Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

October 22, 2021 (unfinished!)

Contents

Math 504 Lecture 14

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

1.1. The centralizer of a matrix

Here is a fairly natural question: Which matrices commute with a given square matrix A ?

Proposition 1.1.1. Let $\mathbb{F}$ be a field. Let $A \in \mathbb{F}^{n\times n}$ be an $n \times n$-matrix. Let f and g be two polynomials in a single variable t over $\mathbb{F}$. Then, $f(A)$ commutes with $g(A)$.

Proof. Write $f(t)$ as $f(t) = \sum\limits_{i=0}^{n} f_i t^i$, and write $g(t)$ as $g(t) = \sum\limits_{j=0}^{m} g_j t^j$. Then,

$$f(A) = \sum_{i=0}^{n} f_i A^i \qquad \text{and} \qquad g(A) = \sum_{j=0}^{m} g_j A^j.$$

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

Thus,

$$f(A) \cdot g(A) = \left(\sum_{i=0}^{n} f_i A^i\right) \cdot \left(\sum_{j=0}^{m} g_j A^j\right) = \sum_{i=0}^{n} \sum_{j=0}^{m} f_i g_j \underbrace{A^i A^j}_{=A^{i+j}} = \sum_{i=0}^{n} \sum_{j=0}^{m} f_i g_j A^{i+j}.$$

A similar computation shows that

$$g(A) \cdot f(A) = \sum_{i=0}^{n} \sum_{j=0}^{m} f_i g_j A^{i+j}.$$

Comparing these two, we obtain $f(A) \cdot g(A) = g(A) \cdot f(A)$, qed. □

Thus, in particular, $f(A)$ commutes with A for any polynomial f (because $A = g(A)$ for $g(t) = t$).

But are there other matrices that commute with A ?

There certainly can be. For instance, if $A = \lambda I_n$ for some $\lambda \in \mathbb{F}$, then **every** $n \times n$-matrix commutes with A (but very few matrices are of the form $f(A)$ for some polynomial f). This is, in a sense, the "best case scentario". Only for $A = \lambda I_n$ is it true that every $n \times n$-matrix commutes with A.

Let us study the general case now.

Definition 1.1.2. Let $A \in \mathbb{F}^{n \times n}$ be an $n \times n$-matrix. The **centralizer** of A is defined to be the set of all $n \times n$-matrices $B \in \mathbb{F}^{n \times n}$ such that $AB = BA$. We denote this set by $\operatorname{Cent} A$.

We thus want to know what $\operatorname{Cent} A$ is.

We begin with some general properties:

Proposition 1.1.3. Let $A \in \mathbb{F}^{n \times n}$ be an $n \times n$-matrix. Then, $\operatorname{Cent} A$ is a subset of $\mathbb{F}^{n \times n}$ that is closed under addition, scaling and multiplication and contains λI_n for all $\lambda \in \mathbb{F}$. In other words:

(a) For any $B, C \in \operatorname{Cent} A$, we have $B + C \in \operatorname{Cent} A$.

(b) For any $B \in \operatorname{Cent} A$ and $\lambda \in \mathbb{F}$, we have $\lambda B \in \operatorname{Cent} A$.

(c) For any $B, C \in \operatorname{Cent} A$, we have $BC \in \operatorname{Cent} A$.

(d) For any $\lambda \in \mathbb{F}$, we have $\lambda I_n \in \operatorname{Cent} A$.

This implies, in particular, that $\operatorname{Cent} A$ is a vector subspace of $\mathbb{F}^{n \times n}$. Furthermore, it shows that $\operatorname{Cent} A$ is an $\mathbb{F}$-subalgebra of $\mathbb{F}^{n \times n}$ (in particular, a subring of $\mathbb{F}^{n \times n}$).

Proof of the Proposition. Let me just show part **(c)**; the other parts are even easier.

(c) Let $B, C \in \operatorname{Cent} A$. Thus, $AB = BA$ and $AC = CA$. Now,

$$\underbrace{AB}_{=BA} C = B \underbrace{AC}_{=CA} = BCA.$$

This shows that $BC \in \operatorname{Cent} A$. Thus, part **(c)** is proved. □

Now, as an example, let us compute $\operatorname{Cent} A$ in the case when A is a single Jordan cell $J_n(0)$. So we fix an $n > 0$, and we set

$$A := J_n(0) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Let $B \in \mathbb{F}^{n \times n}$ be arbitrary. We want to know when $B \in \operatorname{Cent} A$. In other words, we want to know when $AB = BA$.

