Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

September 24, 2021 (unfinished!)

Contents

Math 504 Lecture 1

Chris volunteered to scribe Lecture 3.

1. Unitary matrices ([HorJoh13, §2.1])

1.1. Inner product

For any $z \in \mathbb{C}$, we let $\overline{z}$ be the complex conjugate of z. So $\overline{a+bi} = a - bi$ if $a, b \in \mathbb{R}$.

Definition 1.1.1. For any two vectors $u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{C}^n$ and $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n$, we define the scalar

$$\langle u, v \rangle := u_1 \overline{v_1} + u_2 \overline{v_2} + \cdots + u_n \overline{v_n} \in \mathbb{C}.$$

*Drexel University, Korman Center, 15 S 33rd Street, Philadelphia PA, 19104, USA

This scalar $\langle u, v \rangle$ is called the **inner product** (or **dot product**) of u and v.

For example,

$$\left\langle \begin{pmatrix} 1+i \\ 2+3i \end{pmatrix}, \begin{pmatrix} -i \\ 4+i \end{pmatrix} \right\rangle = (1+i)\,\overline{(-i)} + (2+3i)\,\overline{(4+i)}$$
$$= (1+i)\,i + (2+3i)\,(4-i) = \cdots = 10+11i.$$

Definition 1.1.2. For any column vector $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n$, we define the row vector

$$v^* := \begin{pmatrix} \overline{v_1} & \overline{v_2} & \cdots & \overline{v_n} \end{pmatrix}.$$

Proposition 1.1.3. Let $u, v \in \mathbb{C}^n$. Then:
(a) We have $\langle u, v \rangle = \overline{\langle v, u \rangle}$.
(b) We have $\langle u, v \rangle = v^* u$.
(c) We have $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$ for any $u' \in \mathbb{C}^n$.
(d) We have $\langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle$ for any $v' \in \mathbb{C}^n$.
(e) We have $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$ for any $\lambda \in \mathbb{C}$.
(f) We have $\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle$ for any $\lambda \in \mathbb{C}$.

Proposition 1.1.4. Let $x \in \mathbb{C}^n$. Then:
(a) The number $\langle x, x \rangle$ is a nonnegative real.
(b) If x is nonzero, then $\langle x, x \rangle$ is a positive real.

Proof.

$$\langle x, x \rangle = x_1\overline{x_1} + x_2\overline{x_2} + \cdots + x_n\overline{x_n} = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2,$$

since $z\overline{z} = |z|^2$ for any $z \in \mathbb{C}$. Note that $|x_1|, |x_2|, \ldots, |x_n|$ are reals. Thus, $|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$ is clearly a nonnegative real. Furthermore, if x is nonzero, then at least one x_i is nonzero, and therefore $|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$ is positive. $\square$

Definition 1.1.5. Let $x \in \mathbb{C}^n$. We define the **length** (aka **norm**) of x to be the nonnegative real number

$$||x|| := \sqrt{\langle x, x \rangle}.$$

For example, if $x = (1, 1)$, then $\langle x, x \rangle = 1\overline{1} + 1\overline{1} = 2$, so $||x|| = \sqrt{2}$.

Proposition 1.1.6. For any $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n$, we have $||\lambda x|| = |\lambda| \cdot ||x||$.

1.2. Orthogonality and orthonormality

Definition 1.2.1. Let $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ be two vectors. We say that x is **orthogonal** to y if and only if $\langle x, y \rangle = 0$. The shorthand for this is "$x \perp y$". (\perp)

The relation $\perp$ is symmetric:

Proposition 1.2.2. Let $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ be two vectors. Then, $x \perp y$ if and only if $y \perp x$.

