Note 2j: Complex Linear Algebra

1 Overview and Motivation

Thus far in EECS16A and EECS16B, we covered many linear algebra results while assuming our matrices and vectors had *real* entries. This makes sense in most cases – for instance, we will never have an imaginary resistance or complex-valued current – so the relevant linear systems have real numbers. But sometimes we will just have to use complex vectors and complex matrices. We saw a taste of this when we did diagonalization of matrices which happened to have complex eigenvalues, or even phasor-domain calculations, but complex-domain calculations are really common in signal processing. Thus, we will review all of the linear algebra we covered in 16A and 16B, and show how it generalizes to the complex-valued case.

The main distinction that we emphasize in this note is that of the *complex inner product* and how it differs from the inner product we're familiar with for real vectors, i.e., $\langle \vec{x}, \vec{y} \rangle = \vec{y}^\top \vec{x}$. This is covered in Section 3.

The critical reading for this note is Sections 3 and 4; the remaining sections are not a core part of this course and thus do not need to be internalized, but are there for your reference.

2 Foundations

This is the warmup content. We start with the basic linear algebra definitions from 16A. The main distinction is that the scalars in everything are allowed to be complex numbers.

Key Idea 1 (Foundational Ideas of Complex Linear Algebra)
The most foundational ideas of complex linear algebra are exactly the same as real linear algebra, *except* all scalars are allowed to be complex numbers. Most of the same theorems apply.

2.1 Vectors and Vector Spaces

Definition 2 (Complex Vector Space)
A *complex vector space* V is a set of elements, called *vectors*, which satisfy the following properties relating to vector addition:

(VA1) Associative: $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ for any $\vec{u}, \vec{v}, \vec{w} \in V$.

(VA2) Commutative: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ for any $\vec{u}, \vec{v} \in V$.

(VA3) Additive identity: there exists a so-called *additive identity* $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for any $\vec{v} \in V$.

(VA4) Additive inverse: for any $\vec{v} \in V$, there exists a so-called *additive inverse* $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = \vec{0}$.

(VA5) Closure under vector addition: if $\vec{u}, \vec{v} \in V$ then $\vec{u} + \vec{v} \in V$.

and scalar multiplication:

(SM1) Associative: $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$ for any $\alpha, \beta \in \mathbb{C}$ and $\vec{v} \in V$.

(SM2) Multiplicative identity: there exists a so-called *multiplicative identity* $1 \in \mathbb{C}$ such that $1 \cdot \vec{v} = \vec{v}$ for any $\vec{v} \in V$.

(SM3) Distributive in vector addition: $\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v}$ for any $\alpha \in \mathbb{C}$ and $\vec{u}, \vec{v} \in V$.

(SM4) Distributive in scalar addition: $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$ for any $\alpha, \beta \in \mathbb{C}$ and $\vec{v} \in V$.

(SM5) Closure under scalar multiplication: if $\alpha \in \mathbb{C}$ and $\vec{v} \in V$ then $\alpha\vec{v} \in V$.

Note that this is the same definition of a vector space as we are familiar with, only we are able to use complex scalars.

Some prototypical examples you should keep in mind of complex vector spaces:

- $\mathbb{C}$, the set of complex numbers, is a complex vector space;
- $\mathbb{C}^n$, the set of length-n tuples of complex numbers, is a complex vector space;
- $\mathbb{C}^{m \times n}$, the set of $m \times n$ matrices of complex numbers, is a complex vector space.

Note that neither $\mathbb{R}$ nor $\mathbb{R}^n$ nor $\mathbb{R}^{m \times n}$ are complex vector spaces.

Definition 3 (Complex Subspace)
A *subspace* of a complex vector space V is a subset of V which is also a complex vector space.

Definition 4 (Complex Linear Combination)
A *linear combination* of a finite set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a sum of the form

$$\sum_{i=1}^{n} \alpha_i \vec{v}_i \tag{1}$$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$.

Definition 5 (Complex Linear Independence)
A set S of (complex) vectors is *linearly independent* if the only (complex) linear combinations of finite subsets[a] of S which sum to $\vec{0}$ themselves have all coefficients equal to 0, i.e.,

$$\sum_{i=1}^{n} \alpha_i \vec{v}_i = \vec{0} \implies \alpha_1 = \cdots = \alpha_n = 0 \qquad \text{for all } \alpha_1, \ldots, \alpha_n \in \mathbb{C}, \vec{v}_1, \ldots, \vec{v}_n \in S, \text{ and } n \in \mathbb{N}. \tag{2}$$

Otherwise, we say S is *linearly dependent*.

[a] This condition is to make all sums well-defined when S is infinite; if S is finite, it is equivalent to look at all linear combinations of S itself, which is the more familiar definition we are used to.

Definition 6 (Complex Span)
The *span* of a set of (complex) vectors S is the set of all (complex) linear combinations of finite subsets of S:

$$\text{Span}(S) = \left\{ \sum_{i=1}^{n} \alpha_i \vec{v}_i \middle| \alpha_1, \ldots, \alpha_n \in \mathbb{C}, \vec{v}_1, \ldots, \vec{v}_n \in S, n \in \mathbb{N} \right\} \tag{3}$$

Theorem 7 (Complex Span is Complex Subspace)
For any complex vector space V and subset $S \subseteq V$, we have that $\text{Span}(S)$ is a complex subspace of V.

Definition 8 (Basis)
Let V be a complex vector space. A *basis* B of V is a subset of V such that:

- B is linearly independent; and
- $\text{Span}(B) = V$.

Theorem 9
Let V be a complex vector space. All bases of V have the same number of vectors.

Definition 10 (Dimension)
Let V be a complex vector space. The *dimension* of V is the number of vectors in any basis of V.

