Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

October 13, 2021 (unfinished!)

Contents

Math 504 Lecture 10

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

1.1. Step 3: Strictly upper-triangular matrices redux

Let us fill in what we couldn't back in Lecture 9.

Consider a strictly upper-triangular $n \times n$-matrix A (that is, an upper-triangular matrix whose diagonal entries are 0 as well).

We want to find a basis $(s_1, s_2, \ldots, s_n)$ of $\mathbb{C}^n$ such that for each $i \in [n]$, the vector As_i is either s_{i-1} or 0. (When $i = 1$, this vector has to be 0, since there is no s_0.) In fact, if $(s_1, s_2, \ldots, s_n)$ is such a basis, then the matrix $S := \begin{pmatrix} s_1 & s_2 & \cdots & s_n \end{pmatrix} \in \mathbb{C}^{n \times n}$ is invertible and satisfies

$$\begin{aligned} AS &= (\text{a matrix whose } i\text{-th column is either } s_{i-1} \text{ or } 0 \text{ for each } i) \\ &= S \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (\text{for example}), \end{aligned}$$

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

so that

$$S^{-1}AS = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (\text{a Jordan matrix}) .$$

So A is similar to a Jordan matrix.

How do we find such a basis $(s_1, s_2, \ldots, s_n)$?

(The following proof is due to Terence Tao.)

We define an **orbit** to be a tuple of the form $\left(v, Av, A^2v, \ldots, A^kv\right)$, where $v \in \mathbb{C}^n$ satisfies $A^{k+1}v = 0$. Note that for each $v \in \mathbb{C}^n$, there is an orbit that starts with v, since $A^n = 0$.

The **concatenation** of some tuples $(a_1, a_2, \ldots, a_k)$ and $(b_1, b_2, \ldots, b_\ell)$ and $(c_1, c_2, \ldots, c_m)$ is $(a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_\ell, c_1, c_2, \ldots, c_m)$.

Now, I claim:

Lemma 1.1.1 (orbit basis lemma)**.** There exists a basis of $\mathbb{C}^n$ that is a concatenation of orbits.

Once this lemma is proved, we will be done, because reading such a basis backwards gives us exactly the basis $(s_1, s_2, \ldots, s_n)$ we are looking for. For example, if our basis that is a concatenation of orbits is

$$\left(u, Au, A^2u, \quad v, Av, A^2v, A^3v, \quad w, Aw\right)$$

(with $A^3u = 0$ and $A^4v = 0$ and $A^2w = 0$), then reading it backwards gives

$$\left(Aw, w, \quad A^3v, A^2v, Av, v, \quad A^2u, Au, u\right),$$

which is a basis $(s_1, s_2, \ldots, s_n)$ of $\mathbb{C}^n$ such that for each $i \in [n]$, the vector As_i is either s_{i-1} or 0.

Proof of the Lemma. It is easy to find a finite **spanning set** of $\mathbb{C}^n$ that is a concatenation of orbits. Indeed, we can start with the standard basis $(e_1, e_2, \ldots, e_n)$, and extend it to the list

$$\begin{aligned}
&(e_1, Ae_1, A^2e_1, \ldots, A^{n-1}e_1, \\
&e_2, Ae_2, A^2e_2, \ldots, A^{n-1}e_2, \\
&\ldots, \\
&e_n, Ae_n, A^2e_n, \ldots, A^{n-1}e_n).
\end{aligned}$$

This is clearly a spanning set of $\mathbb{C}^n$ (since $e_1, e_2, \ldots, e_n$ already span $\mathbb{C}^n$), and also a concatenation of orbits (since $A^n = 0$).

Now, we will gradually shorten this spanning set (i.e., replace it by smaller ones) until we get a basis. We have to do this in such a way that it remains a spanning set throughout the process, and that it remains a concatenation of orbits throughout the process.

For the sake of concreteness, let us assume that our spanning set is

$$\left(x, Ax, \ \ y, Ay, A^2y, \ \ z, Az, A^2z, A^3z, \ \ w\right),$$

with $A^2x = 0$ and $A^3y = 0$ and $A^4z = 0$ and $Aw = 0$. If this spanning set is linearly independent, then it is already a basis, and we are done. So assume that it isn't. Thus, there exists some linear dependence relation – say,

$$3x + 4Ax + 5Ay + 6A^2y + 7A^2z + 8w = 0.$$

Apply A to this relation:

$$\begin{aligned} 3Ax + 4A^2x + 5A^2y + 6A^3y + 7A^3z + 8Aw = 0, \qquad & \text{i.e.} \\ 3Ax + 5A^2y + 7A^3z = 0. \end{aligned}$$