Proof. Recall that $\langle u, v \rangle = \overline{\langle v, u \rangle}$. $\square$

Definition 1.2.3. Let $(u_1, u_2, \ldots, u_k)$ be a tuple of vectors in $\mathbb{C}^n$. Then:

(a) We say that the tuple $(u_1, u_2, \ldots, u_k)$ is **orthogonal** if we have

$$u_p \perp u_q \qquad \text{for all } p \neq q.$$

(b) We say that the tuple $(u_1, u_2, \ldots, u_k)$ is **orthonormal** if it is orthogonal and

$$||u_1|| = ||u_2|| = \cdots = ||u_k|| = 1.$$

Example 1.2.4. **(a)** The tuple

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

is orthonormal.

(b) More generally: Let $n \in \mathbb{N}$. Let $e_1, e_2, \ldots, e_n \in \mathbb{C}^n$ be the vectors defined by

$$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{with the 1 being in the } i\text{-th position.}$$

Then, the tuple $(e_1, e_2, \ldots, e_n)$ is orthonormal. It is furthermore a basis of $\mathbb{C}^n$, and is known as the **standard basis**.

(c) The pair $\left(\begin{pmatrix} 1 \\ -i \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2i \\ 1 \end{pmatrix} \right)$ of vectors in $\mathbb{C}^3$ is orthogonal, but not orthonormal.

(d) The pair $\left(\frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -i \\ 2 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2i \\ 1 \end{pmatrix} \right)$ of vectors in $\mathbb{C}^3$ is orthonormal.

Proposition 1.2.5. Let $(u_1, u_2, \ldots, u_k)$ be an orthogonal tuple of nonzero vectors in $\mathbb{C}^n$. Then, the tuple

$$\left(\frac{1}{||u_1||} u_1, \ \frac{1}{||u_2||} u_2, \ \ldots, \ \frac{1}{||u_k||} u_k \right)$$

is orthonormal.

Proposition 1.2.6. Any orthogonal tuple of nonzero vectors in $\mathbb{C}^n$ is linearly independent.

Proof. See notes. $\square$

Lemma 1.2.7. Let $k < n$. Let $a_1, a_2, \ldots, a_k$ be k vectors in $\mathbb{C}^n$. Then, there exists a nonzero vector b that is orthogonal to each of $a_1, a_2, \ldots, a_k$.

Proof. (See notes for details.) Write each vector a_i as $a_i = \begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,n} \end{pmatrix}$. Let $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{C}^n$ be a vector whose entries are so far undetermined. To ensure that b is orthogonal to a_i, we need $\langle b, a_i \rangle = 0$. In other words, we need

$$b_1 \overline{a_{i,1}} + b_2 \overline{a_{i,2}} + \cdots + b_n \overline{a_{i,n}} = 0.$$

This has to hold for each $i \in [k] := \{1, 2, \ldots, k\}$. This is a system of k homogeneous linear equations in the n unknowns $b_1, b_2, \ldots, b_n$. Since there are fewer equations than there are unknowns, there exists a nonzero solution. IOW, there exists a nonzero vector b orthogonal to all of a_i. $\square$

1.3. Conjugate transposes

Generalizing our notation v^*, we can define A^* for any matrix A:

Definition 1.3.1. Let $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}$ be any matrix. Then, we define the matrix

$$A^* = \begin{pmatrix} \overline{a_{1,1}} & \cdots & \overline{a_{n,1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1,m}} & \cdots & \overline{a_{n,m}} \end{pmatrix}.$$

This matrix A^* is called the **conjugate transpose** of A.

IOW, A^* is obtained from A by transposing the matrix and conjugating all entries.

Example 1.3.2.

$$\begin{pmatrix} 1+i & 2-3i & i \\ 6 & 0 & 10+i \end{pmatrix}^* = \begin{pmatrix} 1-i & 6 \\ 2+3i & 0 \\ -i & 10-i \end{pmatrix}.$$

Proposition 1.3.3. **(a)** If $A, B \in \mathbb{C}^{n \times m}$ are two matrices, then $(A+B)^* = A^* + B^*$.
(b) If A is a matrix and $\lambda \in \mathbb{C}$, then $(\lambda A)^* = \overline{\lambda} A^*$.
(c) If A and B are two matrices that can be multiplied, then $(AB)^* = B^* A^*$.
(d) If A is any matrix, then $(A^*)^* = A$.

