LECTURE 19

The Cayley-Hamilton Theorem and the Jordan Decomposition

Let me begin by summarizing the main results of the last lecture. Suppose T is a endomorphism of a vector space V. Then T has a minimal polynomial that factorizes into a product of powers of irreducible polynomials

$$m_T(x) = p_1(x)^{s_1} \cdots p_k(x)^{s_k} \tag{1}$$

These irreducible factors can then be used to construct certain polynomials $f_1(x), \ldots, f_k(x)$ and corresponding operators $E_i \equiv f_i(T)$ which can be used to decompose the vector space V into a direct sum of T-invariant subspaces

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

Morever, we have both $V_i = E_i(V)$ (the image of V under E_i) and $V_i = \ker\left(p_i(T)^{s_i}\right)$ (the kernel of the operator $p_i(T)^{s_i}$). Since each V_i is T-invariant

$$v \in V_i \quad \Rightarrow \quad T(v) \in V_i$$

it follows that if we construct a basis by V by first choosing bases for each subspace V_i and then adjoining these bases to get a basis B for the entire vector space V, then with respect to B, the matrix T will take the block diagonal form

$$\mathbf{A}_{BB} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & & & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \mathbf{A}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_k \end{pmatrix}$$

What we aim to show in this lecture is that the submatrices $\mathbf{A}_i$ (which describe how T operators on the subspace V_i) can chosen to be upper triangular. A little more precisely, we shall show that we can always chose a basis for a given V_i so that the corresponding matrix $\mathbf{A}_i$ is upper triangular.

In what follows below we shall be making a special assumption about the factorization (1) of the minimal polynomial; namely

(2) **Assumption :** Each irreducible factor of m_T is of the form $p_i(x) = x - \xi_i$

where, of course, ξ_i is some eigenvalue of T. We note that whenever we are working over an *algebraically closed field* (like $\mathbb{C}$) this assumption on m_T will hold automatically. So we have

$$m_T(x) = (x - \xi_1)^{s_1} \cdots (x - \xi_k)^{s_k} \qquad , \qquad \text{with } \xi_i \neq \xi_j \text{ if } i \neq j. \tag{2'}$$

Now the first thing to point out is that since the subspaces V_i are also the kernel of the operators $(p_i(T))^{s_i} = (T - \xi_i \mathbf{1})^{s_i}$, if we defined operators

$$N_i \equiv T - \xi_i \mathbf{1}$$

then

$$(N_i)^{s_i} v = \mathbf{0}_V \qquad \forall \;\; v \in V_i = \ker\left((N_i)^{s_i}\right)$$

Thus, when we restrict the action of N_i to its corresponding subspace V_i, the s_i^{th} power of N_i vanishes identically. One says that N_i acts *nilpotently* on V_i.

1. Digression: Nilpotent Transformations

DEFINITION 19.1. *A linear transformation $N : W \to W$ of a vector space W is said to be* ***nilpotent*** *if there exists a positive integer k such that $N^k = \mathbf{0}_{L(W,W)}$. If N is nilpotent and s is the minimal integer such that $N^s = \mathbf{0}_{L(W,W)}$ we say that N is s-nilpotent.*

LEMMA 19.2. *If N is m-nilpotent, then its minimal polynomial is*

$$m_N(x) = x^m \quad .$$

Proof. By the definition of m-nilpotent, we certainly have

$$N^m = 0$$

and so, by Theorem 16.8, the minimal polynomial of N must divide x^m. Since x^m is a power of a single irreducible polynomial $(x-0)$, the only possibilities for $m_N(x)$ are other are polynomials of the form $(x-0)^k$ with $k \leq m$. But we still have to had

$$\mathbf{0}_{L(V,V)} = m_N(N) = N^k$$

but since m is the smallest k such that $N^k = 0$, we must have

$$m_N(x) = (x-0)^m \quad .$$

□

LEMMA 19.3. *Let $N \in L(V,V)$ be a nilpotent endomorphism of a finite-dimensional vector space V over an algebraically closed field. Then W has a basis $\{w_1, w_2, \ldots, w_n\}$ such that*

$$\begin{aligned}
N(v_1) &= \mathbf{0}_V \\
N(v_2) &\in span_{\mathbb{F}}(v_1) \\
N(v_3) &\in span_{\mathbb{F}}(v_1, v_2) \\
&\vdots \\
N(v_n) &\in span_{\mathbb{F}}(v_1, v_2, \ldots, v_{n-1}) \quad .
\end{aligned}$$

