

# Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg\*

October 4, 2021 (unfinished!)

## Contents

1. Schur triangularization (cont'd)	1
1.1. Sylvester's equation (cont'd)	1
2. Jordan canonical (aka normal) form	3
2.1. Similarity redux	3
2.2. Jordan cells	4
2.3. The Jordan canonical form (aka Jordan normal form)	5

## Math 504 Lecture 7

### 1. Schur triangularization (cont'd)

#### 1.1. Sylvester's equation (cont'd)

Last time, we proved the  $\mathcal{V} \implies \mathcal{U}$  direction of the following fact:

**Theorem 1.1.1.** Let  $A$  be an  $n \times n$ -matrix, and let  $B$  be an  $m \times m$ -matrix (both with complex entries). Let  $C$  be an  $n \times m$ -matrix. Then, the following statements are equivalent:

- $\mathcal{U}$ : There is a **unique** matrix  $X \in \mathbb{C}^{n \times m}$  such that  $AX - XB = C$ .
- $\mathcal{V}$ : We have  $\sigma(A) \cap \sigma(B) = \emptyset$ .

---

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

Here,  $\sigma(A)$  denotes the **spectrum** of a matrix  $A$  (that is, the set of all eigenvalues of  $A$ ).

Today, we shall derive a corollary from the above theorem:

**Corollary 1.1.2.** Let  $A \in \mathbb{C}^{n \times n}$ ,  $B \in \mathbb{C}^{m \times m}$  and  $C \in \mathbb{C}^{n \times m}$  be three matrices such that  $\sigma(A) \cap \sigma(B) = \emptyset$ . Then, the two  $(n+m) \times (n+m)$ -matrices

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

(written in block-matrix notation) are similar.

**Example 1.1.3.** Let  $A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$  and  $B = \begin{pmatrix} 2 \end{pmatrix}$  and  $C = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$ . Then, the corollary says that the matrices

$$\begin{pmatrix} 1 & 3 & 7 \\ 0 & 1 & 9 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

are similar.

*Proof of Corollary.* We know from the previous theorem (specifically, from its  $\mathcal{V} \implies \mathcal{U}$  direction) that there exists a matrix  $X \in \mathbb{C}^{n \times m}$  such that  $AX - XB = C$ . Consider this  $X$ .

Now, let  $S = \begin{pmatrix} I_n & X \\ 0 & I_m \end{pmatrix}$ . Now I claim that  $S$  is invertible and that

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = S \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} S^{-1}.$$

Once this claim is proved, the corollary will follow.

To see that  $S$  is invertible, we construct an inverse. Namely, we claim that  $\begin{pmatrix} I_n & -X \\ 0 & I_m \end{pmatrix}$  is an inverse for  $S$ . To show this, we just check that

$$\begin{aligned} S \begin{pmatrix} I_n & -X \\ 0 & I_m \end{pmatrix} &= \begin{pmatrix} I_n & X \\ 0 & I_m \end{pmatrix} \begin{pmatrix} I_n & -X \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} I_n I_n + X \cdot 0 & I_n (-X) + X I_m \\ 0 I_n + I_m 0 & 0 (-X) + I_m I_m \end{pmatrix} \\ &= \begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} = I_{n+m} \end{aligned}$$

and

$$\begin{pmatrix} I_n & -X \\ 0 & I_m \end{pmatrix} S = I_{n+m} \quad (\text{similarly}).$$

So we have shown that  $S$  is invertible. It remains to check that

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = S \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} S^{-1}.$$

To do so, we check the equivalent identity

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} S = S \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

This we do by computing both sides and comparing:

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} S = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I_n & X \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} A & AX \\ 0 & B \end{pmatrix}$$

and

$$S \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} I_n & X \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & C + XB \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & AX \\ 0 & B \end{pmatrix}$$

(since  $AX - XB = C$  entails  $C + XB = AX$ ).

We are done. □

## 2. Jordan canonical (aka normal) form

### 2.1. Similarity redux

**Definition 2.1.1.** Let  $A$  and  $B$  be two  $n \times n$ -matrices. We write  $A \sim B$  to say “ $A$  is similar to  $B$ ”.

We recall some properties of similar matrices: Similarity is an equivalence relation. Furthermore:

**Proposition 2.1.2.** Let  $A$  and  $B$  be two  $n \times n$ -matrices that satisfy  $A \sim B$ . Then:

- (a)  $A^k \sim B^k$  for each  $k \in \mathbb{N}$ .
- (b)  $A - \lambda I_n \sim B - \lambda I_n$  for each  $\lambda \in \mathbb{C}$ .
- (c)  $(A - \lambda I_n)^k \sim (B - \lambda I_n)^k$  for each  $\lambda \in \mathbb{C}$  and  $k \in \mathbb{N}$ .
- (d)  $p_A = p_B$ . (Recall that  $p_M$  stands for the characteristic polynomial of  $M$ .)
- (e)  $\dim \text{Ker } A = \dim \text{Ker } B$ .
- (f)  $\dim \text{Ker } ((A - \lambda I_n)^k) = \dim \text{Ker } ((B - \lambda I_n)^k)$  for each  $\lambda \in \mathbb{C}$  and  $k \in \mathbb{N}$ .