2.2 Matrices and Linear Transformations

Definition 11 (Complex Linear Transformation)
Let V and W be two complex vector spaces. A *linear transformation* $T\colon V \to W$ is a function such that

$$T(\alpha\vec{x} + \beta\vec{y}) = \alpha T(\vec{x}) + \beta T(\vec{y}) \qquad \text{for all } \alpha, \beta \in \mathbb{C} \text{ and } \vec{x}, \vec{y} \in V \tag{4}$$

Definition 12 (Complex Matrix)
An $m \times n$ *matrix* A is an element of $\mathbb{C}^{m \times n}$, i.e., it is an $m \times n$ array of complex numbers:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix} \tag{5}$$

where $\vec{a}_1, \ldots, \vec{a}_n$ are the *columns* of A.

Theorem 13 (Matrix Operations)
Matrices have the following operations:

(MO1) (Matrix addition.) If $A, B \in \mathbb{C}^{m \times n}$, then

$$A + B = \begin{bmatrix} \vec{a}_1 + \vec{b}_1 & \cdots & \vec{a}_n + \vec{b}_n \end{bmatrix}. \tag{6}$$

(MO2) (Scalar multiplication.) If $A \in \mathbb{C}^{m \times n}$ and $\alpha \in \mathbb{C}$, then

$$\alpha A = \begin{bmatrix} \alpha \vec{a}_1 & \cdots & \alpha \vec{a}_n \end{bmatrix}. \tag{7}$$

(MO3) (Matrix-vector multiplication.) If $A \in \mathbb{C}^{m \times n}$ and $\vec{x} \in \mathbb{C}^n$, then

$$A\vec{x} = \sum_{i=1}^{n} x_i \vec{a}_i. \tag{8}$$

(MO4) (Matrix multiplication.) If $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times p}$, then

$$AB = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix}. \tag{9}$$

In this way, $m \times n$ matrices are linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

2.3 Fundamental Subspaces of a Matrix

Definition 14 (Column Space)
Let $A \in \mathbb{C}^{m \times n}$ be a matrix. The *column space* of A is the span of its columns:

$$\operatorname{Col}(A) = \operatorname{Span}(\vec{a}_1, \ldots, \vec{a}_n) \subseteq \mathbb{C}^m. \tag{10}$$

Definition 15 (Null Space)
Let $A \in \mathbb{C}^{m \times n}$ be a matrix. The *null space* of A is the set of all vectors in $\mathbb{C}^n$ which are sent by A to $\vec{0}_m$:

$$\operatorname{Null}(A) = \left\{ \vec{x} \in \mathbb{C}^n \,\middle|\, A\vec{x} = \vec{0}_m \right\} \subseteq \mathbb{C}^n. \tag{11}$$

Theorem 16 (Column Space and Null Space are Subspaces)
Let $A \in \mathbb{C}^{m \times n}$ be a matrix. Then $\operatorname{Col}(A)$ is a subspace of $\mathbb{C}^m$ and $\operatorname{Null}(A)$ is a subspace of $\mathbb{C}^n$.

Definition 17 (Rank)
Let $A \in \mathbb{C}^{m \times n}$ be a matrix. Then the *rank* of A is the dimension of its column space:

$$\operatorname{rank}(A) := \dim(\operatorname{Col}(A)). \tag{12}$$

Theorem 18 (Rank-Nullity Theorem)
Let $A \in \mathbb{C}^{m \times n}$ be a matrix. Then we have

$$\operatorname{rank}(A) + \dim(\operatorname{Null}(A)) = n. \tag{13}$$

Definition 19 (Matrix Inverse)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. An *inverse* of A is a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.

Theorem 20 (Matrix Inverse Needs Full Rank)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix.

(i) A^{-1} exists if and only if $\text{rank}(A) = n$.

(ii) If A^{-1} exists then it is unique.

2.4 Eigenvalues and Eigenvectors

Definition 21 (Eigenvalue and Eigenvector)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. Then $(\lambda, \vec{v})$ is an *eigenvalue-eigenvector* pair for A if and only if

$$A\vec{v} = \lambda\vec{v}. \tag{14}$$

Definition 22 (Characteristic Polynomial)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. The *characteristic polynomial* of A is the $n \times n$ polynomial given by

$$p_A(\lambda) := \det(A - \lambda I_n). \tag{15}$$

Theorem 23 (Roots of Characteristic Polynomial are Eigenvalues)
The n roots of p_A are the n eigenvalues of A.

Definition 24 (Multiplicities)
Let $A \in \mathbb{C}^{n \times n}$ have distinct eigenvalues $\lambda_1, \ldots, \lambda_d$.

- The *algebraic multiplicity* m_A^a of λ_i is the multiplicity of λ_i as a root of the characteristic polynomial $p_A(\lambda)$. In other words, factoring the characteristic polynomial into linear factors:

  $$p_A(\lambda) = \prod_{i=1}^{d} (\lambda - \lambda_i)^{m_i} \tag{16}$$

  we have

  $$m_A^a(\lambda_i) = m_i. \tag{17}$$

- The *geometric multiplicity* $m_A^g(\lambda_i)$ of λ_i is the number of linearly independent eigenvectors of A with eigenvalue λ_i. In other words,

  $$m_A^g(\lambda_i) = \dim(\text{Null}(A - \lambda_i I_n)). \tag{18}$$

Theorem 25 (Results on Multiplicities)
Let $A \in \mathbb{C}^{n \times n}$ have distinct eigenvalues $\lambda_1, \dots, \lambda_d$.

(i) We have $\sum_{i=1}^{d} m_A^a(\lambda_i) = n$.

