Math200b, lecture 7

Golsefidy

Finitely generated modules over a PID

In the previous lecture we proved the fundamental theorem of finitely generated modules over a PID; here is the extended version of it which we proved.

Theorem 1 *Suppose* D *is a PID and* M *is a finitely generated* D-*module. Then there are* $d_1|d_2|\cdots|d_m$ *in* $D\setminus\{0\}$ *such that*

$$M \simeq D^n \oplus \bigoplus_{i=1}^{m} D/\langle d_i\rangle;$$

and n *and the proper non-zero principal ideals* $\langle d_i\rangle$*'s are unique with these properties. Moreover* $\operatorname{rank}(M) = n$*, under the above isomorphism we have* $\operatorname{Tor}(M) \simeq \bigoplus_{i=1}^{m} D/d_iD$*, and* $\forall \mathfrak{m} \in \operatorname{Max}(D)$

and non-negative integer k

$$\dim_{D/\mathfrak{m}} \frac{\mathfrak{m}^k M_\mathfrak{m}}{\mathfrak{m}^{k+1} M_\mathfrak{m}} = |\{i | \nu_\mathfrak{m}(d_i) > k\}|. \quad (1)$$

Note: in class we used irreducible elements to describe (1); since in a PID a non-zero maximal ideal is generated by an irreducible element and vice versa an irreducible element generates a maximal non-zero ideal, these approaches are essentially the same.

We continue with two immediate corollaries of this result. Let's recall that when D is an integral domain and M is a D-module

$$\mathrm{Tor}(M) := \{m \in M | \exists d \in D \setminus \{0\}, dm = 0\}$$

is a submodule of M. A D-module is called torsion free if $dm = 0$ implies either $d = 0$ or $m = 0$. It is clear that a D-module is torsion free if and only if $\mathrm{Tor}(M) = 0$. And a free D-module is torsion free. As a corollary of the fundamental theorem of finitely generated modules over a PID we get the converse.

Corollary 2 *Suppose* D *is a PID and* M *is a finitely generated* D*-module. Then* M *is a free module if and only if* M *is torsion free.*

Proof. ($\Rightarrow$) is clear. To show ($\Leftarrow$), we notice that by Fundamental Theorem of finitely generated of modules over a PID, $M \simeq D^n \oplus \bigoplus_{i=1}^m D/\langle d_i \rangle$ (for some $d_i \in D \setminus (\{0\} \cup D^\times)$) and under this isomorphism $\mathrm{Tor}(M) \simeq \bigoplus_{i=1}^m D/d_i D$. Since M is torsion free, $\mathrm{Tor}(M) = 0$; so $M \simeq D^n$; and claim follows. ■

Note. In the above corollary it is crucial that M is finitely generated. For instance $\mathbb{Q}$ is a torsion free $\mathbb{Z}$-module; but it is not a free $\mathbb{Z}$-module. Let's see why $\mathbb{Q}$ is not a free $\mathbb{Z}$-module.

Method 1. Any two elements of $\mathbb{Q}$ are $\mathbb{Z}$-linearly dependent; and so $\mathrm{rank}(\mathbb{Q}) = 1$. Hence if $\mathbb{Q}$ is a free $\mathbb{Z}$-module, then it should be isomorphic to $\mathbb{Z}$; that means it should be a cyclic abelian group which is a contradiction. (Why?)

Method 2. Let $\theta : \mathbb{Q} \to \bigoplus_{i\in I} \mathbb{Z}$ be a $\mathbb{Z}$-module homomorphism. Then, for any $n \in \mathbb{Z} \setminus \{0\}$ and $a \in \mathbb{Q}$, $nx = \theta(a)$ has a solution in $\bigoplus_{i\in I} \mathbb{Z}$ (let $x := \theta(a/n)$); this implies that all the coordinates of $\theta(a)$ are multiples of n for any non-zero integer n. Hence $\theta(a) = 0$. So $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \bigoplus_{i\in I} \mathbb{Z}) = 0$; in particular $\mathbb{Q}$ is not a free $\mathbb{Z}$-module. (In class we used the above argument only for $n = 2$ and deduced that θ cannot be surjective; as you can see here the same idea gives us much more. And as it was mentioned during lecture the same statement holds for the so called divisible groups (these are groups were $nx = a$ has a solution for any non-zero integer n and $a \in G$.))

