\psi\rangle \langle \psi|)$ use Gram-Schmidt to extend $|\psi\rangle$ to an orthonormal basis $|i\rangle$ which includes $|\psi\rangle$ as the first element. Then, since all except the first basis vector are orthogonal to $|\psi\rangle$,

$$\text{tr}(A|\psi\rangle \langle \psi|) = \sum_i \langle i | A | \psi \rangle \langle \psi | i \rangle = \langle \psi | A | \psi \rangle$$

7.6 Commutator and Anti-commutator

Definition 7.6.1. The **commutator** between two operators A and B is defined to be

$$[A, B] \equiv AB - BA$$

If $[A, B] = 0$, then we say A **commutes** with B . The **anti-commutator** is defined to be

$$\{A, B\} \equiv AB + BA$$

If $\{A, B\} = 0$, then we say A **anti-commutes** with B .

Many important properties of pairs of operators can be deduced from their commutator or anti-commutator. A useful relation is the connection between the commutator and the property of being able to *simultaneously diagonalize* Hermitian operators A and B , that is, write $A = \sum_i a_i |i\rangle \langle i|$ and $B = \sum_i b_i |i\rangle \langle i|$, where $|i\rangle$ is a common set of eigenvectors for A and B .

Theorem 7.6.1 (Simultaneous diagonalization theorem). *Suppose A and B are Hermitian operators. Then $[A, B] = 0$ if and only if there exists an orthonormal basis such that both A and B are diagonal with respect to that basis.*

Proof. Proving if is simple, since if $A = UA_1U^{-1}$ and $B = UB_1U^{-1}$, then clearly $AB - BA = 0$.

To prove only if, let $|a, j\rangle$ be an orthonormal basis for the eigenspace V_a of A with eigenvalue a ; the index j being used to label possible degeneracies. Because we know $AB = BA$, we can write

$$AB|a, j\rangle = BA|a, j\rangle = aB|a, j\rangle$$

which means $B|a, j\rangle$ is an element of the eigenspace V_a . Let P_a denote the projector onto V_a and define $B_a = P_a B P_a$. Since projectors are Hermitian, we know B_a is also Hermitian so it must have a spectral decomposition in terms of an orthonormal set of eigenvectors that span V_a . Let's call these eigenvectors $|a, b, k\rangle$, where a and b label the eigenvalues of A and B_a , and k indexes degeneracies. Note that $B|a, b, k\rangle$ is in V_a , so $B|a, b, k\rangle = P_a B|a, b, k\rangle$. Additionally, we know $P_a B|a, b, k\rangle = |a, b, k\rangle$, so

$$B|a, b, k\rangle = P_a B P_a |a, b, k\rangle = B_a |a, b, k\rangle$$

This means $|a, b, k\rangle$ is an eigenvector of B with eigenvalue b , and therefore $|a, b, k\rangle$ is a spanning orthonormal set of eigenvectors of both A and B . Thus, A and B are simultaneously diagonalizable. ■

Another use of the commutator and anti-commutator is the following identity:

$$AB = \frac{[A, B] + \{A, B\}}{2}$$

Lemma 7.6.2. *Suppose $[A, B] = 0$, $\{A, B\} = 0$, and A is invertible. Then $B = 0$.*

Proof. Using the identity we just defined, we know $AB = \frac{0+0}{2}$. Since A is invertible we left multiply both sides by A^{-1} to get $B = 0$. ■

Three more easily verifiable identities:

1. $[A, B]^\dagger = [B^\dagger, A^\dagger]$
2. $[A, B] = -[B, A]$
3. If A, B are Hermitian, then $i[A, B]$ is also Hermitian.

7.7 Polar Decomposition

We don't understand *general* linear operators incredibly well, but we do have a good understanding of *unitary* and *positive* operators. The polar and singular value decompositions allow us to take advantage of this by breaking general linear operators into products of unitary and positive operators.

Theorem 7.7.1 (Polar decomposition). *Let A be a linear operator on V . Then there exists unitary U and positive operators J and K such that*

$$A = UJ = KU$$

where $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$. Additionally, if A is invertible, then U is unique.

Proof. We know J (which is the modulus of A) is a positive operator (which we know are normal), so we can give it a spectral decomposition,

$$J = \sum_i \lambda_i |i\rangle \langle i| \quad (\lambda_i \geq 0)$$

Define $|\psi_i\rangle = A|i\rangle$, so that $\langle \psi_i | \psi_i \rangle = \lambda_i^2$.

Now consider only the positive eigenvalues. With these eigenvalues, define $|e_i\rangle = |\psi_i\rangle / \lambda_i$, so the $|e_i\rangle$ are normalized and are orthogonal, since if $i \neq j$, then $\langle e_i | e_j \rangle = \langle i | A^\dagger A | j \rangle / \lambda_i \lambda_j = 0$.

Now we use Gram-Schmidt to extend the orthonormal set $|e_i\rangle$ to form an orthonormal basis, which we'll also call $|e_i\rangle$.

We define the unitary operator $U = \sum_i |e_i\rangle \langle i|$. When $\lambda_i \neq 0$, we have $UJ|i\rangle = \lambda_i |e_i\rangle = |\psi_i\rangle = A|i\rangle$. When $\lambda_i = 0$, we have $UJ|i\rangle = 0 = |\psi_i\rangle$.

Since A and UJ agree on the basis $|i\rangle$, we know $A = UJ$.

J is unique since left-multiplying $A = UJ$ by $A^\dagger = JU^\dagger$ gives $A^\dagger A = J^2$, from which we get $J = \sqrt{A^\dagger A}$. The equation $A = UJ$ also means that if A is invertible, so is J , which means U is uniquely determined by $U = AJ^{-1}$.

Proving the right polar decomposition $A = KU$ follows since

$$A = UJ = UJU^\dagger U = KU$$

where $K = UJU^\dagger$ is a positive operator. And since $AA^\dagger = KU U^\dagger K = K^2$, we have $K = \sqrt{AA^\dagger}$. ■

7.8 Singular Value Decomposition

The singular value decomposition combines polar decomposition and the spectral theorem.

Theorem 7.8.1 (Singular value decomposition). *Let A be a linear operator on V . Then there exist unitary matrices U and V , and a diagonal matrix D with non-negative entries such that*

$$A = UDV$$

The diagonal entries of D are called the singular values of A .

Proof. By the polar decomposition, $A = SJ$, for unitary S and positive operator J . By the spectral theorem, $J = TDT^\dagger$, for unitary T and diagonal D with non-negative entries.

Now we just set $U = ST$ and $V = T^\dagger$ to get

$$A = SJ = STD T^\dagger = UDV$$

■