(ii) We have $m_A^a(\lambda_i) \geq m_A^g(\lambda_i)$ for every i.

(iii) The following are equivalent:

(a) For all eigenvalues λ of A, we have $m_A^a(\lambda) = m_A^g(\lambda)$;

(b) $\sum_{i=1}^{d} m_A^g(\lambda_i) = n$.

3 Geometry

Here is where the picture gets a bit different than the real vector case. Remember that in the real vector case, most of our geometry – that is, lengths and angles – comes from our ideas of inner products and norms.

Recall that the inner product on $\mathbb{R}^n$ (in this note, denoted by $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$) is given by

$$\langle \vec{x}, \vec{y} \rangle_{\mathbb{R}^n} = \vec{y}^\top \vec{x} = \sum_{i=1}^{n} x_i y_i \tag{19}$$

and that the norm on $\mathbb{R}^n$ (in this note, denoted by $\|\cdot\|_{\mathbb{R}^n}$) is given by

$$\|\vec{x}\|_{\mathbb{R}^n}^2 = \langle \vec{x}, \vec{x} \rangle = \sum_{i=1}^{n} x_i^2. \tag{20}$$

We know that we want the norm of any vector to be a non-negative real number denoting its length – this is where we get most of our geometric ideas from. But if we blindly apply these definitions to the case of $\mathbb{C}^n$, we get weird results. Suppose, for example, we apply the definitions to $\begin{bmatrix} 1 \\ 2\mathrm{j} \end{bmatrix} \in \mathbb{C}^2$:

$$\begin{aligned} \left\| \begin{bmatrix} 1 \\ 2\mathrm{j} \end{bmatrix} \right\|_{\mathbb{R}^2}^2 &= 1^2 + (2\mathrm{j})^2 && (21) \\ &= 1 - 4 && (22) \\ &= -3. && (23) \end{aligned}$$

This means that the definition of norm on $\mathbb{R}^2$ applied to the vector $\begin{bmatrix} 1 \\ 2\mathrm{j} \end{bmatrix}$ gives paradoxical results – after all, it is impossible to have negative squared length. Since the norm comes from the inner product, this choice of inner product and norm does not yield a useful geometry for $\mathbb{C}^n$.

3.1 Norm

With this in mind, we define what a norm is on a general complex vector space, and then give a new norm on $\mathbb{C}^n$ specifically.

Definition 26 (Complex Norm)
Let V be a complex vector space. A *norm* $\|\cdot\|_V$ on V is any function $V \to \mathbb{R}$ such that:

(N1) Triangle inequality:
$$\|\vec{x}+\vec{y}\|_V \leq \|\vec{x}\|_V + \|\vec{y}\|_V \qquad \text{for all } \vec{x},\vec{y} \in V. \tag{24}$$

(N2) Homogeneity:
$$\|\alpha\vec{x}\|_V = |\alpha|\|\vec{x}\|_V \qquad \text{for all } \alpha \in \mathbb{C} \text{ and } \vec{x} \in V. \tag{25}$$

(N3) Positive definiteness:
$$\|\vec{x}\|_V = 0 \implies \vec{x} = \vec{0}. \tag{26}$$

Concept Check: Show that $\|\cdot\|_{\mathbb{R}^n}$ is not a complex norm.

Now, instead of taking the sum of squared entries as a measure of squared length, we can use the more geometrically motivated idea of taking the sum of the squared *magnitudes* of each entry. Since the magnitude is already a measure of length of a complex number, the Pythagorean theorem validates this idea as reasonable.

Definition 27 (Norm on $\mathbb{C}^n$)
The norm $\|\cdot\|$ on $\mathbb{C}^n$ is given by

$$\|\vec{x}\|^2 = \sum_{i=1}^{n} |x_i|^2 \qquad \text{for all } \vec{x} \in \mathbb{C}^n. \tag{27}$$

Concept Check: Show that the norm on $\mathbb{C}^n$ given by Definition 27 fills all the properties of Definition 26, and also agrees with the real vector norm $\|\cdot\|_{\mathbb{R}^n}$ on vectors of entirely real numbers.

Using this idea of norm on our earlier example, we have

$$\left\|\begin{bmatrix}1\\2\mathrm{j}\end{bmatrix}\right\|^2 = |1|^2 + |2\mathrm{j}|^2 \tag{28}$$
$$= 1^2 + 2^2 \tag{29}$$
$$= 5. \tag{30}$$

This is a positive number and thus possibly a reasonable notion of length of the vector $\begin{bmatrix}1\\2\mathrm{j}\end{bmatrix}$. Indeed, since our norm fulfills all the criteria of Definition 27, it will give a good geometry for $\mathbb{C}^n$.

This idea also allows us to define a corresponding complex Frobenius norm.

Definition 28 (Frobenius Norm on $\mathbb{C}^{m\times n}$)
The complex Frobenius norm $\|\cdot\|_F$ on $\mathbb{C}^{m\times n}$ is given by

$$\|A\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 \qquad \text{for all } A \in \mathbb{C}^{m\times n}. \tag{31}$$

Concept Check: Show that the norm on $\mathbb{C}^{m\times n}$ given by Definition 28 fills all the properties of Definition 26, and also agrees with the real Frobenius norm on matrices of entirely real numbers.

3.2 Inner Products

We can also define what an inner product is on a complex vector space, and then on $\mathbb{C}^n$ in particular.