Recall the concepts of algebraic and geometric multiplicities:

**Definition 2.1.3.** Let  $A$  be an  $n \times n$ -matrix, and let  $\lambda \in \mathbb{C}$ .

(a) The **algebraic multiplicity** of  $\lambda$  as an eigenvalue of  $A$  is defined to be the multiplicity of  $\lambda$  as a root of  $p_A$ . (If  $\lambda$  is not an eigenvalue of  $A$ , then this is 0.)

(b) The **geometric multiplicity** of  $\lambda$  as an eigenvalue of  $A$  is defined to be  $\dim \text{Ker } (A - \lambda I_n)$ . In other words, it is the maximum number of linearly independent eigenvectors for eigenvalue  $\lambda$ . (If  $\lambda$  is not an eigenvalue of  $A$ , then this is 0.)

The geometric multiplicity is always  $\leq$  to the algebraic multiplicity. Sometimes, it is strictly smaller. For instance, the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

has only 1 as its eigenvalue, with algebraic multiplicity 3 and geometric multiplicity is 2.

## 2.2. Jordan cells

**Definition 2.2.1.** A **Jordan cell** is a  $k \times k$ -matrix of the form

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \quad \text{for some } k > 0 \text{ and some } \lambda \in \mathbb{C}.$$

So its diagonal entries are  $\lambda$ ; its entries directly above the diagonal are 1; all its other entries are 0. In formal terms, it is the  $k \times k$ -matrix  $A$  whose entries are given by the rule

$$A_{i,j} = \begin{cases} \lambda, & \text{if } i = j; \\ 1, & \text{if } i = j - 1; \\ 0, & \text{otherwise.} \end{cases}$$

Specifically, this matrix is called the **Jordan cell of size  $k$  at eigenvalue  $\lambda$** . It is denoted by  $J_k(\lambda)$ .

For example, the Jordan cell of size 3 at eigenvalue  $-5$  is

$$\begin{pmatrix} -5 & 1 & 0 \\ 0 & -5 & 1 \\ 0 & 0 & -5 \end{pmatrix}.$$

**Proposition 2.2.2.** Let  $k > 0$  and  $\lambda \in \mathbb{C}$ . The only eigenvalue of  $J_k(\lambda)$  is  $\lambda$ , and it has algebraic multiplicity  $k$  and geometric multiplicity 1.

*Proof.* Look at the matrix:

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

This matrix is upper-triangular, so its characteristic polynomial is

$$p_{J_k(\lambda)} = (t - \lambda)(t - \lambda) \cdots (t - \lambda) = (t - \lambda)^k.$$

Thus, its only eigenvalue is  $\lambda$ , and this eigenvalue has algebraic multiplicity  $k$ . Its geometric multiplicity is

$$\dim \operatorname{Ker} (J_k(\lambda) - \lambda I_k) = \dim \operatorname{Ker} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = 1.$$

□

**Proposition 2.2.3.** Let  $k > 0$  and  $\lambda \in \mathbb{C}$ . Let  $B = J_k(\lambda) - \lambda I_k$ . For any  $p \in \mathbb{N}$ , we have

$$\dim \operatorname{Ker} (B^p) = \begin{cases} p, & \text{if } p \leq k; \\ k, & \text{if } p > k. \end{cases}$$

*Proof.* As we have just seen,

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

For example, if  $k = 4$ , then

$$\begin{aligned}
 B &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \Rightarrow B^2 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \Rightarrow B^3 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \Rightarrow B^4 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \Rightarrow B^p &= 0 \text{ for all } p \geq 4.
 \end{aligned}$$

This generalizes to arbitrary  $k$ . Indeed, it is easy to see (by induction on  $p$ ) that

$$(B^p)_{i,j} = \begin{cases} 1, & \text{if } i = j - p; \\ 0, & \text{otherwise} \end{cases} \quad \text{for all } i, j.$$

Now, the claim about  $\dim \text{Ker}(B^p)$  easily follows. □

### 2.3. The Jordan canonical form (aka Jordan normal form)

**Definition 2.3.1.** A **Jordan matrix** is a block-diagonal matrix whose diagonal blocks are Jordan cells. In other words, it is a matrix of the form

$$\begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_k}(\lambda_k) \end{pmatrix},$$

where  $n_1, n_2, \dots, n_k$  are positive integers and  $\lambda_1, \lambda_2, \dots, \lambda_k$  are scalars in  $\mathbb{C}$  (not necessarily distinct, but not necessarily equal either).

We claim the following:

**Theorem 2.3.2** (existence of the Jordan canonical form). Every  $n \times n$ -matrix over  $\mathbb{C}$  is similar to a Jordan matrix.