Proof. Since N is nilpotent, for sufficient large k we must have $N^k = \mathbf{0}_{L(W,W)}$. Let m be the mininal such k (so N is m-nilpotent). By the preceding lemma, the minimal polynomial of N is then

$$m_N(x) = x^m \quad .$$

Now consider the following family of subspaces of V

$$\begin{aligned}
V_0 \quad &: \quad = \{\mathbf{0}\} \\
V_1 \quad &: \quad = \ker(N) \\
V_2 \quad &: \quad = \ker\left(N^2\right) \\
&\vdots \\
V_{m-1} \quad &= \quad \ker\left(N^{m-1}\right) \\
V_m \quad &= \quad \ker(N^m) = V
\end{aligned}$$

I claim V_i is a proper subspace of V_{i+1} for $i = 0, \ldots, m-1$. Indeed, if $v \in V_i$, then $N^i v = \mathbf{0}$ and so also $N^{i+1} v = \mathbf{0}$ as well; so certainly $V_i \subseteq V_{i+1}$. To see that V_i is, in fact, a proper subspace of V_{i+1}, we argue as follows. Since m is the minimal power of N such that $N^m = \mathbf{0}_{L(V,V)}$ there has to be a vector w that survives N^{m-1};

$$N^i w \neq \mathbf{0} \quad \text{if } i < m \qquad \text{but} \quad N^m w = \mathbf{0} \quad .$$

Set

$$w_{i+1} = N^{m-i-1} w$$

Then

$$N^i w_{i+1} = N^{m-1} w \neq \mathbf{0}$$

but

$$N^{i+1} w_{i+1} = N^m w = \mathbf{0} \quad .$$

So

$$w_{i+1} \notin V_i \quad \text{but} \quad w_{i+1} \in V_{i+1}.$$

Thus, each V_i is a proper subset of V_{i+1}.

And so we have the following chain of proper subspaces

$$\{\mathbf{0}_W\} \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_{m-1} \subsetneq V_m = V \quad .$$

Next we note that

$$N(V_i) \subset V_{i-1}$$

Indeed, if $v \in V_i$, we have $N^i v = 0$. But then Nv will satisfy $N^{i-1}(Nv) = \mathbf{0}$, and so $Nv \in V_{i-1}$.

Ok. We are now ready to construct a basis B of V that has the property

$$N(v_i) \subset \mathit{span}(v_1, \ldots, v_{i-1}) \ .$$

To do this we just apply Theorem 5.4 repeatedly. Let $k_i = \dim V_i$. Since V_1 is a subspace it has a basis $B_1 = \left\{ v_1^{(1)}, \ldots, v_{k_1}^{(1)} \right\}$. Since V_1 is a subspace of V_2, by Theorem 5.4, V_2 has a basis B_2 that extends the basis B_1 of V_1.

$$B_2 = \left\{ v_1^{(1)}, \ldots, v_k^{(1)}, v_{k_1+1}^{(2)}, \ldots, v_{k_2}^{(2)} \right\} \quad .$$

Then, in a similar fashion, we can extend B_2 to a basis B_3 of V_3, B_3 to a basis B_4 of V_4, and so on. In the end, we'll arrive at a basis

$$B = \left\{ v_1^{(1)}, \ldots, v_{k_1}^{(1)}, v_{k_1+1}^{(2)}, \ldots, v_n^{(m)} \right\} \text{ of } V_m = V \quad .$$