Definition 29 (Complex Inner Product)
Let V be a complex vector space. An *inner product* $\langle \cdot, \cdot \rangle_V$ on V is any function $V \times V \to \mathbb{C}$ such that:

(IP1) Conjugate symmetry:
$$\langle \vec{x}, \vec{y} \rangle_V = \overline{\langle \vec{y}, \vec{x} \rangle_V} \qquad \text{for all } \vec{x}, \vec{y} \in V. \tag{32}$$

(IP2) Linearity in the first argument:
$$\langle \alpha\vec{x} + \beta\vec{y}, \vec{z} \rangle_V = \alpha \langle \vec{x}, \vec{z} \rangle_V + \beta \langle \vec{y}, \vec{z} \rangle_V \qquad \text{for all } \alpha, \beta \in \mathbb{C} \text{ and } \vec{x}, \vec{y}, \vec{z} \in V. \tag{33}$$

(IP3) Positive definiteness:
$$\langle \vec{x}, \vec{x} \rangle_V \in \mathbb{R}_{\geq 0} \qquad \text{for all } \vec{x} \in V \tag{34}$$
$$\text{if} \quad \langle \vec{x}, \vec{x} \rangle_V = 0 \quad \text{then} \quad \vec{x} = \vec{0}. \tag{35}$$

Concept Check: Show that $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ is not a complex inner product.

To get a notion of inner product on $\mathbb{C}^n$, recall that we want to find an inner product such that $\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2 = \sum_{i=1}^n |x_i|^2$. Recall that $|x_i|^2 = x_i \cdot \overline{x_i}$. So, it follows that
$$\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2 = \sum_{i=1}^n |x_i|^2 = \sum_{i=1}^n x_i \cdot \overline{x_i} = \left(\overline{\vec{x}}\right)^\top \vec{x}. \tag{36}$$

This motivates us to define the conjugate transpose as its own operation, and then the complex inner product.

Definition 30 (Conjugate Transpose)
If $\vec{x} \in \mathbb{C}^n$ is a complex column vector, the *conjugate transpose* $\vec{x}^*$ of $\vec{x}$ is the length-n row vector given by
$$\vec{x}^* := \left(\overline{\vec{x}}\right)^\top. \tag{37}$$

If $A \in \mathbb{C}^{m \times n}$ is a complex matrix, its conjugate transpose $A^* \in \mathbb{C}^{n \times m}$ is given by
$$A^* := \left(\overline{A}\right)^\top. \tag{38}$$

With this notation, we see that
$$\langle \vec{x}, \vec{x} \rangle = \vec{x}^* \vec{x}. \tag{39}$$

This motivates the following inner product.

Definition 31 (Inner Product on $\mathbb{C}^n$)
The inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^n$ is given by
$$\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i \cdot \overline{y_i} = \vec{y}^* \vec{x} \qquad \text{for all } \vec{x}, \vec{y} \in \mathbb{C}^n. \tag{40}$$

Warning 32 (Order of Vectors in Inner Product)
Note that $\langle \vec{x}, \vec{y} \rangle = \vec{y}^* \vec{x}$, *not* $\vec{x}^* \vec{y}$. The two expressions are complex conjugates of each other.

Concept Check: Show that the inner product on $\mathbb{C}^n$ given by Definition 31 fills all the properties of Definition 29, agrees with the real vector inner product $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ on vectors of entirely real numbers, and also obeys $\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$ for all $\vec{x} \in \mathbb{C}^n$.

We can also define an inner product on $\mathbb{C}^{m \times n}$ which begets the complex Frobenius norm.

Definition 33 (Frobenius Inner Product on $\mathbb{C}^{m \times n}$)
The Frobenius inner product $\langle \cdot, \cdot \rangle_F$ on $\mathbb{C}^{m \times n}$ is given by

$$\langle A, B \rangle_F = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \cdot \overline{b_{ij}} = \operatorname{tr}(B^* A) \qquad \text{for all } A, B \in \mathbb{C}^{m \times n}. \tag{41}$$

Concept Check: Show that the inner product on $\mathbb{C}^{m \times n}$ given by Definition 33 fills all the properties of Definition 29, agrees with the real Frobenius inner product on vectors of entirely real numbers, and also obeys $\langle A, A \rangle_F = \|A\|_F^2$ for all $A \in \mathbb{C}^{m \times n}$.

Essentially, we substituted the conjugate transpose $*$ for the transpose $\top$ in our definitions. Because inner products and norms are so important, this substitution ends up being the primary difference between theorems in the real case and in the complex case.

Key Idea 34
When converting a theorem or proof from the real case to the complex case, it is often enough to substitute conjugate transpose in for transpose everywhere in the theorem or proof. The rest usually works out the same.

3.3 Orthogonality and Projections

Now that we have developed the theory of complex norms and inner products, the projection story is *exactly* the same. In previous notes, we were very careful to use inner product notation, and only use generic inner product properties which also hold for the complex inner product. So, all the orthogonality proofs remain the same (and indeed would remain the same for generic inner product spaces).

3.3.1 Orthogonality and Orthonormality

Definition 35 (Orthogonal Vectors)

- Let $\vec{x}, \vec{y} \in \mathbb{C}^n$. Then $\vec{x}$ and $\vec{y}$ are *orthogonal* if and only if

$$\langle \vec{x}, \vec{y} \rangle = 0. \tag{42}$$

- Let $S \subseteq \mathbb{C}^n$ be a set of vectors. Then S is an *orthogonal set* if and only if every pair of distinct vectors in S is orthogonal, i..e,

$$\langle \vec{x}, \vec{y} \rangle = 0 \qquad \text{for all } \vec{x}, \vec{y} \in S \text{ with } \vec{x} \neq \vec{y}. \tag{43}$$

Theorem 36
Suppose $S \subseteq \mathbb{C}^n$ is an orthogonal set of nonzero vectors. Then S is linearly independent.

