Math200b, lecture 8

Golsefidy

Char polynomial of companion matrices

In the previous lecture we proved that the characteristic polynomial of a matrix (with entries in a field k) is the product of its invariant factors; this had been done modulo the fact that the characteristic polynomial of the companion matrix of a monic polynomial $g(x) \in k[x]$ is $g(x)$.

Lemma 1 *Suppose* $g(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_0 \in k[x]$ *and* $c(g)$ *is the companion matrix of* g. *Then* $f_{c(g)}(x) = g(x)$ *where* $f_{c(g)}(x)$ *is the characteristic polynomial of* $c(g)$.

Proof. We proceed by induction on $\deg g$. Base of induction is clear; so we focus on the induction step:

$$f_{c(g)}(x) = \det(xI - c(g)) = \det \begin{pmatrix} x & 0 & \cdots & 0 & c_0 \\ -1 & x & \cdots & 0 & c_1 \\ 0 & -1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & x + c_{n-1} \end{pmatrix}.$$

$$= x \det \begin{pmatrix} x & \cdots & 0 & c_1 \\ -1 & \cdots & 0 & c_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -1 & x + c_{n-1} \end{pmatrix} + (-1)^{n+1} c_0 \det \begin{pmatrix} -1 & x & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}$$

by expanding determinant with respect to the first row. Notice that the first matrix is $xI - c(x^{n-1} + c_{n-1}x^{n-2} + \cdots + c_1)$; and so by the induction hypothesis the first term is just

$$x(x^{n-1} + c_{n-1}x^{n-2} + \cdots + c_1).$$

The matrix in the second term is an upper-triangular matrix and so its determinant is the product of its diagonal entries and so the second term is $(-1)^{n+1}c_0(-1)^{n-1} = c_0$. Overall we get

$$f_{c(g)}(x) = x(x^{n-1} + c_{n-1}x^{n-2} + \cdots + c_1) + c_0 = g(x);$$

and claim follows. ■

Char polynomial of nilpotent matrices

Let's see how the theorem that we proved about the connections between characteristic polynomial, minimal polynomial and invariant factors can help us to get a better understanding of nilpotent matrices.

Proposition 2 *Suppose* k *is a field and* $N \in M_n(k)$ *is a nilpotent matrix. Then* $N^n = 0$.

Proof. Since N is nilpotent, $N^m = 0$ for some $m \in \mathbb{Z}^+$. Hence $m_N(x)|x^m$ where $m_N(x)$ is the minimal polynomial of N; and so $m_N(x) = x^l$ for some positive integer l. Since any irreducible factor of the characteristic polynomial $f_N(x)$ is also an irreducible factor of $m_N(x)$ and x is the only irreducible factor of $m_N(x)$, we deduce that x is the only irreducible factor of $f_N(x)$; hence $f_N(x)$ is also a power of x. As $\deg f_N(x) = n$, $f_N(x) = x^n$. Therefore by the Cayley-Hamilton theorem $N^n = 0$. ■

Notice that all n-by-n nilpotent matrices have the same characteristic polynomial; but they are not necessarily similar, for instance one can be zero and the other non-zero. Even if $m_{N_1}(x) = m_{N_2}(x)$, we cannot deduce that they are similar. By Rational Canonical Form, we need to know all the invariant factors in order to get similarity; and $m_N(x)$ and $f_N(x)$ cannot give us all the invariant factors unless we were told that there are at most two invariant factors or $\deg m_N(x) = \deg f_N(x)$.

Jordan form

Can we get a better understanding of a matrix up to similarity assuming all of its eigenvalues are in k? For instance over $\mathbb{C}$ we know any polynomial can be written as a product of degree one terms; and so all the eigenvalues of a given complex matrix is in $\mathbb{C}$. Or all the eigenvalues of a nilpotent matrix are 0. Can this be used to get a better understanding of the similarity class of a matrix A? We have already seen that the similarity class of A can be completely understood by looking at the $k[x]$-module V_A. And if $f_1|f_2|\cdots|f_m$ are invariant factors of A, then