Minimum number of generators. How can we find the minimum number $d(M)$ of generators of a finitely generated D-module M when D is a PID? Suppose

$$M \simeq D^n \oplus \bigoplus_{i=1}^{m} D/\langle d_i \rangle$$

for $d_1|d_2|\cdots|d_m \in D \setminus (\{0\} \cup D^\times)$. Suppose p is an irreducible factor of d_1. As $d_1|d_i$, $p|d_i$ for any i. Therefore

$$pM \simeq (pD)^n \oplus \bigoplus_{i=1}^{m} (pD/d_i D);$$

and so

$$\frac{M}{pM} \simeq \left(\frac{D}{pD}\right)^n \oplus \bigoplus_{i=1}^{m} \frac{(D/d_i D)}{(pD/d_i D)} \simeq \left(\frac{D}{pD}\right)^{n+m}.$$

Since D/pD is a field, by linear algebra we know that

$$\underbrace{d(M/pM)}_{\text{as a } D/pD\text{-module}} = \dim_{D/pD} M/pM = m + n.$$

Since the scalar multiplication of M/pM as a D/pD-module is the same as the scalar multiplication of M/PM as a D-module, we have that the minimum numbers of generators of M/pM as a D-module and as a D/pD-module are the same.

We also notice that, if X generates a module M, then the image of X under the quotient map $M \to M/N$, for any submodule N, generates M/N; and so $d(M) \ge d(M/N)$ for any submodule N. Hence

$$d(M) \ge d(M/pM) = m + n.$$

Conversely $\{e_i\}_{i=1}^{m+n}$ generates $(pD)^n \oplus \bigoplus_{i=1}^{m}(pD/d_iD)$ as a D-module where e_i has 1 in its i-th coordinate and 0 in the rest (here 1 refers either to the identity of D or the identity of D/d_jD for some j); and so $d(M) \le m + n$. So overall we get for $d_1|d_2|\cdots|d_m \in D \setminus (\{0\} \cup D^\times)$ and $M = D^n \oplus \bigoplus_{i=1}^{m} D/\langle d_i \rangle$, we have

$$\mathrm{rank}(M) = n, \text{ and } d(M) = m + n;$$

in particular, $d(M) = \mathrm{rank}(M)$ if and only if M is a free D-module. In your HW assignment you will prove that the same statement holds for an arbitrary unital commutative ring D.

Rational canonical form. Suppose k is a field and $A \in M_n(k)$. As we discussed earlier in the course, we can view k^n

as a $k[x]$-module by defining $x \cdot v := Av$; more explicitly for any polynomial $f(x) := \sum_{i=0}^{\infty} a_i x^i$ and $v \in k^n$ we have

$$f(x) \cdot v := \sum_{i=0}^{\infty} a_i A^i v.$$

In order to remember that the above $k[x]$-module structure of k^n depends on A we denote it by V_A. The first question that we address is how much the module structure of V_A depends on A.

Lemma 3 $V_A \simeq V_B$ *if and only if* A *and* B *are similar; that means* $\exists g \in \mathrm{GL}_n(k)$ *such that* $A = g^{-1}Bg$.