We have

$$AB = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & B_{2,3} & \cdots & B_{2,n} \\ B_{3,1} & B_{3,2} & B_{3,3} & \cdots & B_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & B_{n,3} & \cdots & B_{n,n} \end{pmatrix}$$

$$= \begin{pmatrix} B_{2,1} & B_{2,2} & B_{2,3} & \cdots & B_{2,n} \\ B_{3,1} & B_{3,2} & B_{3,3} & \cdots & B_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & B_{n,3} & \cdots & B_{n,n} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

and

$$BA = \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & B_{2,3} & \cdots & B_{2,n} \\ B_{3,1} & B_{3,2} & B_{3,3} & \cdots & B_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & B_{n,3} & \cdots & B_{n,n} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & B_{1,1} & B_{1,2} & \cdots & B_{1,n-1} \\ 0 & B_{2,1} & B_{2,2} & \cdots & B_{2,n-1} \\ 0 & B_{3,1} & B_{3,2} & \cdots & B_{3,n-1} \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & B_{n,1} & B_{n,2} & \cdots & B_{n,n-1} \end{pmatrix}.$$

Thus, $AB = BA$ holds if and only if

$$\begin{pmatrix} B_{2,1} & B_{2,2} & B_{2,3} & \cdots & B_{2,n} \\ B_{3,1} & B_{3,2} & B_{3,3} & \cdots & B_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & B_{n,3} & \cdots & B_{n,n} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & B_{1,1} & B_{1,2} & \cdots & B_{1,n-1} \\ 0 & B_{2,1} & B_{2,2} & \cdots & B_{2,n-1} \\ 0 & B_{3,1} & B_{3,2} & \cdots & B_{3,n-1} \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & B_{n,1} & B_{n,2} & \cdots & B_{n,n-1} \end{pmatrix},$$

i.e., if

$$\begin{aligned}
B_{2,j} &= B_{1,j-1} && \text{for all } j \in [n] \text{ (where } B_{1,0} := 0\text{);} \\
B_{3,j} &= B_{2,j-1} && \text{for all } j \in [n] \text{ (where } B_{2,0} := 0\text{);} \\
B_{4,j} &= B_{3,j-1} && \text{for all } j \in [n] \text{ (where } B_{3,0} := 0\text{);} \\
&\ldots; \\
B_{n,j} &= B_{n-1,j-1} && \text{for all } j \in [n] \text{ (where } B_{n-1,0} := 0\text{);} \\
0 &= B_{n,j} && \text{for all } j \in [n-1].
\end{aligned}$$

The latter system of equations can be restated as follows:

$$\begin{aligned}
&\ldots; \\
B_{n,n-2} &= B_{n-1,n-3} = B_{n-2,n-4} = \cdots = B_{3,1} = 0; \\
B_{n,n-1} &= B_{n-1,n-2} = B_{n-2,n-3} = \cdots = B_{2,1} = 0; \\
B_{n,n} &= B_{n-1,n-1} = B_{n-2,n-2} = \cdots = B_{1,1}; \\
B_{n-1,n} &= B_{n-2,n-1} = B_{n-3,n-2} = \cdots = B_{1,2}; \\
B_{n-2,n} &= B_{n-3,n-1} = B_{n-4,n-2} = \cdots = B_{1,3}; \\
&\ldots.
\end{aligned}$$

In other words, it means that the matrix B looks as follows:

$$B = \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{n-1} \\ & b_0 & b_1 & \cdots & b_{n-2} \\ & & b_0 & \cdots & b_{n-3} \\ & & & \ddots & \vdots \\ & & & & b_0 \end{pmatrix}$$

(where the empty cells have entries equal to 0). This is called an **upper-triangular Toeplitz matrix**. We can also rewrite it as

$$B = b_0 I_n + b_1 A + b_2 A^2 + \cdots + b_{n-1} A^{n-1}.$$

So we have proved the following:

Theorem 1.1.4. Let $n > 0$. Let $A = J_n(0)$. Then,

$$\begin{aligned}
\operatorname{Cent} A &= \left\{ \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{n-1} \\ & b_0 & b_1 & \cdots & b_{n-2} \\ & & b_0 & \cdots & b_{n-3} \\ & & & \ddots & \vdots \\ & & & & b_0 \end{pmatrix} \ \middle| \ b_0, b_1, \ldots, b_{n-1} \in \mathbb{F} \right\} \\
&= \left\{ b_0 I_n + b_1 A + b_2 A^2 + \cdots + b_{n-1} A^{n-1} \ \middle| \ b_0, b_1, \ldots, b_{n-1} \in \mathbb{F} \right\} \\
&= \left\{ f(A) \ \mid \ f \in \mathbb{F}[t] \text{ is a polynomial of degree } \leq n-1 \right\}.
\end{aligned}$$

So this is the worst-case scenario: The only matrices commuting with A are the matrices of the form $f(A)$ (which, as we recall, must always commute with A).