$$V_A \simeq k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle. \tag{1}$$

By our assumption there are distinct λ_i's in k such that

$$f_A(x) = \prod_{i=1}^{l}(x-\lambda_i)^{n_i}.$$

Since $f_A(x) = \prod_{i=1}^{m} f_i(x)$, there are $n_{ij} \in \mathbb{Z}^{\geq 0}$ such that

$$f_j(x) = \prod_{i=1}^{l}(x-\lambda_i)^{n_{ij}}.$$

We notice that, since λ_i's are distinct, $(x-\lambda_i)^{n_{ij}}$ are pairwise coprime for a fixed j and $1 \leq i \leq l$. And so by Chinese Remainder Theorem for $k[x]$ we have that

$$k[x]/\langle f_j(x)\rangle \xrightarrow{\phi} \bigoplus_{i=1}^{l} k[x]/\langle (x-\lambda_i)^{n_{ij}}\rangle,$$

$$\phi(p(x)+\langle f_j(x)\rangle := (p(x)+\langle (x-\lambda_i)^{n_{ij}}\rangle)_{i=1}^{l}, \tag{2}$$

is a $k[x]$-module isomorphism (and also ring isomorphism).

Let's quickly prove the Chinese Remainder Theorem for PIDs. What we will prove holds for any unital commutative ring; but here for the sake of brevity we refrain from going to the general case.

Theorem 3 (Chinese Remainder Theorem for PIDs) *Suppose* D *is a PID,* $\mathfrak{a}_i \trianglelefteq D$*, and* $\mathfrak{a}_i + \mathfrak{a}_j = D$ *is* $i \neq j$ *(co-primeness). Then*

$$\phi : D/\bigcap_{i=1}^{n} \mathfrak{a}_i \to \bigoplus_{i=1}^{n} D/\mathfrak{a}_i, \phi(a + \bigcap_{i=1}^{n} \mathfrak{a}_i) := (a + \mathfrak{a}_i)_{i=1}^{n} \tag{3}$$

is an D*-module and ring isomorphism.*

Proof. Let $\widetilde{\phi} : D \to \bigoplus_{i=1}^{n} D/\mathfrak{a}_i, \widetilde{\phi}(a) := (a + \mathfrak{a}_i)_{i=1}^{n}$. Then clearly $\widetilde{\phi}$ is a ring and D-module homomorphism. So by the first isomorphism theorem (in ring theory and module theory), we have that

$$\phi : D/\ker \widetilde{\phi} \to \bigoplus_{i=1}^{n} D/\mathfrak{a}_i, \phi(a) := (a + \mathfrak{a}_i)_{i=1}^{n}$$

is a well-defined injective ring and D-module homomorphism. It is easy to see that $\ker \widetilde{\phi} = \bigcap_{i=1}^{n} \mathfrak{a}_i$; and so ϕ given in (3) is a well-defined injective D-module and ring homomorphism. To finish the proof, we need to show that ϕ is surjective. To do so it is enough to show that $(0,\cdots,0,\underbrace{1}_{i-th},0,\cdots,0)$ is in the image of ϕ for any i. This means we need to find $a \in D$ such that $a + \mathfrak{a}_i = 1 + \mathfrak{a}_i$ and $a \in \bigcap_{j\neq i}\mathfrak{a}_j$; this is equivalent to say that $\mathfrak{a}_i + \bigcap_{j\neq i}\mathfrak{a}_j = D$.

Since D is a PID, there are $a_j \in D$ such that $\mathfrak{a}_j = \langle a_j\rangle$. As $\mathfrak{a}_i + \mathfrak{a}_j = D$, we have that $\gcd(a_i, a_j) = [1]$; that means that a_i and a_j do not have any common irreducible factor. Notice that, since D is a PID, $\bigcap_{j\neq i}\mathfrak{a}_j$ is generated by $\operatorname{lcm}(a_j)_{j\neq i}$; and, as a_j's are pairwise co-prime,

$$\operatorname{lcm}(a_j)_{j\neq i} = \prod_{j\neq i} a_j \text{ and } \gcd(a_i, \prod_{j\neq i} a_j) = [1].$$

Hence $\mathfrak{a}_i + \bigcap_{j\neq i}\mathfrak{a}_j = \langle a_i\rangle + \langle\prod_{j\neq i} a_j\rangle = D$; and claim follows. ■

Remark. We used the PID condition only in the last paragraph; and this part can be proved without the PID assumption.