Proof. $(\Rightarrow)$ Suppose $\theta : V_A \to V_B$ is a $k[x]$-module isomorphism. Then for any $v \in k^n$ we have

$$\theta(\underbrace{x \cdot v}_{\text{in } V_A}) = \underbrace{x \cdot \theta(v)}_{\text{in } V_B};$$

and so $\theta(Av) = B\theta(v)$. Suppose $g \in M_n(k)$ is the matrix representation of θ in the standard basis; that means for any $v \in k^n$ (in a column form) we have $\theta(v) = gv$. Since θ is a bijection, $g \in \mathrm{GL}_n(k)$. Therefore we have that for any $v \in k^n$

$$gAv = \theta(Av) = B\theta(v) = Bgv, \text{ which implies } A = g^{-1}Bg.$$

($\Leftarrow$) Let $\theta : k^n \to k^n, \theta(v) := gv$. Then for any $v \in V_A$ and $f(x) := \sum_{i=0}^{\infty} c_i x^i \in k[x]$ we have

$$\begin{aligned}\theta(\underbrace{f(x)\cdot v}_{\text{in } V_A}) &= g\left(\sum_{i=0}^{\infty} c_i A^i\right) v = g\left(\sum_{i=0}^{\infty} c_i (g^{-1}Bg)^i\right) v \\ &= g\left(\sum_{i=0}^{\infty} c_i g^{-1} B^i g\right) v = \left(\sum_{i=0}^{\infty} c_i B^i\right) \underbrace{gv}_{\theta(v)} \\ &= \underbrace{f(x)\cdot\theta(v)}_{\text{in } V_B};\end{aligned}$$

and so $\theta : V_A \to V_B$ is a $k[x]$-module isomorphism. ∎

Next we notice that $\mathrm{rank}(V_A) = 0$ as a $k[x]$-module; and so:

Proposition 4 *There are unique monic positive degree polynomials* $f_1 | f_2 | \cdots | f_m \in k[x]$ *such that*

$$V_A \simeq k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle$$

as $k[x]$*-modules.*

Proof. If $\mathrm{rank}(V_A) \neq 0$, then $k[x]$ can be embedded into V_A (as a $k[x]$-module); this implies that $\dim_k V_A = \infty$, which is a contradiction. Since $\dim_k V_A = n < \infty$, V_A is a finitely generated

$k[x]$-module. Since $k[x]$ is a PID, by Theorem 1 and having $\operatorname{rank}(V_A) = 0$ we deduce that there are polynomials $f_1|f_2|\cdots|f_m \in k[x] \setminus (\{0\} \cup k[x]^\times)$ such that

$$V_A \simeq k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle$$

as $k[x]$-module. As $k[x]^\times = k^\times$, we have that $\deg f_i \geq 1$; and after multiplying f_i's by some units we can assume that f_i's are monic. And uniqueness follows from the uniqueness part of Theorem 1 and the fact that two different monic polynomials cannot generate the same principal ideal. ■

Next we would like to see if $k[x]/\langle f(x)\rangle$ is isomorphic to V_C for some $C \in M_n(k)$.

Lemma 5 *Suppose* $f(x) \in k[x]$ *and* $\deg f = n > 0$. *Then* $\{\overline{1}, \ldots, \overline{x^{n-1}}\}$ *is a* k*-basis of* $k[x]/\langle f(x)\rangle$, *where* $\overline{x^i} := x^i + \langle f(x)\rangle$; *in particular* $\dim_k k[x]/\langle f(x)\rangle = \deg f$.

Proof. This is an immediate corollary of long division: for $a(x) \in k[x]$, let $q(x)$ and $r(x)$ be the quotient and remainder of $a(x)$ divided by $f(x)$, respectively. So $a(x) = f(x)q(x) + r(x)$ and $r(x) = \sum_{i=0}^{n-1} c_i x^i$; thus

$$a(x) + \langle f(x)\rangle = r(x) + \langle f(x)\rangle = \sum_{i=0}^{n-1} c_i \overline{x^i}.$$

This implies that the k-span of $\{\overline{x^i}\}_{i=0}^{n-1}$ is $k[x]/\langle f(x)\rangle$.