Now because

$$NV_i \subset V_{i-1}$$

we will have to have

$$N\left(v_j^{(i)}\right) \in V_{i-1} = \mathit{span}\left(v_1^{(1)}, \ldots, v_{k_{i-1}}^{(i-1)}\right)$$

And in fact, by the way we ordered to lower indices on the basis elements $v_j^{(i)}$, we have $k_{i-1} < j \leq k_i$. Thus,

$$N\left(v_j^{(i)}\right) \in \mathit{span}\left(v_1^{(1)}, \ldots, v_{k_{i-1}}^{(i-1)}\right) \subseteq \mathit{span}\left(v_1^{(1)}, \ldots, v_{k_{i-1}}^{(i-1)}, v_{k_{i-1}+1}^{(i)}, \ldots, v_{j-1}^{(i)}\right)$$

which, is the statement we sought to prove (by simply ignoring the upper indices, which after all are just a notational device to indicate how we constructed the basis B). $\square$

Remark 19.4. Another way of framing this result is as follows. If N is a nilpotent operator acting a finite-dimensional vector space W, then there exists a basis B for W such the matrix of N with respect to B is an upper triangular matrix with 0's along the diagonal:

$$\mathbf{N}_{BB} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \quad .$$

To see this, consider the j^{th} column vector $\mathbf{c}_j$ of N_{BB}.

$$\mathbf{c}_j = N(v_j)_B \quad .$$

Since $N(v_j) \in \mathit{span}(v_1, \ldots, v_{j-1})$, the *coordinates* of $N(v_j)$ with respect to $v_j, v_{j+1}, \ldots, v_n$ must all be 0. So we have

$$(\mathbf{N}_{BB})_{ij} = (\mathbf{c}_j)_i = 0 \qquad \text{if} \qquad i \geq j \quad .$$

So if we say that $n \times n$ matrix $\mathbf{A}$ is **strictly upper triagular** whenever its entries satisfy

$$A_{ij} = 0 \qquad \text{whenever } j \leq i$$

then the preceding theorem can be rephrased more succinctly as saying:

- If $N \in L(V, V)$ is a nilpotent transfomation, there exists a basis for V such that matrix representing N is strictly upper triangular.

REMARK 19.5. Another observation on can make is that if $N \in L(W, W)$ is nilpotent and s is the minimal integer such that $N^s = \mathbf{0}_{L(W,W)}$. Then W has to be at least s dimensional (because each application of N confines a vector to a smaller and smaller subspace of W. Cf. equation (3)).

THEOREM 19.6. *Let $T \in L(V, V)$ and suppose T has a minimal polynomial of the form*

$$m_T(x) = (x - \alpha_1)^{s_1} \cdots (s - \alpha_k)^{s_k} \quad .$$

Then there exists a basis $\{v_1, \ldots, v_n\}$ of V such that the matrix $\mathbf{A}_T$ of T with respect to this basis has a block diagonal form

$$\mathbf{A}_T = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & & & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \mathbf{A}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_k \end{pmatrix}$$

and, moreover, submatrix A_i along the diagonal is of the form

$$\mathbf{A}_i = \begin{pmatrix} \alpha_i & * & \cdots & * \\ 0 & \alpha_i & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_i \end{pmatrix}$$

with only zero entries appearing below the diagonal. Moreover, the sizes of these submatrices $\mathbf{A}_i$ is $\geq s_i$.

Proof. We have already remarked at the beginning of this lecture how (via Theorem 17.17) how the minimal polynomial m_T of T leads to a direct sum decomposition of V into T-invariants subspaces

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

each subspace corresponding to a particular irreducible factor occuring in the minimal polynomial. In fact, any basis of V constructed by adjoining bases of the direct summands $V_1, \ldots, V_k$ will cast $\mathbf{A}$ in this block diagonal form. What we need to show is that we can adopt bases B_i of the individual subspace V_i so that when T is restriced to V_i its matrix (with respect to the basis B_i of V_i) is upper triagular.