Going back to understanding the similarity class of A, by (1) and (2) we have

$$V_A \simeq \bigoplus_{i=1}^{l}\bigoplus_{j=1}^{m} k[x]/\langle(x-\lambda_i)^{n_{ij}}\rangle. \tag{4}$$

To get back to linear algebra, we need to have a "nice" matrix representation of $x \times \cdot$ (multiplication by x) in $k[x]/\langle(x-\lambda_i)^{n_{ij}}\rangle$;

this is needed as A is a matrix representation of the multiplication by x in V_A. We can take the companion matrix of $(x-\lambda_i)^{n_{ij}}$; but then binomial coefficients will be needed which makes it hard to work with the given matrix. If $\lambda_i = 0$, then the companion matrix is easy to work with. So first we shift and then look at the matrix representation:

$\widetilde{\theta} : k[x] \to k[y], \widetilde{\theta}(f(x)) := f(y+\lambda)$ is a k-linear ring isomorphism (we say it is a k-algebra isomorphism); and so we get a k-algebra isomorphism $\theta : k[x]/\langle (x-\lambda)^n\rangle \to k[y]/\langle y^n\rangle$. Hence we get the following commuting diagram:

$$\begin{array}{ccc} k[x]/\langle (x-\lambda)^n\rangle & \xrightarrow{\theta} & k[y]/\langle y^n\rangle \\ \big\downarrow{\scriptstyle \times x} & & \big\downarrow{\scriptstyle \times(y+\lambda)} \\ k[x]/\langle (x-\lambda)^n\rangle & \xrightarrow{\theta} & k[y]/\langle y^n\rangle \end{array}$$

In the right column, multiplication by y can be represented by the companion matrix of y^n; so $\times(y+\lambda)$ can be represented by

$$J_n(\lambda) := \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}.$$

The above commuting diagram implies that multiplication by x in $k[x]/\langle (x-\lambda)^n \rangle$ can be also represented by $J_n(\lambda)$; we call $J_n(\lambda)$ a Jordan block. Altogether we get the following:

Lemma 4 $k[x]/\langle (x-\lambda)^n \rangle \simeq V_{J_n(\lambda)}$ *as* $k[x]$*-modules.*

By Lemma 4 and (4) we have

$$V_A \simeq \bigoplus_{i=1}^{l} \bigoplus_{j=1}^{m} V_{J_{n_{ij}}(\lambda_i)} \simeq V_{\mathrm{diag}(J_{n_{ij}}(\lambda_i))_{i,j}}$$

as $k[x]$-modules; and so A is similar to the matrix $\mathrm{diag}(J_{n_{ij}}(\lambda_i))_{i,j}$ that has Jordan blocks $J_{n_{ij}}(\lambda_i)$ on its diagonal. This is called a Jordan Form of A.

Theorem 5 (Jordan Form) *Suppose* k *is a field,* $A \in M_n(k)$*, all the eigenvalues of* A *are in* k*, and* $\lambda_1, \ldots, \lambda_l$ *are distinct eigenvalues of* A*. Then there are unique increasing sequences (with finitely many terms) of positive integers* $n_{1j} \le n_{2j} \le \cdots$ *for* $1 \le i \le l$ *such that* A *is similar to* $\mathrm{diag}(J_{n_{ij}}(\lambda_i))_{i,j}$.

Proof. We have already proved the existence part; so we focus on the uniqueness part. Suppose A is similar to $\mathrm{diag}(J_{n_{ij}}(\lambda'_i))_{i,j}$; comparing eigenvalues of both sides we deduce that λ'_i's are a reordering of λ_i's. So after relabelling, if needed, we can and will assume that $\lambda_i = \lambda'_i$. To show the uniqueness of n_{ij}'s, similar to our approach in the uniqueness part of Rational Canonical Form, we will show that n_{ij}'s can be determined by the $k[x]$-module structure of V_A; and the latter is determined uniquely by the similarity class of A. So we get that Jordan form can be determined by the similarity class of A (up to a reordering of its Jordan blocks).