If $\sum_{i=0}^{n-1} c_i\overline{x^i} = 0$, then $\sum_{i=0}^{n-1} c_i x^i \in \langle f(x)\rangle$. Since the only multiple of $f(x)$ that has degree less than degree of $f(x)$ is zero, we get that $\sum_{i=0}^{n-1} c_i x^i$; this implies that $c_i = 0$ for any i, which means $\{\overline{x^i}\}_{i=0}^{n-1}$ consists of k-linearly independent elements. ∎

To find $C \in M_n(k)$ in a way that $k[x]/\langle f(x)\rangle$ becomes isomorphic to V_C as a $k[x]$-module, we have to focus on the k-linear map of multiplying by x:

$$k[x]/\langle f(x)\rangle \xrightarrow{\times x} k[x]/\langle f(x)\rangle;$$

(recall that in V_C multiplication by x is given by C). We will write down the matrix representation of $\times x$ in the basis $\{\overline{x^i}\}_{i=0}^{n-1}$: to find the i-th column we have to multiply $\overline{x^i}$ by x and then write it as a linear combination of elements of $\{\overline{x^i}\}_{i=0}^{n-1}$. So for $0 \le i < n-1$, we simply have $x \cdot \overline{x^i} = \overline{x^{i+1}}$, and for the last column we have

$$x \cdot \overline{x^{n-1}} = \overline{x^n} = -c_0 - c_1\overline{x} - \cdots - c_{n-1}\overline{x^{n-1}}$$

where $f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_0$. Hence the associated matrix is

$$c(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}.$$

$c(f)$ is called the <u>companion matrix</u> of f. Overall we proved:

Proposition 6 *Suppose* $f(x) \in k[x]$ *is a monic positive degree polynomial. Then*

$$k[x]/\langle f(x) \rangle \simeq V_{c(f)}$$

as $k[x]$*-module.*

Theorem 7 (Rational canonical form) *Suppose* k *is a field and* $A \in M_n(k)$*. Then there are unique monic positive degree polynomials* $f_1 | f_2 | \cdots | f_m \in k[x]$ *such that* A *is similar to*

$$\begin{pmatrix} c(f_1) & 0 & \cdots & 0 \\ 0 & c(f_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c(f_m) \end{pmatrix}.$$

Proof. (Existence) By Proposition 4 there are unique monic positive degree polynomials $f_1|f_2|\cdots|f_m \in k[x]$ such that

$$\begin{aligned} V_A &\simeq k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle \\ &\simeq V_{c(f_1)} \oplus \cdots \oplus V_{c(f_m)} \qquad \text{(by Proposition 6)} \\ &\simeq V_{\mathrm{diag}(c(f_1),\dots,c(f_m))}; \end{aligned}$$

and so, by Lemma 3, A and $\mathrm{diag}(c(f_1),\dots,c(f_m))$ are similar which gives us the existence part.

(Uniqueness) Suppose A is similar to $\mathrm{diag}(c(f_1),\dots,c(f_m))$ and $\mathrm{diag}(c(f'_1),\dots,c(f'_{m'}))$ for some monic positive degree polynomials $f_1|f_2|\cdots|f_m$ and $f'_1|f'_2|\cdots|f'_{m'}$. So by Lemma 3

$$V_{\mathrm{diag}(c(f_1),\dots,c(f_m))} \simeq V_{\mathrm{diag}(c(f'_1),\dots,c(f'_{m'}))};$$

therefore

$$V_{c(f_1)} \oplus \cdots \oplus V_{c(f_m)} \simeq V_{c(f'_1)} \oplus \cdots \oplus V_{c(f'_{m'})}.$$

Hence by Proposition 6

$$k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle \simeq k[x]/\langle f'_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f'_{m'}(x)\rangle,$$

as $k[x]$-modules. Therefore by the uniqueness part of Theorem 1 claim follows. ∎

The word rational refers to the fact that this form works for an arbitrary field.

