MATH 593 - Linear Algebra on a Ring

AResseges Stery

October 31, 2023

Contents

1 Linear Transformations on a Ring

Recall the two versions of Structural Theorem of finitely generated modules over PID:

Theorem 1.1 (Sturcture, v1). *Let M be a finitely generated R-module, with R a PID. Then there exists $a_1, \dots, a_m \in R \setminus \{0\}$ s.t. $a_1 \mid a_2 \mid \cdots \mid a_m$.*

Theorem 1.2 (Structure, v2). *Let M be a finitely generated R-module, with R a PID. Then there exists primes p_is and $n_i \in \mathbb{N}$ s.t. $M \simeq R^r \oplus R/(p_1^{n_1}) \oplus \cdots \oplus R/(p_r^{n_r})$.*

Example 1.1. Let $R = \mathbb{Z}$, with M an R-module. Since defining scalar multiplication on modules is equivalent to defining maps from R to endomorphisms on M, it is sufficient for M to be an abelian group. By Structural Theorem, there exists p_is and m_is s.t. $M \simeq \mathbb{Z}^r \oplus \mathbb{Z}/(p_1^{m_1}) \oplus \cdots \oplus \mathbb{Z}/(p_k^{m_k})$. M is finite (as a group) if and only if $r = 0$.

The direct sum allows describing it with a basis, which gives a generalization of linear algebra defined on ring (module) structure.

Parenthesis 1.1. Linear maps between elements in free modules can be represented as invertible matrices.

Proof. The reasoning is similar to that under the context of vector spaces. Fix R to be a commutative ring, with M a finitely generated R-module. Let $n = \operatorname{rank}(M)$. Then $M \simeq R^n$. Choose $B = (e_1, \dots, e_n)$ to be a basis of M.

For all $u \in M$, there exists a unique decomposition of u into the basis, i.e. there exists $a_1, \dots, a_n$ s.t. $u = \sum_{k=1}^n a_k e_k$. Denote $M_B(u) = (a_1, \dots, a_k)^T$.

Now consider change of basis. Suppose that $B' = (e_1', \dots, e_n')$ is another basis of M. There exists b_{ik}s s.t. $e_i = \sum_{k=1}^n b_{ik} e_k'$; and there exists c_{ik}s s.t. $e_i' = \sum_{k=1}^n b_{ik} e_k$. Apply the substitution twice gives

$$e_i = \sum_{j=1}^n b_{ij} e_j' = \sum_{j=1}^n b_{ij} \left(\sum_{k=1}^n c_{jk} e_k \right) = \sum_{k=1}^n \sum_{j=1}^n (b_{ij} c_{jk}) e_k \implies \left(\sum_{k=1}^n \sum_{j=1}^n (b_{ij} c_{jk}) e_k \right) - e_i = 0$$

Since e_is give a basis, this implies that $\sum_{j=1}^n (b_{ij} c_{jk}) = \delta_{ik}$. Let $V = (b_{ij}) \in M_n(R)$ to be the transition matrix from B to B', abd $U = (c_{ij}) \in M_n(R)$ the transition matrix from B' to B. Conducting this concurrently gives $UV = \mathrm{Id}_B$. Similarly $VU = \mathrm{Id}_{B'}$. □

Proposition 1.1. *The converse of the above also holds, i.e. If (c_{kl}) is invertible in $M_n(R)$, then for $e_k' = \sum_{l=1}^n c_{kl} e_l$, e_k's also give a basis.*

Proof. It suffices to verify that e_k's are R-linearly independent, and they span the whole module:

- If there exists λ_is that are not all zero, that $\sum_{i=1}^n \lambda_i e_i' = 0$, then $\sum_{i=1}^n \lambda_i \sum_{k=1}^n c_{ik} e_k = 0$ which implies that e_k are not R-linearly independent, which is a contradiction.
- Since (c_{kl}) is invertible, there exists some (b_{kl}) s.t. $e_k = \sum_{l=1}^n b_{kl} e_l$. Then, for all $u \in M$ with decomposition into the original basis $u = \sum_{i=1}^n u_i e_i$, there exists a decomposition into e_k's: $u = \sum_{i=1}^n u_i \sum_{j=1}^n b_{ij} e_j$.