Since A is similar to $\mathrm{diag}(J_{n_{ij}}(\lambda_i))_{i,j}$, we have

$$V_A \simeq V_{\mathrm{diag}(J_{n_{ij}}(\lambda_i))_{i,j}} \simeq \bigoplus_{i,j} V_{J_{n_{ij}}(\lambda_i)} \simeq \bigoplus_{i,j} k[x]/\langle (x-\lambda_i)^{n_{ij}} \rangle$$

as $k[x]$-modules. Similar to the proof of the uniqueness part of Rational Canonical Form, we will consider $\frac{(x-\lambda_i)^s V_A}{(x-\lambda_i)^{s+1} V_A}$; to be precise in the proof of the Rational Canonical Form, we first localized

Note that in a PID D, if $\gcd(a,b)=1$, then there are $r,s \in D$ such that $ar+bs=1$; and so $(a+\langle b\rangle)(r+\langle b\rangle) = 1+\langle b\rangle$ which implies that $a+\langle b\rangle \in (D/\langle b\rangle)^{\times}$.

By the above fact, $(x-\lambda_r)^s + \langle (x-\lambda_j)^{n_{ij}}\rangle$ is a unit in the ring $k[x]/\langle (x-\lambda_j)^{n_{ij}}\rangle$ for any positive integer s and $r \neq j$; and so multiplication by $(x-\lambda_r)^s$ does not change $k[x]/\langle (x-\lambda_j)^{n_{ij}}\rangle$.

Therefore for any non-negative integer s we have

$$(x-\lambda_r)^s V_A \simeq \bigoplus_{n_{ir}>s} \frac{(x-\lambda_r)^s k[x]}{(x-\lambda_r)^{n_{ir}} k[x]} \oplus \bigoplus_{i,j\neq r} \frac{k[x]}{(x-\lambda_j)^{n_{ij}} k[x]};$$

and so

$$\begin{aligned} \frac{(x-\lambda_r)^s V_A}{(x-\lambda_r)^{s+1} V_A} &\simeq \bigoplus_{n_{ir}>s} \frac{(x-\lambda_r)^s k[x]/(x-\lambda_r)^{n_{ir}} k[x]}{(x-\lambda_r)^{s+1} k[x]/(x-\lambda_r)^{n_{ir}} k[x]} \\ &\simeq \bigoplus_{n_{ir}>s} \frac{(x-\lambda_r)^s k[x]}{(x-\lambda_r)^{s+1} k[x]} \simeq \bigoplus_{n_{ir}>s} \frac{k[x]}{(x-\lambda_r) k[x]} \qquad (5) \end{aligned}$$

(To see why the last isomorphism hold, consider

$$k[x] \xrightarrow{\theta} \frac{(x-\lambda_r)^s k[x]}{(x-\lambda_r)^{s+1} k[x]}, \theta(p) := (x-\lambda_r)^s p + \langle (x-\lambda_r)^{s+1} \rangle;$$

it is easy to see that θ is a surjective $k[x]$-module homomorphism and its kernel is $(x-\lambda_r)k[x]$; thus by the first isomorphism theorem claim follows.)