What happens for an arbitrary A ? Is the answer closer to the best-case scenario or to the worst-case scenario? The answer is that the worst-case scenario holds for a randomly chosen matrix, but we can actually answer the question "what is $\operatorname{Cent} A$ exactly" if we know the Jordan canonical form of A.

We start with simple propositions:

Proposition 1.1.5. Let $A \in \mathbb{F}^{n\times n}$ and $\lambda \in \mathbb{F}$. Then, $\operatorname{Cent}(A - \lambda I_n) = \operatorname{Cent} A$.

Proof. Exercise (diff. [1]). □

Proposition 1.1.6. Let A, B and S be three $n \times n$-matrices such that S is invertible. Then,

$$(B \in \operatorname{Cent} A) \iff \left(SBS^{-1} \in \operatorname{Cent}\left(SAS^{-1}\right)\right).$$

Proof. Exercise (diff. [1]). □

Thus, if A is a matrix with complex entries, and if we want to compute $\operatorname{Cent} A$, it suffices to compute $\operatorname{Cent} J$, where J is the JCF of A.

Therefore, we now focus on centralizers of Jordan matrices.

Proposition 1.1.7. Let $A_1, A_2, \ldots, A_k$ be square matrices with complex entries. Assume that the spectra of these matrices are disjoint – i.e., if $i \neq j$, then $\sigma(A_i) \cap \sigma(A_j) = \varnothing$.

Then,

$$\operatorname{Cent}\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix} = \left\{ \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix} \mid B_i \in \operatorname{Cent}(A_i) \text{ for each } i \in [k] \right\}.$$

Proof. The $\supseteq$ inclusion is obvious. We thus need to prove the $\subseteq$ inclusion only.

Let A_i be an $n_i \times n_i$-matrix for each $i \in [k]$.

Let $B \in \operatorname{Cent}\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$. We want to show that B has the form

$\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}$ where $B_i \in \operatorname{Cent}(A_i)$ for each $i \in [k]$.

Write B as a block matrix

$$B = \begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix},$$

where each $B(i,j)$ is an $n_i \times n_j$-matrix. Then, by the rule for multiplying block matrices, we have

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix} \begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix}$$
$$= \begin{pmatrix} A_1B(1,1) & A_1B(1,2) & \cdots & A_1B(1,k) \\ A_2B(2,1) & A_2B(2,2) & \cdots & A_2B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ A_kB(k,1) & A_kB(k,2) & \cdots & A_kB(k,k) \end{pmatrix}$$

and

$$\begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix} \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$
$$= \begin{pmatrix} B(1,1)A_1 & B(1,2)A_2 & \cdots & B(1,k)A_k \\ B(2,1)A_1 & B(2,2)A_2 & \cdots & B(2,k)A_k \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1)A_1 & B(k,2)A_2 & \cdots & B(k,k)A_k \end{pmatrix}.$$

However, these two matrices must be equal, since $\begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix} \in$

$\operatorname{Cent}\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$. Thus, we have

$$\begin{pmatrix} A_1 B(1,1) & A_1 B(1,2) & \cdots & A_1 B(1,k) \\ A_2 B(2,1) & A_2 B(2,2) & \cdots & A_2 B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ A_k B(k,1) & A_k B(k,2) & \cdots & A_k B(k,k) \end{pmatrix} = \begin{pmatrix} B(1,1) A_1 & B(1,2) A_2 & \cdots & B(1,k) A_k \\ B(2,1) A_1 & B(2,2) A_2 & \cdots & B(2,k) A_k \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) A_1 & B(k,2) A_2 & \cdots & B(k,k) A_k \end{pmatrix}.$$

Comparing blocks, we can rewrite this as

$$A_i B(i,j) = B(i,j) A_j \qquad \text{for all } i,j \in [k].$$

Now, let $i,j \in [k]$ be distinct. Consider this equality $A_i B(i,j) = B(i,j) A_j$. We can rewrite it as $A_i B(i,j) - B(i,j) A_j = 0$. Thus, $B(i,j)$ is an $n_i \times n_j$-matrix X satisfying $A_i X - X A_j = 0$. However, because $\sigma(A_i) \cap \sigma(A_j) = \varnothing$, a theorem we proved before (the Sylvester matrix equation) tells us that there is a **unique** $n_i \times n_j$-matrix X satisfying $A_i X - X A_j = 0$. Clearly, this unique matrix X must be the 0 matrix (since the 0 matrix satisfies $A_i 0 - 0 A_j = 0$). So we conclude that $B(i,j)$ is the 0 matrix. In other words, $B(i,j) = 0$.