The monic polynomials $f_1, \ldots, f_m$ in Theorem 7 (Rational Canonical Form) are called invariant factors of A. Notice that if f_i's are invariant factors of A, then

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

Let's recall that the characteristic polynomial $f_A(x)$ of A is

$$f_A(x) := \det(xI - A).$$

A monic polynomial $m_A(x)$ is called the minimal polynomial of A if $m_A(A) = 0$ and $p(A) = 0$ for $p(x) \in k[x]$ implies that $m_A(x)|p(x)$; alternatively $m_A(x)$ is a monic polynomial such that

$$\langle m_A(x)\rangle = \{p(x) \in k[x] | \, p(A) = 0\}.$$

Note that one can easily check that the RHS of the above equality is an ideal of $k[x]$; and so there is such $m_A(x)$. We will show the existence by proving that f_m (the last invariant factor) satisfies these properties. Convince yourself that if there is a minimal polynomial it is unique.

Theorem 8 (Invariant factors and minimal polynomial) *Suppose* $f_1|f_2|\cdots|f_m$ *are invariant factors of* A. *Then* f_m *is the minimal polynomial of* A.

Proof. We know that

$$V_A \simeq k[x]/\langle f_1(x)\rangle \oplus \cdots \oplus k[x]/\langle f_m(x)\rangle,$$

as $k[x]$-module. Since $f_i(x)|f_m(x)$, $f_m(x)$ times the RHS is zero. Hence $f_m(x)\cdot V_A = 0$, which means $f_m(A)k^n = 0$. Therefore $f_m(A) = 0$.

Suppose $p(A) = 0$. Then for any $v \in k^n$ we have $p(A)v = 0$; and so $\underbrace{p(x)\cdot v}_{\text{in } V_A} = 0$. Therefore $p(x)$times the RHS is zero; in particular

$$p(x)\left(k[x]/\langle f_m(x)\rangle\right) = 0.$$

This implies that $f_m(x)|p(x)$. And claim follows. ∎

Next we prove a stronger result which implies Cayley-Hamilton Theorem.

Theorem 9 (Cayley-Hamilton Theorem) $f_A(A) = 0$ *where* f_A *is the characteristic polynomial of* A.

To prove Cayley-Hamilton Theorem it is enough to show that $m_A(x)|f_A(x)$. So Cayley-Hamilton Theorem is a corollary of the next theorem.

Theorem 10 *Suppose* $f_1|f_2|\cdots|f_m$ *are invariant factors of* A. *Then* $m_A(x) = f_m(x)$ *and* $f_A(x) = f_1(x)f_2(x)\cdots f_m(x)$; *in particular,* $m_A(x)|f_A(x)$ *and any irreducible factor of* $f_A(x)$ *is an irreducible factor of* $m_A(x)$.

Proof. By Rational Canonical Form theorem, there is g in $\mathrm{GL}_n(k)$ such that

$$g^{-1}Ag = \mathrm{diag}(c(f_1), \cdots, c(f_m)).$$

Hence

$$g^{-1}(xI - A)g = xI - g^{-1}Ag = \mathrm{diag}(xI - c(f_1), \cdots, xI - c(f_m)).$$

Therefore

$$f_A(x) = \det(xI - A) = \prod_{i=1}^{m} \det(xI - c(f_i)) = \prod_{i=1}^{m} f_{c(f_i)}(x).$$

In the next Lemma we will prove that $f_{c(f)}(x) = f(x)$ for a monic positive degree polynomial; for now we will assume this and continue the proof. And so we have

$$f_A(x) = f_1(x)f_2(x)\cdots f_m(x).$$

By Theorem 8 we know that $m_A(x) = f_m(x)$; and so $m_A(x)|f_A(x)$.

Now suppose $p(x)$ is an irreducible factor of $f_A(x)$; then $p(x)$ divides $\prod_{i=1}^{m} f_i(x)$. Since $p(x)$ is prime ($k[x]$ is a PID), $p(x)$ divides $f_i(x)$ for some i. As $f_i(x)|f_m(x)$ for any i, $p(x)$ divides $f_m(x) = m_A(x)$; and claim follows. ∎

In the next lecture we will prove $f_{c(f)} = f(x)$ and the Jordan Canonical Form; and then we get to the more general theory of modules.