Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

October 17, 2021 (unfinished!)

Contents

Math 504 Lecture 9

1. Jordan canonical (aka normal) form (cont'd)

1.1. Step 3: Strictly upper-triangular matrices

Last time, we have proved the uniqueness of the JCF (Jordan canonical form) of a matrix $A \in \mathbb{C}^{n \times n}$.

We also started proving its existence. The first 2 steps we made were:

- We brought A to an upper-triangular form with diagonal entries in contiguous blocks. (This was just careful Schur triangularization.)
- We cleaned out the space "between" distinct diagonal entries: $\begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 2 \end{pmatrix} \to \begin{pmatrix} 1 & * & \\ & 1 & \\ & & 2 \end{pmatrix}$.

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

Thus, after these steps, our matrix has become a block-diagonal matrix $\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$, where each A_i is an upper-triangular matrix with all its diagonal entries equal.

Now it remains to decompose each A_i into Jordan cells. To be precise, we want to show that each A_i is similar to a Jordan matrix. (Indeed, this will easily cause the total matrix to be similar to a Jordan matrix.)

For each $i \in [k]$, the matrix A_i has all its diagonal entries equal. Let's say they all equal μ_i. Thus, $A_i - \mu_i I$ is a strictly upper-triangular matrix. (Recall: a strictly **upper-triangular** matrix is an upper-triangular matrix whose diagonal entries are 0.)

So we need to find a way to decompose a strictly upper-triangular matrix into Jordan cells.

Call the upper-triangular matrix A instead of $A_i - \mu_i I$. Forget about the big matrix.

So we have a strictly upper-triangular matrix $A \in \mathbb{C}^{n \times n}$. We want to prove that A is similar to a Jordan matrix.

We begin by trying some examples:

Example 1.1.1. Let $n = 2$. Then, $A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ for some $a \in \mathbb{C}$.

We are looking for an invertible matrix $S \in \mathbb{C}^{2\times 2}$ such that $S^{-1}AS$ is a Jordan matrix.

If $a = 0$, then this is obvious (just take $S = I_2$), since $A = \begin{pmatrix} J_1(0) & \\ & J_1(0) \end{pmatrix}$ is already a Jordan matrix.

Now assume $a \neq 0$.

Consider our unknown invertible matrix S. Let s_1 and s_2 be its columns. Then, s_1 and s_2 are linearly independent (since S is invertible). Moreover, we want $S^{-1}AS = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. In other words, we want $AS = S \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. However, $S = \begin{pmatrix} s_1 & s_2 \end{pmatrix}$ (in block-matrix notation), so $S \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & s_1 \end{pmatrix}$. Thus our equation $AS = S \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is equivalent to

$$\begin{pmatrix} As_1 & As_2 \end{pmatrix} = \begin{pmatrix} 0 & s_1 \end{pmatrix}.$$

In other words, $As_1 = 0$ and $As_2 = s_1$.

So we are looking for two linearly independent vectors $s_1, s_2 \in \mathbb{C}^2$ such that $As_1 = 0$ and $As_2 = s_1$.

One way to do so is to pick some nonzero vector $s_1 \in \operatorname{Ker} A$, and then define s_2 to be some preimage of s_1 under A. (It can be shown that such preimage exists.) This way, however, does not generalize to higher n.

Another (better) way is to start by picking $s_2 \in \mathbb{C}^2 \setminus \operatorname{Ker} A$ and then setting $s_1 = As_2$. We claim that s_1 and s_2 are linearly independent, and that $As_1 = 0$.

To show that $As_1 = 0$, we just observe that $As_1 = \underbrace{AA}_{=A^2=0} s_2 = 0$.

To show that s_1 and s_2 are linearly independent, we argue as follows: Let $\lambda_1, \lambda_2 \in \mathbb{C}$ be such that $\lambda_1 s_1 + \lambda_2 s_2 = 0$. Applying A to this, we obtain $A \cdot (\lambda_1 s_1 + \lambda_2 s_2) = A \cdot 0 = 0$. However,

$$A \cdot (\lambda_1 s_1 + \lambda_2 s_2) = \lambda_1 \underbrace{As_1}_{=0} + \lambda_2 \underbrace{As_2}_{=s_1} = \lambda_2 s_1,$$

so this becomes $\lambda_2 s_1 = 0$. However, $s_1 \neq 0$ (because $s_1 = As_2$ but $s_2 \notin \operatorname{Ker} A$). Hence, $\lambda_2 = 0$. Now, $\lambda_1 s_1 + \lambda_2 s_2 = 0$ becomes $\lambda_1 s_1 = 0$. Since $s_1 \neq 0$, this yields $\lambda_1 = 0$. Now both λ_is are 0, qed.

Example 1.1.2. Let $n = 3$ and $A = \begin{pmatrix} & 1 & 1 \\ & & 0 \\ & & \end{pmatrix}$.

Our first method above doesn't work, because most vectors in $\operatorname{Ker} A$ do not have preimages under A.

However, our second method can be made to work:

We pick a vector $s_3 \notin \operatorname{Ker} A$. To wit, we pick $s_3 = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Then, $As_3 = e_1$.

Set $s_2 = As_3 = e_1$. Note that $s_2 \in \operatorname{Ker} A$. Let s_1 be another nonzero vector in $\operatorname{Ker} A$, namely $e_2 - e_3$. These three vectors s_1, s_2, s_3 are linearly independent and satisfy $As_1 = 0$ and $As_2 = 0$ and $As_3 = s_2$.

So $S = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. And indeed, $S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ is a Jordan matrix.

So what is the general algorithm here? Can we always find n linearly independent vectors $s_1, s_2, \ldots, s_n$ such that each As_i is either 0 or s_{i-1} ?