Apply A to this relation:

$$\begin{aligned} 3A^2x + 5A^3y + 7A^4z = 0, \qquad & \text{i.e.} \\ 0 = 0. \end{aligned}$$

We have gone too far, so let us revert to the previous equation:

$$3Ax + 5A^2y + 7A^3z = 0.$$

So this is a linear dependence relation between the **final** vectors of the orbits in our spanning set. ("Final" means the last vector in the orbit.) Factoring out an A in this relation, we obtain

$$A\left(3x + 5Ay + 7A^2z\right) = 0.$$

So the 1-tuple $\left(3x + 5Ay + 7A^2z\right)$ is an orbit.

Now, let us replace the orbit (x, Ax) in our spanning set $\left(x, Ax, \ \ y, Ay, A^2y, \ \ z, Az, A^2z, A^3z, \ \ w\right)$ by the orbit $\left(3x + 5Ay + 7A^2z\right)$. We get

$$\left(3x + 5Ay + 7A^2z, \ \ y, Ay, A^2y, \ \ z, Az, A^2z, A^3z, \ \ w\right).$$

This is still a concatenation of orbits, since the 1-tuple $\left(3x + 5Ay + 7A^2z\right)$ is an orbit. Furthermore, this is still a spanning set of $\mathbb{C}^n$; why? Because we removed the dependent vector Ax (this is a combination of the other vectors, because $3Ax + 5A^2y + 7A^3z = 0$) and we replaced x by $3x + 5Ay + 7A^2z$ (which does not change the span, because Ay and A^2z are still in the spanning set).

This example generalizes. In the general case, you have a spanning set that is a concatenation of orbits:

$$\left(v_1, Av_1, \ldots, A^{m_1}v_1, \ v_2, Av_2, \ldots, A^{m_2}v_2, \ \ldots, \ v_k, Av_k, \ldots, A^{m_k}v_k\right).$$

If it is a basis, you are done. If not, you pick a linear dependence relation:

$$\sum_{i,j} \lambda_{i,j} A^j v_i = 0.$$

By multiplying this by A an appropriate amount of times (namely, you keep multiplying until it becomes $0 = 0$, and then you take a step back), you obtain a linear dependence relation that involves only the **final** vectors of the orbits (i.e., the vectors $A^{m_1}v_1,\ A^{m_2}v_2,\ \ldots,\ A^{m_k}v_k$). So it will look like this:

$$\mu_1 A^{m_1} v_1 + \mu_2 A^{m_2} v_2 + \cdots + \mu_k A^{m_k} v_k = 0.$$

Assume WLOG that the first p of the $\mu_1, \mu_2, \ldots, \mu_k$ are nonzero, while the remaining $k - p$ are 0. So the relation becomes

$$\mu_1 A^{m_1} v_1 + \mu_2 A^{m_2} v_2 + \cdots + \mu_p A^{m_p} v_p = 0,$$

with $\mu_1, \mu_2, \ldots, \mu_p$ being nonzero. Assume WLOG that $m_1 = \min \{m_1, m_2, \ldots, m_p\}$, and factor out A^{m_1} from this relation. This yields

$$A^{m_1} \left(\mu_1 v_1 + \mu_2 A^{m_2 - m_1} v_2 + \cdots + \mu_p A^{m_p - m_1} v_p \right) = 0.$$

Now, set $w_1 = \mu_1 v_1 + \mu_2 A^{m_2 - m_1} v_2 + \cdots + \mu_p A^{m_p - m_1} v_p$. Thus, $A^{m_1} w_1 = 0$. Hence, $\left(w_1, Aw_1, A^2 w_1, \ldots, A^{m_1 - 1} w_1\right)$ is an orbit of length m_1. Now, replace the orbit $(v_1, Av_1, \ldots, A^{m_1} v_1)$ in the spanning set

$$(v_1, Av_1, \ldots, A^{m_1} v_1,\ v_2, Av_2, \ldots, A^{m_2} v_2,\ \ldots,\ v_k, Av_k, \ldots, A^{m_k} v_k)$$

by the shorter orbit $\left(w_1, Aw_1, A^2 w_1, \ldots, A^{m_1 - 1} w_1\right)$. The resulting list

$$\left(w_1, Aw_1, A^2 w_1, \ldots, A^{m_1 - 1} w_1,\ v_2, Av_2, \ldots, A^{m_2} v_2,\ \ldots,\ v_k, Av_k, \ldots, A^{m_k} v_k\right)$$