Now recall the subspaces V_i can be identified with the kernel of the operator $(p_i(T))^{s_i} = (T - \alpha_i \mathbf{1})^{s_i}$. This means the operator $(T - \alpha_i \mathbf{1})$ is a nilpotent operator on V_i. By the preceding lemma then, there exists a basis for V_i such that $T - \alpha_i \mathbf{1}$ is upper triagular with zeros along the diagonal. Thus

$$(\mathbf{T} - \alpha_i \mathbf{I})_{BB} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

and so

$$\mathbf{T}_{BB} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} + \begin{pmatrix} \alpha_i & 0 & \cdots & 0 \\ 0 & \alpha_i & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_i \end{pmatrix} = \begin{pmatrix} \alpha_i & * & \cdots & * \\ 0 & \alpha_i & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_i \end{pmatrix}$$

Finally, we point out that since $(T - \alpha_i \mathbf{1})$ is s_i-nilpotent on V_i, the dimension of V_i has to be at least s_i. □

COROLLARY 19.7. *Let V be a vector space over an algebraically closed field and let $T \in L(V,V)$. Then the minimal polynomial of T divides the characteristic polynomial of T.*

Proof. Let

$$m_T(x) = (x-\alpha_1)^{s_1} \cdots (x-\alpha_k)^{s_k} \tag{4}$$

be the minimal polynomial of T (guaranteed to be of this form since the underlying field is algebraically closed). By the preceding theorem we can find a basis B for V such that the matrix for T with respect to B takes an upper triangular form, with entries $\alpha_1, \dots \alpha_k$ along the diagonal. Then $(\mathbf{T}_{BB} - x\mathbf{1})$ will also be upper triangular, but with entries $\alpha_i - x$ along the diagonal. Since the determinant of an upper triangular matrix is just the product of diagonal elements, we will have

$$p_T(x) = \det(\mathbf{T}_{BB} - x\mathbf{1}) = (\alpha_1 - x)^{d_1} \cdots w (\alpha_k - x)^{d_k}$$

where d_i is the size of the submatrix $\mathbf{A}_i$ ($d_i = \dim V_i$). As remarked at the end of the proof of Theorem 18.5, we have $d_i \geq s_i$. Thus, minimal polynomial will divide the characteristic polynomial. □

COROLLARY 19.8 (Cayley-Hamilton). *If $T \in L(V,V)$, and $p_T(x)$ is its characteristic polynomial, then $p_T(T) = \mathbf{0}_{L(V,V)}$.*

Proof. By preceding corollary,

$$p_T(x) = q(x)\, m_t(x)$$

for some polynomial $q(x)$. But then

$$p_T(\backslash T) = q(T\ )\, m_T(T) = q(T\ )\, \mathbf{0}_{L(V,V)} = \mathbf{0}_{L(V,V)} \quad .$$

□

2. The Jordan Decomposition

Let V be a vector space over an algebraically closed field. We have seen that the minimal polynomial m_T of a linear transformation $T \in L(V,V)$ provides us a natural direct sum decomposition of a vector space V

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

such that T acts invariantly and, in fact, upper-triagularly on each subspace V_i. We'll now see that T itself decomposes in a particular nice, predictable way.

LEMMA 19.9. *If $T \in L(V,V)$ is diagonalizable. Then its restriction to any T-invariant subspace W of T is diagonalizable.*

Proof. If T is diagonalizable, then its minimal polynomial is of the form

$$m_T(x) = (x-\alpha_1)\cdots(x-\alpha_k) \qquad , \qquad \alpha_i \neq \alpha_j \text{ if } i \neq j \quad , \tag{4}$$

where $\alpha_1, \dots, \alpha_k$ are the eigenvalues of T. Suppose W is a T-invariant subspace of V. Then any power of T, or indeed, any polynomial in T will preserve W. Because of this,

$$f\left(T|_W\right) = f(T)|_W = f(T|)_W$$

that is to say the restriction of any polynomial in T to W makes sense and, in fact, it amounts to the same polynomial f "evaluated" at the restriction $T|_W$ of T to W. In particular, we'll have

$$\mathbf{0}_{L(V,V)} = m_T(T) \quad \Rightarrow \quad \mathbf{0}_{L(W,W)} = m_T\left(T|_W\right) \quad .$$