To approach this question, we recall a theorem from basic linear algebra:

Theorem 1.1.3. Let V be a finite-dimensional vector space. Let $(v_1, v_2, \ldots, v_k)$ be any linearly independent tuple of vectors in V. Then, this tuple can be extended to a basis of V. In other words, we can define further vectors $v_{k+1}, v_{k+2}, \ldots, v_m$ (where $m = \dim V$) such that $(v_1, v_2, \ldots, v_m)$ is a basis of V.

Now, we are considering a strictly upper-triangular matrix $A \in \mathbb{C}^{n \times n}$.

The "entries ballooning upwards" argument shows that $A^n = 0$. Thus,

$$0 = \operatorname{Ker}\left(A^0\right) \subseteq \operatorname{Ker}\left(A^1\right) \subseteq \operatorname{Ker}\left(A^2\right) \subseteq \cdots \subseteq \operatorname{Ker}\left(A^{n-1}\right) \subseteq \operatorname{Ker}\left(A^n\right) = \mathbb{C}^n.$$

This is a chain of subspaces of $\mathbb{C}^n$ (although some of the $\subseteq$ inclusions can be equalities).

Now, we begin by picking a basis $(v_1, v_2, v_3, \ldots)$ of $\operatorname{Ker}\left(A^{n-1}\right)$ (this list is actually finite; we just don't want to give a name to its last entry), and extending it to a basis $(v_1, v_2, v_3, \ldots, s_{n,1}, s_{n,2}, s_{n,3}, \ldots)$ of $\operatorname{Ker}(A^n)$ (by the theorem above, since it is a linearly independent list of vectors in $\operatorname{Ker}(A^n)$). Now we throw away the $v_1, v_2, v_3, \ldots$ and only keep the $s_{n,1}, s_{n,2}, s_{n,3}, \ldots$.

Then, $As_{n,1}, As_{n,2}, As_{n,3}, \ldots$ belong to $\operatorname{Ker}\left(A^{n-1}\right)$ (indeed, more generally, if $w \in \operatorname{Ker}\left(A^k\right)$, then $Aw \in \operatorname{Ker}\left(A^{k-1}\right)$), and are linearly independent (to be proved later). Extend the list $(As_{n,1}, As_{n,2}, As_{n,3}, \ldots)$ to

==> TO BE CONTINUED NEXT WEDNESDAY

Now to something completely different...

1.2. The Cauchy–Schwarz inequality

Recall the inner product $\langle u, v\rangle = v^* u = u_1 \overline{v_1} + u_2 \overline{v_2} + \cdots + u_n \overline{v_n}$ of two vectors $u, v \in \mathbb{C}^n$.

Theorem 1.2.1 (Cauchy–Schwarz inequality)**.** Let $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ be two vectors. Then:

(a) The inequality

$$||x|| \cdot ||y|| \geq |\langle x, y\rangle| \qquad \text{holds.}$$

(b) This inequality becomes an equality if and only if the pair (x, y) is linearly dependent.

Proof. If $x = 0$, then this is obvious. So WLOG assume that $x \neq 0$. Thus, $\langle x, x\rangle > 0$.

Let

$$a = \langle x, x\rangle = ||x||^2 \in \mathbb{R} \qquad \text{and} \qquad b = \langle y, x\rangle \in \mathbb{C}.$$

Now, recall that $\langle u, u\rangle \geq 0$ for any $u \in \mathbb{C}^n$. Thus,

$$\langle bx - ay,\ bx - ay\rangle \geq 0.$$

Since

$$\begin{aligned}
\langle bx - ay,\ bx - ay \rangle &= \langle bx - ay,\ bx \rangle - \langle bx - ay,\ ay \rangle \\
&= \langle bx,\ bx \rangle - \langle ay,\ bx \rangle - \langle bx,\ ay \rangle + \langle ay,\ ay \rangle \\
&= b\overline{b} \langle x, x \rangle - a\overline{b} \langle y, x \rangle - b\overline{a} \langle x, y \rangle + a\overline{a} \langle y, y \rangle \\
&= b\overline{b} \underbrace{\langle x, x \rangle}_{=a} - a\overline{b} \underbrace{\langle y, x \rangle}_{=b} - ba \underbrace{\langle x, y \rangle}_{\substack{=\overline{b} \\ (\text{since } \langle x,y \rangle = \overline{\langle y,x \rangle})}} + aa \langle y, y \rangle \\
&\qquad (\text{since } a \in \mathbb{R} \text{ and thus } \overline{a} = a) \\
&= b\overline{b}a - a\overline{b}b - ba\overline{b} + aa \langle y, y \rangle = a \left(a \langle y, y \rangle - b\overline{b} \right),
\end{aligned}$$

this rewrites as

$$a \left(a \langle y, y \rangle - b\overline{b} \right) \geq 0.$$

We can divide by a (since $a = ||x||^2 > 0$), and obtain $a \langle y, y \rangle - b\overline{b} \geq 0$. In other words,

$$a \langle y, y \rangle \geq b\overline{b} = |b|^2 = |\langle x, y \rangle|^2 \qquad \left(\text{since } b = \langle y, x \rangle = \overline{\langle x, y \rangle} \right).$$

Since $a = \langle x, x \rangle = ||x||^2$ and $\langle y, y \rangle = ||y||^2$, this rewrites as

$$||x||^2 \cdot ||y||^2 \geq |\langle x, y \rangle|^2 .$$

Now take square roots and be done with **(a)**.

(b) Take a look at the above proof of **(a)** and think about it. See notes for details (§1.1). □

Corollary 1.2.2 (triangle inequality)**.** For any $x, y \in \mathbb{C}^n$, we have $||x|| + ||y|| \geq ||x + y||$.

Proof. Exercise 1.1.2. □