is still a concatenation of orbits. Also, it still spans $\mathbb{C}^n$, because

$$\begin{aligned}
w_1 &= \underbrace{\mu_1}_{\neq 0} v_1 + \mu_2 A^{m_2 - m_1} v_2 + \cdots + \mu_p A^{m_p - m_1} v_p; \\
Aw_1 &= \underbrace{\mu_1}_{\neq 0} Av_1 + \mu_2 A^{m_2 - m_1 + 1} v_2 + \cdots + \mu_p A^{m_p - m_1 + 1} v_p; \\
&\ldots; \\
A^{m_1} v_1 &= -\left(\mu_2 A^{m_2} v_2 + \cdots + \mu_p A^{m_p} v_p\right) \\
&\qquad \left(\text{since } \mu_1 A^{m_1} v_1 + \mu_2 A^{m_2} v_2 + \cdots + \mu_p A^{m_p} v_p = 0\right).
\end{aligned}$$

So we have found a spanning set of $\mathbb{C}^n$ that is still a concatenation of orbits, but is shorter (it has one less vector). Doing this repeatedly, we will eventually obtain a basis (since we cannot keep making a finite list shorter and shorter indefinitely). This proves the lemma. □

As we said, the lemma gives us a basis $(s_1, s_2, \ldots, s_n)$ such that As_i is either s_{i-1} or 0; and that shows that A is similar to a Jordan matrix. This completes the proof of the existence part of the Jordan canonical form.

Example 1.1.2. Let

$$A = \begin{pmatrix} 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 1 & -2 & 2 & -2 \\ 0 & 1 & 0 & -1 & 2 & -2 \\ 0 & 1 & 0 & -1 & 2 & -2 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & -1 & 1 & -1 \end{pmatrix}.$$

This is not strictly upper-triangular, but it is nilpotent, with $A^3 = 0$, so the above argument goes equally well with this A.

Let us try to find a basis of $\mathbb{C}^6$ that is a concatenation of orbits.

We begin with the spanning set

$$\left(e_1, Ae_1, A^2e_1,\ \ e_2, Ae_2, A^2e_2,\ \ \ldots,\ \ e_6, Ae_6, A^2e_6\right).$$

It has lots of linear dependencies. For one, $Ae_1 = 0$. Multiplying it by A gives $A^2e_1 = 0$, so we can replace $\left(e_1, Ae_1, A^2e_1\right)$ by (e_1, Ae_1). So our spanning set becomes

$$\left(e_1, Ae_1,\ \ e_2, Ae_2, A^2e_2,\ \ \ldots,\ \ e_6, Ae_6, A^2e_6\right).$$

One more step of the same form gives

$$\left(e_1,\ \ e_2, Ae_2, A^2e_2,\ \ \ldots,\ \ e_6, Ae_6, A^2e_6\right).$$

Now, observe that $Ae_3 = e_2$. That is, $e_2 - Ae_3 = 0$. Multiplying it by A^2, we obtain $A^2e_2 = 0$ (since $A^2 \cdot Ae_3 = A^3e_3 = 0$). So we replace the orbit $\left(e_2, Ae_2, A^2e_2\right)$ by (e_2, Ae_2). So we get the spanning set

$$\left(e_1,\ \ e_2, Ae_2,\ \ e_3, Ae_3, A^2e_3,\ \ e_4, Ae_4, A^2e_4,\ \ e_5, Ae_5, A^2e_5,\ \ e_6, Ae_6, A^2e_6\right).$$

We observe that

$$Ae_2 = e_1 + e_2 + e_3 + e_4 + e_5 + e_6.$$

In other words,

$$Ae_2 - e_1 - e_2 - e_3 - e_4 - e_5 - e_6 = 0.$$

Multiplying this by A^2, we obtain

$$-A^2e_3 - A^2e_4 - A^2e_5 - A^2e_6 = 0.$$

In other words,

$$A^2\left(-e_3 - e_4 - e_5 - e_6\right) = 0.$$

Thus, we set $w_1 := -e_3 - e_4 - e_5 - e_6$, and we replace (e_3, Ae_3, A^2e_3) by (w_1, Aw_1). So we get the spanning set

$$\left(e_1, \quad e_2, Ae_2, \quad w_1, Aw_1, \quad e_4, Ae_4, A^2e_4, \quad e_5, Ae_5, A^2e_5, \quad e_6, Ae_6, A^2e_6\right).$$

Keep making these steps. Eventually, there will be no more linear dependencies, so we will have a basis.