Definition 37 (Orthogonality of Sets)
Let $S_1, S_2 \subseteq \mathbb{C}^n$ be two sets of vectors. Then S_1 and S_2 are *orthogonal to each other* if and only if, for all $\vec{x} \in S_1$ and $\vec{y} \in S_2$, we have that $\vec{x}$ and $\vec{y}$ are orthogonal.

Proposition 38
Let B_1 be a basis for the complex subspace S_1, and B_2 be a basis for the complex subspace S_2. Then S_1 and S_2 are orthogonal to each other if and only if B_1 and B_2 are orthogonal to each other.

Definition 39 (Normalized Vectors)
Let $\vec{x} \in \mathbb{C}^n$. Then $\vec{x}$ is *normalized* (i.e., has *unit norm*) if and only if $\|\vec{x}\| = 1$.

Definition 40 (Orthonormal Vectors)

- Let $\vec{x}, \vec{y} \in \mathbb{C}^n$. Then $\vec{x}$ and $\vec{y}$ are orthonormal if and only if:
  - $\vec{x}$ and $\vec{y}$ are orthogonal; and
  - both $\vec{x}$ and $\vec{y}$ are unit norm.
- Let $S \subseteq \mathbb{C}^n$. Then S is an *orthonormal set* if and only if S is an orthogonal set and all vectors in S have unit norm.

Theorem 41
Let $S \subseteq \mathbb{C}^n$ be a set of vectors. Then S is an orthonormal set if and only if

$$\langle \vec{x}, \vec{y} \rangle = \begin{cases} 1 & \text{if } \vec{x} = \vec{y} \\ 0 & \text{if } \vec{x} \neq \vec{y} \end{cases} \qquad \text{for all } \vec{x}, \vec{y} \in S. \tag{44}$$

3.3.2 Projections

Definition 42 (Projection)
Let $\vec{x} \in \mathbb{C}^n$ and let $S \subseteq \mathbb{C}^n$ be a subspace. Then we define *the projection of $\vec{x}$ onto S* to be the unique point

$$\operatorname{proj}_S(\vec{x}) := \underset{\vec{y} \in S}{\operatorname{argmin}} \|\vec{x} - \vec{y}\|. \tag{45}$$

Theorem 43 (Orthogonality Principle)
Let $S \subseteq \mathbb{R}^n$ be a linear subspace, and $\vec{x} \in \mathbb{R}^n$ be any vector. Then the following are equivalent:

(i) $\vec{y} = \operatorname{proj}_S(\vec{x})$;

(ii) $\vec{y} - \vec{x}$ is orthogonal to S.

Theorem 44 (Projection onto Orthogonal or Orthonormal Basis of Subspace)
Let $S \subseteq \mathbb{C}^n$ be a subspace and $\{\vec{s}_1, \ldots, \vec{s}_\ell\}$ be a basis of orthogonal vectors for this subspace. Then

$$\operatorname{proj}_S(\vec{x}) = \sum_{i=1}^{\ell} \frac{\langle \vec{x}, \vec{s}_i \rangle}{\|\vec{s}_i\|^2} \vec{s}_i = \sum_{i=1}^{\ell} \frac{\langle \vec{x}, \vec{s}_i \rangle}{\langle \vec{s}_i, \vec{s}_i \rangle} \vec{s}_i. \tag{46}$$

If $\{\vec{s}_1, \ldots, \vec{s}_\ell\}$ is an orthonormal set, then

$$\operatorname{proj}_S(\vec{x}) = \sum_{i=1}^{\ell} \langle \vec{x}, \vec{s}_i \rangle \vec{s}_i. \tag{47}$$

3.3.3 Gram-Schmidt Orthonormalization and its Consequences

Algorithm 45 Gram-Schmidt Orthogonalization Algorithm

Input: A set of vectors $\{\vec{a}_1, \ldots, \vec{a}_\ell\} \subseteq \mathbb{C}^n$.

Output: An orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_l\} \subseteq \mathbb{C}^n$ which spans the same set.

```
function GRAMSCHMIDT({a_1, ..., a_ℓ})
    for i = 1, ..., ℓ do
        z_i := a_i − Σ_{k=1}^{i−1} ⟨a_i, p_k⟩ p_k        ▷ If i = 1 then there are no k to sum over, so z_1 = a_1.
        if z_i = 0_n then
            p_i = 0_n
        else
            p_i := z_i / ||z_i||
        end if
    end for
    return {q_1, ..., q_l} := {p_1, ..., p_ℓ} with the ℓ − l zero vectors discarded.
end function
```

Theorem 46 (Correctness of Gram-Schmidt Orthogonalization Algorithm)
Let $\{\vec{a}_1, \ldots, \vec{a}_\ell\} \subseteq \mathbb{C}^n$ be a set of vectors and $\{\vec{q}_1, \ldots, \vec{q}_l\} := \textsc{GramSchmidt}(\{\vec{a}_1, \ldots, \vec{a}_\ell\})$. Then

(i) $\{\vec{q}_1, \ldots, \vec{q}_l\}$ is an orthonormal set of vectors.

(ii) $\text{Span}(\vec{q}_1, \ldots, \vec{q}_l) = \text{Span}(\vec{a}_1, \ldots, \vec{a}_\ell)$.

Further, if $\{\vec{a}_1, \ldots, \vec{a}_\ell\}$ are linearly independent then

(iii) $\ell = l$. (That is, no vectors are discarded.)

(iv) $\text{Span}(\vec{q}_1, \ldots, \vec{q}_i) = \text{Span}(\vec{a}_1, \ldots, \vec{a}_i)$ for all $1 \le i \le \ell$.

Theorem 47 (Existence of Orthonormal Basis)
Let $S \subseteq \mathbb{C}^n$ be a subspace. There exists an orthonormal basis for S.