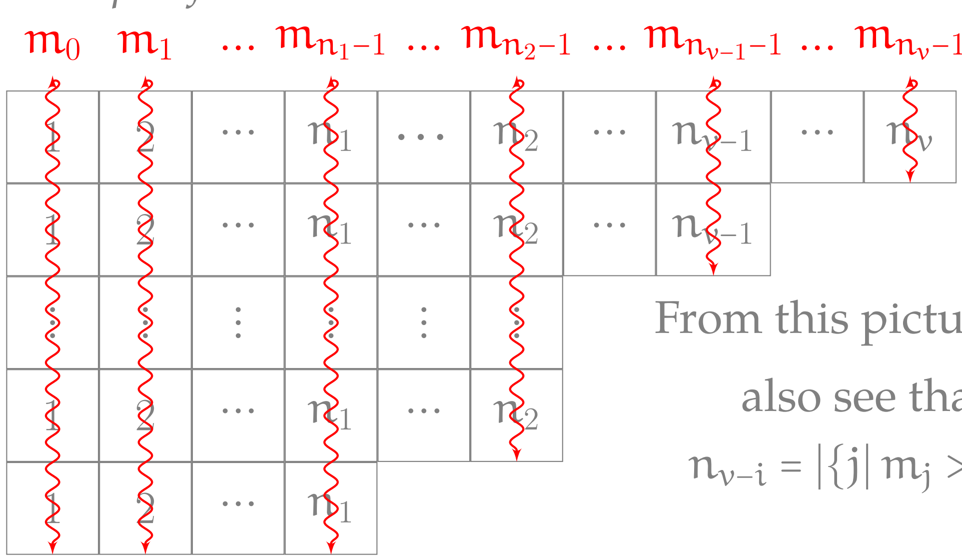
We also know that $\frac{k[x]}{(x-\lambda_r)k[x]} \simeq V_{[\lambda_r]}$; and so $\dim_k \frac{k[x]}{(x-\lambda_r)k[x]} = 1$. Therefore by (5) we deduce

$$\dim_k \frac{(x-\lambda_r)^s V_A}{(x-\lambda_r)^{s+1} V_A} = |\{i | \, n_{ir} > s\}|. \qquad (6)$$

The above equation implies that $|\{i | \, n_{ir} > s\}|$ only depends on the module structure of V_A; and so they just depend on the similarity class of A. Next we observe that the sequence $\{|\{i|\, n_{ir} > s\}|\}_s$ uniquely determines $\{n_{ir}\}_i$; and the uniqueness of Jordan form follows. This part of argument is identical to what we have done in the proof of rational canonical form theorem. For a possible future use we write it as a separate lemma.

Lemma 6 *Suppose $n_1 \le n_2 \le \cdots \le n_\nu$ is an increasing sequence of positive integers. Let $m_s := \{i|\, n_i > s\}$ for non-negative integers s. Then $\{n_i\}$ is uniquely determined by $\{m_s\}_s$.*

Pictorial proof.



From this picture we also see that $n_{\nu-i} = |\{j|\, m_j > i\}|$.

[1] ■ ■

[1]Special thanks go to B. Touri for teaching me how to create this picture!

Simple modules

Now that we have seen how important and instrumental module theory is, we try to study them a bit more systematically. As in group theory, we can start with *simplest* A-modules and try to build all the modules out of them.

Definition 7 *We say an A-modules M is a simple A-module if 0 and M are its only submodules and $M \neq 0$.*

Lemma 8 *(a) Suppose M_1 and M_2 are simple A-modules. Then $\mathrm{Hom}_A(M_1, M_2) \neq 0$ if and only if $M_1 \simeq M_2$.*

(b) (Schur's lemma) Suppose M is a simple A-module. Then $\mathrm{End}_A(M)$ is a division ring.

Proof (a) ($\Rightarrow$) Suppose $\theta \in \mathrm{Hom}_A(M_1, M_2)$. Then $\ker\theta$ is a submodule of M_1. Since M_1 is a simple A-module, $\ker\theta$ is either 0 or M_1. As θ is not zero, we deduce that $\ker\theta = 0$; and so θ is injective. We also know that $\mathrm{Im}\theta$ is a submodule of M_2. Since M_2 is a simple A-module, $\mathrm{Im}\theta$ is either 0 or M_2. Since θ is not zero, we deduce that θ is surjective. Overall we get that θ is a bijective A-module homomorphism; and so it is an isomorphism, which implies that $M_1 \simeq M_2$.

($\Leftarrow$) If $\theta : M_1 \to M_2$ is an isomorphism, then $\theta \neq 0$ (M_i's are not zero) and $\theta \in \mathrm{Hom}_A(M_1, M_2)$.

(b) Suppose $\theta \in \mathrm{End}_A(M) \setminus \{0\}$. By the above argument θ is an isomorphism; and so $\theta^{-1} \in \mathrm{End}_A(M)$; and claim follows. ■

Later we will see how this helps us to detect submodules of a *completely reducible* module $\bigoplus_{i=1}^n M_i$ that are isomorphic to a given simple A-module M. This is an important tool in the proof of Artin-Wedderburn theorem.