This theorem is useful partly (but not only) because it allows to reduce questions about general square matrices to questions about Jordan matrices. And the latter can usually further be reduced to questions about Jordan cells, because a block-diagonal matrix “behaves like its diagonal blocks are separate”.

The above existence theorem has a matching uniqueness statement:

**Theorem 2.3.3** (Jordan canonical form theorem). Let  $A$  be an  $n \times n$ -matrix over  $\mathbb{C}$ . Then, there exists a Jordan matrix  $J$  such that  $A \sim J$ . Furthermore, this  $J$  is unique up to the order of the diagonal blocks.

**Definition 2.3.4.** The matrix  $J$  in this theorem is called a **Jordan canonical form** of  $A$  (or a **Jordan normal form** of  $A$ ).

**Example 2.3.5.** The Jordan canonical form of  $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 1 \end{pmatrix}$  is  $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ . Indeed,  $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$  is a Jordan matrix (it can be written as  $\begin{pmatrix} J_1(2) & 0 \\ 0 & J_2(1) \end{pmatrix}$ ) and it can be checked that

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 5 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

We can, of course, swap the two Jordan cells in this Jordan matrix, and obtain another Jordan canonical form of the same matrix.

**Example 2.3.6.** The Jordan canonical form of  $I_n$  is  $I_n$  itself. Indeed, each of its diagonal entries is itself a little Jordan cell  $J_1(1)$ , so it is already a Jordan matrix.

Now, we shall approach the proof of the Jordan canonical form theorem step by step, beginning with the uniqueness part.

**Example 2.3.7.** Suppose that  $A \sim J$  with

$$J = \begin{pmatrix} J_2(1) & 0 & 0 \\ 0 & J_3(1) & 0 \\ 0 & 0 & J_2(2) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

What can we say about  $A$ ?

First, since  $A \sim J$ , we have  $\sigma(A) = \sigma(J) = \{1, 2\}$  (since  $J$  is upper-triangular). Actually, from  $A \sim J$ , we also obtain

$$p_A = p_J = (t-1)^5 (t-2)^2 \quad \left( \begin{array}{c} \text{since } J \text{ and therefore also } tI_n - J \\ \text{is upper-triangular} \end{array} \right).$$

Thus, the algebraic multiplicities of the eigenvalues 1 and 2 of  $A$  are 5 and 2, respectively.

Now, what about the geometric multiplicities? What is  $\dim \text{Ker}(A - 1I_n)$ ?

In general, if  $B_1, B_2, \dots, B_k$  are square matrices, then

$$\dim \text{Ker} \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k \end{pmatrix} = \dim \text{Ker } B_1 + \dim \text{Ker } B_2 + \cdots + \dim \text{Ker } B_k,$$

since a block-structured vector  $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix}$  lies in  $\text{Ker} \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k \end{pmatrix}$  if and

only if each of its pieces  $v_i$  lies in  $\text{Ker } B_i$ . Thus,

$$\begin{aligned} & \dim \text{Ker}(A - 1I_n) \\ &= \dim \text{Ker}(J - 1I_n) \quad (\text{since } A \sim J) \\ &= \dim \text{Ker} \left( \begin{pmatrix} J_2(1) & 0 & 0 \\ 0 & J_3(1) & 0 \\ 0 & 0 & J_2(2) \end{pmatrix} - 1I_n \right) \\ &= \dim \text{Ker} \begin{pmatrix} J_2(1) - 1I_2 & 0 & 0 \\ 0 & J_3(1) - 1I_3 & 0 \\ 0 & 0 & J_2(2) - 1I_2 \end{pmatrix} \\ &= \underbrace{\dim \text{Ker}(J_2(1) - 1I_2)}_{=1 \text{ (by proposition above)}} + \underbrace{\dim \text{Ker}(J_3(1) - 1I_3)}_{=1 \text{ (by proposition above)}} + \underbrace{\dim \text{Ker}(J_2(2) - 1I_2)}_{=0 \text{ (since } J_2(2) - 1I_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix})} \\ &= 1 + 1 + 0 = 2. \end{aligned}$$

In general, this reasoning yields the following:

**Proposition 2.3.8.** Let  $A$  be an  $n \times n$ -matrix, and let  $J = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_k}(\lambda_k) \end{pmatrix}$  be its Jordan canonical form. Then:



(a) We have  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ .

(b) The geometric multiplicity of an eigenvalue  $\lambda$  of  $A$  is the number of Jordan cells of  $A$  at eigenvalue  $\lambda$ . In other words, it is the number of  $i \in [k]$  satisfying  $\lambda_i = \lambda$ .

(c) The algebraic multiplicity of an eigenvalue  $\lambda$  of  $A$  is the **sum** of the sizes of all Jordan cells of  $A$  at eigenvalue  $\lambda$ . In other words, it is  $\sum_{\substack{i \in [k]; \\ \lambda_i = \lambda}} n_i$ .