Theorem 48 (Extending an Orthonormal Basis)
Let $S \subseteq T \subseteq \mathbb{C}^n$ be subspaces. Then for every orthonormal basis B_S of S, there exists an orthonormal basis B_T for T that contains B_S.

4 Structured Matrices

In the real case, we had four types of structure in our matrices we were concerned about: diagonal matrices, upper-triangular matrices, matrices with orthonormal rows or columns, and symmetric matrices. They all have equivalents in the complex domain which are obtained by interchanging transpose $\top$ with conjugate transpose $*$.

4.1 Diagonal Matrices

The definition for diagonal matrix is exactly the same as the real case.

Definition 49 (Diagonal Matrix)
A matrix $A \in \mathbb{C}^{m \times n}$ is *diagonal* if and only if its off-diagonal entries are zero:

$$a_{ij} = 0 \qquad \text{for all } i \neq j, 1 \le i \le m, 1 \le j \le n. \tag{48}$$

Complex diagonal matrices also have their eigenvalues on the diagonal, etc.

Theorem 50 (Eigenvalues of Square Diagonal Matrix)
Let $A \in \mathbb{C}^{n \times n}$ be diagonal. The characteristic polynomial of A is

$$p_A(\lambda) = \prod_{i=1}^{n} (\lambda - a_{ii}) \tag{49}$$

meaning that its eigenvalues are exactly $a_{11}, \ldots, a_{nn}$.

4.2 Upper-Triangular Matrices

Again, the definition for upper triangular matrix is exactly the same as the real case.

Definition 51 (Upper Triangular Matrix)
A matrix $A \in \mathbb{C}^{m \times n}$ is *upper triangular* if and only if its entries below its diagonal are zero:

$$a_{ij} = 0 \qquad \text{for all } i > j, 1 \leq i \leq m, 1 \leq j \leq n. \tag{50}$$

Complex upper triangular matrices also have their eigenvalues on the diagonal, etc.

Theorem 52 (Eigenvalues of Square Upper Triangular Matrix)
Let $A \in \mathbb{C}^{n \times n}$ be upper triangular. The characteristic polynomial of A is

$$p_A(\lambda) = \prod_{i=1}^{n}(\lambda - a_{ii}) \tag{51}$$

meaning that its eigenvalues are exactly $a_{11}, \ldots, a_{nn}$.

4.3 Unitary Matrices

Unitary matrices are the complex analogue of square orthonormal matrices. We can also define the non-square matrices with orthonormal rows and columns similarly to how we defined them in the first place.

Definition 53 (Orthonormal and Unitary Matrices)

- A square matrix $Q \in \mathbb{C}^{n \times n}$ whose columns or rows form an orthonormal set is called a *unitary matrix*.
- A tall matrix $Q \in \mathbb{C}^{m \times n}$ with $m \geq n$ whose columns form an orthonormal set is said to have *orthonormal columns*.
- A wide matrix $Q \in \mathbb{C}^{m \times n}$ with $m \leq n$ whose rows form an orthonormal set is said to have *orthonormal rows*.

Similarly to the real case, the unitary matrices have nice expressions in terms of matrix products.

Theorem 54

(i) If $Q \in \mathbb{C}^{n \times n}$ is a square matrix, then Q is unitary if and only if $Q^* = Q^{-1}$.

(ii) If $Q \in \mathbb{C}^{m \times n}$ with $m \geq n$ is a tall matrix, then Q has orthonormal columns if and only if $Q^*Q = I_n$.

(iii) If $Q \in \mathbb{C}^{m \times n}$ with $m \leq n$ is a wide matrix, then Q has orthonormal rows if and only if $QQ^* = I_m$.

Corollary 55
If $Q \in \mathbb{C}^{n \times n}$ is a square matrix, then it is unitary if and only if it has both orthonormal rows and orthonormal columns.

And, multiplication by unitary matrices does not affect norms.

Theorem 56 (Matrices with Orthonormal Columns are Reflections and Rotations)

(i) Let $Q \in \mathbb{C}^{m \times n}$ with $m \geq n$ have orthonormal columns, and $x \in \mathbb{C}^n$. Then $\|Q\vec{x}\| = \|x\|$.

(ii) Let $Q \in \mathbb{C}^{m \times n}$ with $m \geq n$ have orthonormal columns, and $X \in \mathbb{C}^{n \times p}$. Then $\|QX\|_F = \|X\|_F$.

(iii) Let $Q \in \mathbb{C}^{m \times n}$ with $m \leq n$ have orthonormal rows, and $X \in \mathbb{C}^{p \times m}$. Then $\|XQ\|_F = \|X\|_F$.

4.4 Hermitian Matrices

Hermitian matrices are the complex analogue of symmetric matrices.

Definition 57 (Hermitian Matrix)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. Then A is *Hermitian* if and only if

$$A = A^*. \tag{52}$$

Like symmetric matrices, they obey a spectral theorem.

Theorem 58 (Spectral Theorem for Complex Hermitian Matrices)
Let $A \in \mathbb{C}^{n \times n}$ be Hermitian. Then the following hold:

(i) The eigenvalues of A are real.

(ii) A is diagonalizable.

(iii) There is an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A.

5 Matrix Factorizations

Using unitary and Hermitian matrices in our matrix decompositions, such as diagonalization and SVD, results in versions of these decompositions that can work in the complex domain as well.

5.1 Diagonalization

Definition 59 (Diagonalization)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. A *diagonalization* of A is a decomposition

$$A = V\Lambda V^{-1} \tag{53}$$

where $\Lambda \in \mathbb{C}^{n \times n}$ is a complex diagonal matrix and $V \in \mathbb{C}^{n \times n}$ is a complex invertible matrix.