1.4. Isometries

Definition 1.4.1. An $n \times k$-matrix A is said to be an **isometry** if $A^* A = I_k$. (The notation I_k means the $k \times k$ identity matrix.)

Proposition 1.4.2. An $n \times k$-matrix A is an isometry if and only if its columns form an orthonormal tuple of vectors.

Proof. Let A be an $n \times k$-matrix, and let $a_1, a_2, \ldots, a_k$ be its k columns. Thus,

$$A = \begin{pmatrix} | & & | \\ a_1 & \cdots & a_k \\ | & & | \end{pmatrix} \qquad \text{and therefore} \qquad A^* = \begin{pmatrix} - & a_1^* & - \\ & \vdots & \\ - & a_k^* & - \end{pmatrix}.$$

Hence,

$$A^*A = \begin{pmatrix} - & a_1^* & - \\ & \vdots & \\ - & a_k^* & - \end{pmatrix} \begin{pmatrix} | & & | \\ a_1 & \cdots & a_k \\ | & & | \end{pmatrix}$$
$$= \begin{pmatrix} a_1^*a_1 & \cdots & a_1^*a_k \\ \vdots & \ddots & \vdots \\ a_k^*a_1 & \cdots & a_k^*a_k \end{pmatrix} = \begin{pmatrix} ||a_1||^2 & \cdots & \langle a_1, a_k \rangle \\ \vdots & \ddots & \vdots \\ \langle a_k, a_1 \rangle & \cdots & ||a_k||^2 \end{pmatrix}.$$

On the other hand,

$$I_k = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}.$$

Thus, $A^*A = I_k$ holds if and only if we have

$$\langle a_p, a_q \rangle = 0 \qquad \text{for all } p \neq q$$

and

$$||a_p||^2 = 1 \qquad \text{for each } p.$$

IOW, it holds if and only if we have

$$a_p \perp a_q \qquad \text{for all } p \neq q$$

and

$$||a_1|| = ||a_2|| = \cdots = ||a_k|| = 1.$$

IOW, it holds if and only if $(a_1, a_2, \ldots, a_k)$ is orthonormal. Qed. □

Isometries are called isometries because they preserve lengths:

Proposition 1.4.3. Let $A \in \mathbb{C}^{n \times k}$ be an isometry. Then, each $x \in \mathbb{C}^k$ satisfies $||Ax|| = ||x||$.

Proof. We have $A^*A = I_k$ (since A is an isometry). Let $x \in \mathbb{C}^k$. By definition, $||Ax|| = \sqrt{\langle Ax, Ax \rangle}$, so that

$$||Ax||^2 = \langle Ax, Ax \rangle = \underbrace{(Ax)^*}_{=x^*A^*} Ax \qquad (\text{since } \langle u, v \rangle = v^*u)$$
$$= x^* \underbrace{A^*A}_{=I_k} x = x^*x = \langle x, x \rangle = ||x||^2.$$

Hence, $||Ax|| = ||x||$. □

1.5. Unitary matrices

Definition 1.5.1. A matrix $U \in \mathbb{C}^{n \times k}$ is said to be **unitary** if both U and U^* are isometries.

Example 1.5.2. **(a)** The matrix $A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is unitary. Indeed, it is easy to see that A is an isometry, but A^* is therefore also an isometry, since $A^* = A$. Thus, A is unitary.

(b) A 1×1-matrix $\begin{pmatrix} \lambda \end{pmatrix}$ is unitary if and only if $|\lambda| = 1$.

(c) For any $n \in \mathbb{N}$, the identity matrix I_n is an isometry and thus unitary.