Now it should be pointed out that this does not imply that m_T is also the minimal polynomial of $T|_W$, indeed, it might not be (e.g. if W were comprised of a single eigenspace then $m_{T|_W} = (x-\alpha_i))$, it does imply that the minimal polynomial of $T|_W$ must divide m_T. Because of the factorization (4) of m_T, it must then be that the minimal polynomial of $T|_W$ contain the same kind of factors. In other words, the minimal polynomial of $T|_W$ will also have the form (4), except with possibly some factors missing. But then having a minimal polynomial of the form (4) means, there exists a basis B_W of W for which $T|_W$ is diagonalizable. □

Lemma 19.10. *If $S, T \in L(V,V)$ are diagonalizable transformations such that $ST = TS$, then there exists a basis for V in which both T and S act diagonally.*

Proof. Let $\{\lambda_1, \ldots, \lambda_k\}$, $\{\alpha_1, \ldots, \alpha_\ell\}$ be the eigenvalues of, respectively, T and S. Suppose we decompose V into its T-eigenspaces

$$V_i = \{v \in V \mid Tv = \lambda_i v\}$$

Then each subspace V_i is preserved by S. For if $v_i \in V_i$

$$T(S(v_i)) = S(T(v_i)) = S(\lambda_i v_i) = \lambda_i S(v_i) \quad \Rightarrow \quad S(v_i) \text{ is an eigenvector of } T \text{ with eigenvalue } \lambda_i \quad \Rightarrow \quad S(v_i) \in V_i$$

But then by the preceding lemma, on each of the S-invariant subspace V_i, S is diagonalizable. Therefore, each V_i has an S-eigenspace decompositio

$$V_i = V_{i,1} \oplus V_{i,2} \oplus \cdots \oplus V_{i,\ell} \qquad ; \qquad V_{i,j} = \{v \in V_i \mid Sv = \alpha_j v\}.$$

Now we can choose any bases we want for the subspace $V_{i,j}$, $i = 1, \ldots, k$, $j = 1, \ldots, \ell$, and adjoin these bases to get a basis B for V. Each of the basis vectors in B will live in one particular $V_{i,j}$ and so will be simultaneously an eigenvector for T and S:

$$v \in V_{i,j} \quad \Rightarrow \quad T(v) = \lambda_i v \quad \text{and} \quad S(v) = \alpha_j v \quad .$$

Thus, with respect to the basis so constructed, both S and T will act diagonally. □

Theorem 19.11. *Let V be a vector space over an algebraically closed field and let $T \in L(V,V)$. Then*

$$T = D + N \qquad , \qquad D, N \in L(V,V)$$

where

(i) *$D \in L(V,V)$ is a diagonalizable linear transformation;*
(ii) *$N \in L(V,V)$ is a nilpotent transformation;*
(iii) *There exist polynomials $f(x)$ and $g(x) \in \mathbb{F}[x]$ such that $D = f(T)$ and $N = g(T)$.*
(iv) *The transformations D and N commute: $D \circ N = N \circ D$.*
(v) *The transformations D and N are uniquely determined in the sense that if $T = D' + N'$ with D' diagonalizable, N' nilpotent and $D' \circ N' = N' \circ D'$, then $D' = D$ and $N' = N$.*

Proof. Since $\mathbb{F}$ is algebraically closed, the irreducible polynomials in $\mathbb{F}[x]$ are all of the form $x - \alpha_i$, and so the minimal polynomial has the form

$$m_T(x) = (x - \alpha_1)^{s_1} \cdots (x - \alpha_k)^{s_k}$$

and we have via Theorem 17.17 a corresponding direct sum decomposition of V

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k \tag{4}$$

into T invariant subspaces where

$$V_i = E_i V$$

where the operators E_i are of the form $E_i = f_i(T)$ for some polynomal $f \in \mathbb{F}[x]$, and satisfy

$$E_i E_j = \begin{cases} E_i & \text{if } i = j \\ \mathbf{0}_{L(V,V)} & \text{if } i \neq j \end{cases}$$