In this case, we can read off the eigenvectors and eigenvalues of A from its diagonalization.

Theorem 60 (Eigenvalues and Eigenvectors of Diagonalizable Matrix)
Suppose $A = V\Lambda V^{-1}$ is a diagonalization of $A \in \mathbb{R}^{n \times n}$. Then for each $i \in \{1, \ldots, n\}$, we have that $(\lambda_i, \vec{v}_i)$ is an eigenvalue-eigenvector pair of A.

The conditions for diagonalization remain the same as in the real case.

Theorem 61 (Conditions for Diagonalization)
Let $A \in \mathbb{R}^{n \times n}$ be a square complex matrix. The following are equivalent:

1. A is diagonalizable.
2. A has a basis of eigenvectors.
3. The conditions of Theorem 25 item (iii) hold.

Finally, we can write Theorem 58 in terms of a particular diagonalization.

Theorem 62 (Spectral Theorem for Complex Hermitian Matrices as a Diagonalization)
Let $A \in \mathbb{C}^{n \times n}$ be Hermitian. Then A can be written as

$$A = V\Lambda V^{*} \tag{54}$$

where $\Lambda \in \mathbb{R}^{n \times n}$ is a real diagonal matrix of eigenvalues, and $V \in \mathbb{C}^{n \times n}$ is the corresponding unitary matrix of eigenvectors. In short, A may be *unitarily diagonalized*.

5.2 Schur Decomposition

Now we are able to upper-triangularize matrices which even have complex eigenvalues.

Theorem 63 (Existence of Schur Decomposition)
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. Then A can be written as

$$A = UTU^{*} \tag{55}$$

where $U \in \mathbb{C}^{n \times n}$ is unitary and $T \in \mathbb{C}^{n \times n}$ is upper triangular. Furthermore, the eigenvalues of A are on the diagonal of T according to their algebraic multiplicity.

The algorithm to do this is also included here:

Algorithm 64 Schur Decomposition

Input: A square matrix $A \in \mathbb{C}^{n \times n}$.

Output: A unitary matrix $U \in \mathbb{C}^{n \times n}$ and an upper-triangular matrix $T \in \mathbb{C}^{n \times n}$ such that $A = UTU^*$.

function $\textsc{SchurDecomposition}(A)$
    if A is 1×1 **then**
        return $\begin{bmatrix} 1 \end{bmatrix}, A$
    end if
    $(\vec{q}_1, \lambda_1) := \textsc{FindEigenvectorEigenvalue}(A)$
    $Q := \textsc{ExtendBasis}(\{\vec{q}_1\}, \mathbb{C}^n)$ ▷ Extend $\{\vec{q}_1\}$ to a basis of $\mathbb{C}^n$ using Gram-Schmidt
    Unpack $Q := \begin{bmatrix} \vec{q}_1 & \widetilde{Q} \end{bmatrix}$
    Compute and unpack $Q^* A Q = \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^{*} \\ \vec{0}_{n-1} & \widetilde{A}_{22} \end{bmatrix}$
    $(P, \widetilde{T}) := \textsc{SchurDecomposition}(\widetilde{A}_{22})$
    $U := \begin{bmatrix} \vec{q}_1 & \widetilde{Q} P \end{bmatrix}$
    $T := \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^{*} P \\ \vec{0}_{n-1} & \widetilde{T} \end{bmatrix}$
    return (U, T)
end function

5.3 Singular Value Decomposition and Moore-Penrose Pseudoinverse

We first define the complex singular value decomposition, then the pseudoinverse.

Definition 65 (Full SVD)
Let $A \in \mathbb{C}^{m \times n}$ have rank $r \leq \min\{m, n\}$. A *(full) SVD* of A is a decomposition

$$A = U \Sigma V^* = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix} \begin{bmatrix} V_r^* \\ V_{n-r}^* \end{bmatrix} \tag{56}$$

where

(I) $U = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix}$ is a matrix of *left singular vectors*; $U \in \mathbb{C}^{m \times m}$, $U_r \in \mathbb{C}^{m \times r}$, $U_{m-r} \in \mathbb{C}^{m \times (m-r)}$;

(II) $V = \begin{bmatrix} V_r & V_{n-r} \end{bmatrix}$ is a matrix of *right singular vectors*; $V \in \mathbb{C}^{n \times n}$, $V_r \in \mathbb{C}^{n \times r}$, $V_{n-r} \in \mathbb{C}^{n \times (n-r)}$;

(III) $\Sigma = \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$ is a diagonal matrix of *singular values*, where $\Sigma \in \mathbb{R}^{m \times n}$ and $\Sigma_r \in \mathbb{R}^{r \times r}$ is diagonal with real and positive diagonal entries;

such that the following holds:

(i) U is an orthonormal matrix of eigenvectors of AA^*;

(ii) V is an orthonormal matrix of eigenvectors of A^*A;

(iii) $\Sigma\Sigma^\top$ is the matrix of eigenvalues of AA^*;

(iv) $\Sigma^\top \Sigma$ is the matrix of eigenvalues of $A^* A$;

(v) $\mathrm{Col}(U_r) = \mathrm{Col}(A)$;

(vi) $\mathrm{Col}(U_{m-r}) = \mathrm{Null}(A^*)$;

(vii) $\mathrm{Col}(V_r) = \mathrm{Col}(A^*)$;

(viii) $\mathrm{Col}(V_{n-r}) = \mathrm{Null}(A)$.

Theorem 66 (Existence of Full SVD)
Let $A \in \mathbb{C}^{m \times n}$. There exists a full SVD $A = U\Sigma V^*$.