So we have shown that $B(i,j) = 0$ whenever i and j are distinct. Thus,

$$B = \begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix} = \begin{pmatrix} B(1,1) & & & \\ & B(2,2) & & \\ & & \ddots & \\ & & & B(k,k) \end{pmatrix}.$$

This shows that B is block-diagonal. Now, applying the equation

$$A_i B(i,j) = B(i,j) A_j \qquad \text{for all } i,j \in [k]$$

to $j = i$, we obtain $A_i B(i,i) = B(i,i) A_i$, which of course means that $B(i,i) \in \operatorname{Cent}(A_i)$. Thus, B has the form $\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}$ where $B_i \in \operatorname{Cent}(A_i)$ for each $i \in [k]$. Proof complete. □

So we only need to compute $\operatorname{Cent} J$ when J is a Jordan matrix with only one eigenvalue.

We can WLOG assume that this eigenvalue is 0, since we know that $\operatorname{Cent}(A - \lambda I_n) = \operatorname{Cent} A$.

So we only need to compute $\operatorname{Cent} J$ when J is a Jordan matrix with zeroes on its diagonal.

If J is just a single Jordan cell, we already know the result (by the above theorem which describes $\operatorname{Cent} A$ for $A = J_n(0)$). In the general case, we have the following:

Proposition 1.1.8. Let J be a Jordan matrix whose Jordan blocks are

$$J_{n_1}(0), \quad J_{n_2}(0), \quad \ldots, \quad J_{n_k}(0).$$

Let B be an $n \times n$-matrix, written as a block matrix

$$B = \begin{pmatrix} B(1,1) & B(1,2) & \cdots & B(1,k) \\ B(2,1) & B(2,2) & \cdots & B(2,k) \\ \vdots & \vdots & \ddots & \vdots \\ B(k,1) & B(k,2) & \cdots & B(k,k) \end{pmatrix},$$

where each $B(i,j)$ is an $n_i \times n_j$-matrix. Then, $B \in \operatorname{Cent} J$ if and only if each of the k^2 blocks $B(i,j)$ is an **upper-triangular Toeplitz matrix in the wide sense**.

Here, we say that a matrix is an **upper-triangular Toeplitz matrix in the wide sense** if it

- has the form $\begin{pmatrix} 0 & U \end{pmatrix}$, where U is an upper-triangular Toeplitz (square) matrix and 0 is a zero matrix, or
- has the form $\begin{pmatrix} U \\ 0 \end{pmatrix}$, where U is an upper-triangular Toeplitz (square) matrix and 0 is a zero matrix.

(The zero matrices are allowed to be empty.)

Proof. Essentially the same argument that we used to prove the theorem about $J_n(0)$, just with a lot more bookkeeping involved. $\square$

We can summarize our results into a single theorem:

Theorem 1.1.9. Let $A \in \mathbb{C}^{n \times n}$ be an $n \times n$-matrix with Jordan canonical form J. Then, $\operatorname{Cent} A$ is a vector subspace of $\mathbb{C}^{n \times n}$ with dimension

$$\sum_{\lambda \in \sigma(A)} g_\lambda(A).$$

Here, for each eigenvalue λ of A, the number $g_\lambda(A)$ is a nonnegative integer defined as follows: Let $n_1, n_2, \ldots, n_k$ be the sizes of the Jordan blocks at eigenvalue λ that appear in J; then, we set

$$g_\lambda(A) := \sum_{i=1}^{k} \sum_{j=1}^{k} \min\{n_i, n_j\}.$$

Proof. Combine our above results and count the degrees of freedom. $\square$

Now, let us return to the worst-case scenario: When is $\operatorname{Cent} A = \{f(A) \mid f \in \mathbb{C}[t]\}$? We can answer this, too, although the proof takes longer.

Definition 1.1.10. An $n \times n$-matrix $A \in \mathbb{F}^{n \times n}$ is said to be **nonderogatory** if $q_A = p_A$ (that is, the minimal polynomial of A equals the characteristic polynomial of A).

A randomly chosen matrix with complex entries will be nonderogatory with probability 1; but there are exceptions. It is easy to see that if a matrix A has n distinct eigenvalues, then A is nonderogatory, but this is not an "if and only if"; a single Jordan cell is also nonderogatory.

Proposition 1.1.11. An $n \times n$-matrix $A \in \mathbb{C}^{n \times n}$ is nonderogatory if and only if its Jordan canonical form has exactly one Jordan block for each eigenvalue.

Proof. HW (difficulty [2]). □

Theorem 1.1.12. Let $A \in \mathbb{C}^{n \times n}$ be an $n \times n$-matrix. Then,

$$\operatorname{Cent} A = \{f(A) \mid f \in \mathbb{C}[t]\}$$

if and only if f is nonderogatory. Moreover, in this case,

$$\operatorname{Cent} A = \{f(A) \mid f \in \mathbb{C}[t] \text{ is a polynomial of degree } \leq n-1\}.$$

Proof. Later or exercises? □