Consider the operator

$$D = \alpha_1 E_1 + \alpha_2 E_2 + \cdots + \alpha_k E_k$$

Note that

$$D = \alpha_1 f_1(T) + \alpha_2 f_2(T) + \cdots + \alpha_k f_k(T)$$

and so is a polynomial in T as desired. Morever, because it is a polynomial in T it will preserve any T-invariant subspace. So it makes sense to restrict it to any of the subspaces V_i

Let $v \in V$. Its component v_i in V_i will be $E_i(v)$. But then

$$\begin{aligned} D(v_i) &= (\alpha_1 E_1 \cdots + \alpha_k E_k) \circ E_i(v) = \alpha_1 E_1 E_i(v) + \cdots + \alpha_i E_i E_i(v) + \cdots + \alpha_k E_k E_i(v) \\ &= \mathbf{0}_V + \cdots + \mathbf{0}_v + \alpha_i E_i(v) + \mathbf{0}_V \ + \cdots + \mathbf{0}_V \\ &= \alpha_i v_i \end{aligned}$$

and so on each of the subspaces V_i, D will simply act by scalar multiplication by α_i. Hence D will be diagonaliable, with eigenvalues α_i.

Now let

$$N = T - D \quad .$$

Since D is a polynomial in T so will be N. We have

$$N v_i = (T - D) v_i = (T - \alpha_i) v_i$$

Now recall that Theorem 17.17 tells us also that the subspaces V_i in the decomposition (4) are also identifiable as $\ker\left((T - \alpha_i)^{s_i}\right)$. This implies

$$N^{s_i} v_i = \mathbf{0}_V$$

and if we choose $n = \max(s_1, \ldots, s_k)$ then we'll have

$$N^n v_i = \mathbf{0}_V \qquad i = 1, 2, \ldots, k$$

and thus, $N^n(v) = \mathbf{0}_V$ for all $v \in V$. Thus, N is nilpotent.

Note also, that since both N and D are polynomials in T we will have automatically that $N \circ D = D \circ N$.

It remains to prove the uniqueness of N and D. Suppose that N' and D' satisfy

$$\begin{aligned} T &= D' + N' \\ D'N' &= N'D' \\ & \quad D' \text{ is diagonalizable} \\ & \quad N' \text{ is nilpotent} \end{aligned}$$

Then we have

$$TD' = (D' + N') D' = D'D' + N'D' = D'D' + D'N' = D'(D' + N') = D'T$$

and similarly $TN' = N'T$. But then $D'D = DD'$ and $N'N = NN'$ since D and N are polynomial in T. From

$$D + N = T = D' + N'$$

we also have

$$D' - D = N - N'$$

Now since N and N' commute, we can use the binomial theorem to expand powers of $(N' - N)$

$$(N - N')^m = \sum_{k=0}^{m} \begin{pmatrix} m \\ k \end{pmatrix} (N')^{m-k} (N)^k$$

Now because N' and N' are nilpotent there exist integers n and n' such that $(N)^n = \mathbf{0}_{L(V,V)}$ and $(N')^{n'} = \mathbf{0}_{L(V,V)}$. Therefore if we choose m larger than say $\max(n, n')/2$ then all the terms in $(N - N')^m$ will vanish. Hence, $N - N'$ is nilpotent. On the other hand, since the matrices D and D' commute, they are simultaneoulsy diagonalizable. And so $D - D'$ can be diagonalized, and in its diagonalizing basis must take the form

$$N - N' = D' - D \sim \begin{pmatrix} \beta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \beta_n \end{pmatrix}$$

But then for $(N - N')^m = 0$, we will need each $\beta_i^m = 0$, Hence $D = D'$, and hence $N = N'$.