Algorithm 67 Constructing the SVD

```
function FullSVD(A ∈ C^{m×n})
    r := Rank(A)
    (V, Λ) := Diagonalize(A* A)                 ▷ Sorted so that Λ_11 ≥ ··· ≥ Λ_nn
    Unpack V := [V_r  V_{n-r}]
    Unpack Λ := [[Λ_r, 0_{r×(n-r)}], [0_{(n-r)×r}, 0_{(n-r)×(n-r)}]]
    Σ_r := Λ_r^{1/2}
    Pack Σ := [[Σ_r, 0_{r×(n-r)}], [0_{(m-r)×r}, 0_{(m-r)×(n-r)}]]
    U_r := A V_r Σ_r^{-1}
    U := ExtendBasis(U_r, C^m)
    return (U, Σ, V)
end function
```

The compact SVD and outer product form of the SVD are defined, shown to exist, and constructed in similar ways.

Finally, we re-introduce the Moore-Penrose pseudoinverse.

Definition 68 (Moore-Penrose Psuedoinverse)
Suppose $A \in \mathbb{C}^{m \times n}$ has rank $r \leq \min\{m, n\}$. Let $A = U\Sigma V^*$ be an SVD of A. The *Moore-Penrose pseudoinverse* of A is a matrix $A^\dagger \in \mathbb{C}^{n \times m}$ given by

$$A^\dagger := V\Sigma^\dagger U^* \qquad \text{where} \qquad \Sigma^\dagger = \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}^\dagger = \begin{bmatrix} \Sigma_r^{-1} & 0_{r \times (m-r)} \\ 0_{(n-r) \times r} & 0_{(n-r) \times (m-r)} \end{bmatrix}. \tag{57}$$

The compact Moore-Penrose pseudoinverse can be defined similarly as $V_r \Sigma_r^{-1} U_r^*$. It has the following properties:

Proposition 69 (Pseudoinverse Identities)

(i) If A is invertible, i.e., A^{-1} exists, then $A^{\dagger} = A^{-1}$ (inverse is pseudoinverse);

(ii) $(A^{\dagger})^{\dagger} = A$ (taking pseudoinverse twice does nothing);

(iii) $(A^{\top})^{\dagger} = (A^{\dagger})^{\top}$ (pseudoinverse commutes with transpose);

(iv) For $\alpha \in \mathbb{C} \setminus \{0\}$, $(\alpha A)^{\dagger} = \alpha^{-1} A^{\dagger}$ (scalar distributivity);

(v) $AA^{\dagger}A = A$ (weak left inverse property);

(vi) $A^{\dagger}AA^{\dagger} = A^{\dagger}$ (weak right inverse property);

(vii) $AA^{\dagger} = U_r U_r^*$ ($AA^{\dagger}$ is projection onto $\operatorname{Col}(A)$);

(viii) $A^{\dagger}A = V_r V_r^*$ ($A^{\dagger}A$ is projection onto $\operatorname{Col}(A^*)$).

6 Least-Squares and Least-Norm Problems

We use the pseudoinverse to solve least squares and least norm problems, but also give closed-form expressions in terms of A when possible.

First, we have a generalization of the least-norm and least-squares problems.

Theorem 70 (Pseudoinverse Solves Least-Norm Least-Squares)
Let $A \in \mathbb{C}^{m \times n}$ have rank $r \leq \min\{m, n\}$, and let $\vec{b} \in \mathbb{C}^m$. Let S be the set of least squares solutions:

$$S := \operatorname*{argmin}_{\vec{z} \in \mathbb{C}^n} \left\| A\vec{z} - \vec{b} \right\|^2. \tag{58}$$

The solution of the optimization problem

$$\min_{\vec{x} \in S} \|\vec{x}\|^2 \tag{59}$$

is unique and given by $\vec{x}^{\star} = A^{\dagger}\vec{b}$.

From this follows the least-squares and least-norm solutions.

Corollary 71 (Pseudoinverse Solves Least-Squares)
Let $A \in \mathbb{C}^{m \times n}$ with $m \geq n$ have full column rank, and let $\vec{b} \in \mathbb{C}^m$.

(i) The solution to the least-squares problem

$$\min_{\vec{x} \in \mathbb{C}^n} \left\| A\vec{x} - \vec{b} \right\|^2 \tag{60}$$

is given by $\vec{x}^{\star}_{\text{LS}} = A^{\dagger}\vec{b}$.

(ii) The pseudoinverse $A^{\dagger}$ has the formula $A^{\dagger} = (A^*A)^{-1}A^*$.

Corollary 72 (Pseudoinverse Solves Least-Norm)
Let $A \in \mathbb{C}^{m \times n}$ with $m \leq n$ have full row rank, and let $\vec{b} \in \mathbb{C}^m$.

(i) The solution to the least-norm problem

$$\min_{\vec{x} \in \mathbb{C}^n} \quad \|\vec{x}\|^2 \tag{61}$$
$$\text{s.t.} \quad A\vec{x} = \vec{b} \tag{62}$$

is given by $\vec{x}^\star_{\mathrm{LN}} = A^\dagger \vec{b}$.

(ii) The pseudoinverse $A^\dagger$ has the formula $A^\dagger = A^*(AA^*)^{-1}$.

7 Final Comments

That was a lot of material! But as it turns out, many operations over the complex domain are just one-step generalizations of things we already know how to do. Mainly, the results and proofs are the same, just replacing transpose with conjugate transpose, and allowing complex scalars. Hopefully this connection was made clear over the course of this note.

In addition, hopefully this note was a good review of all the major linear algebra we covered in 16A and 16B.

Contributors:

- Druv Pai.