□

Remark 1.1. The transition matrix is compatible with representation of an element in the basis. Let $M \ni u = \sum_{i=1}^{n} u_i e_i$, with $U = (b_{ij})$ the transition matrix from $B = (e_i)$ to $B' = (e'_i)$. Then

$$u = \sum_{i=1}^{n} u_i e_i = \sum_{i=1}^{n} \left(u_i \sum_{j=1}^{n} b_{ij} e'_i \right) \implies M_{B'}(u) = U \cdot M_B(u)$$

Remark 1.2. Using such formalization the operations are represented in the identical way as that in vector spaces:

1. *Applying a linear map.* If $T : F \to G$ is not an endomorphism and T is specified via specifying the image of the basis $T(e_j) = \sum_{i=1}^{n} a_{ij} f_j$, where F and G are finitely generated free R-modules; and $B_F = (e_i), B_G = (f_i)$ give a basis in the corresponding module. Then the matrix representation of T under such bases is $M_{B_F B_G}(T) = (a_{ij})$. It acts in the same way as matrices acting on vectors, as for $M_{B_F}(u) = (b_1, \ldots, b_u)^T$

$$\begin{aligned} T(u) = T\left(\sum_{j=1}^{n} b_j e_j \right) = \sum_{j=1}^{n} b_j T(e_j) = \sum_{j=1}^{n} b_j \left(\sum_{i=1}^{n} a_{ij} f_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} b_j) f_j \\ \implies M_{B_G}(T(u)) = M_{B_F B_G}(T) \cdot M_{B_F}(u) \end{aligned}$$

2. *Composition of linear maps.* Consider $T : F \to G$ and $S : G \to H$ where $B_F = (e_i), B_G = (f_i)$ and $B_H = (g_i)$. To specify the linear maps, it suffices to specify where the elements of the basis is mapped to. Suppose that $T(e_i) = \sum_{j=1}^{n} a_{ji} f_j; S(f_i) = \sum_{j=1}^{n} b_{ji} h_j$. For $F \ni u = \sum_{i=1}^{n} u_i e_i$, considering $g \circ f$ gives

$$\begin{aligned} (S \circ T)(u) = (S \circ T) \left(\sum_{i=1}^{n} u_i e_i \right) = S \left(\sum_{i=1}^{n} u_i \sum_{j=1}^{n} a_{ji} f_j \right) = \sum_{i=1}^{n} u_i \sum_{j=1}^{n} a_{ji} S(f_j) \\ = \sum_{i=1}^{n} u_i \sum_{j=1}^{n} a_{ji} \sum_{k=1}^{n} b_{kj} h_j = \sum_{i=1}^{n} u_i \sum_{j=1}^{n} \left(\sum_{k=1}^{n} (a_{ji} b_{kj}) h_j \right) \\ \implies M_{B_F B_H}(S \circ T) = B_{B_G B_H}(S) \cdot M_{B_F B_G}(T) \end{aligned}$$

   where the elements of $M_{B_F B_H}(S \circ T)$ is specified by $\sum_{j=1}^{n} \sum_{k=1}^{n} (a_{ji} b_{kj})$.

3. *Change of basis.* Now consider change of basis under the context of a linear transformation. Let $T : F \to G$ be an R-linear map, with $M_{B_F B_G}(T)$ the matrix representation of T under B_F and B_G. Now consider change of basis maps $U : B_F \to B'_F$ and $\tilde{U} : B_G \to B'_G$. We are interested in the corresponding map $\tilde{T}$ of T after applying the change of basis:

$$\begin{array}{ccc} B'_F & \xrightarrow{\tilde{T}} & B'_G \\ {\scriptstyle U}\uparrow & & \uparrow{\scriptstyle \tilde{U}} \\ B_F & \xrightarrow{T} & B_G \end{array}$$

   As proven above it is valid to express linear transformation and change of basis using matrices, and matrices corresponding to change of basis are invertible, we have for $u \in F$,

$$M_{B'_G}(T(u)) = M_{B_G B'_G}(\tilde{U}) M_{B_F B_G}(T) M_{B'_F B_F}(U) = M_{B_G B'_G}(\tilde{U}) M_{B_F B_G}(T) (M_{B_F B'_F}(U))^{-1}$$

4. *Change of basis on endomorphisms.* Then the equality above becomes

$$M_{B'_G}(T(u)) = M_{B_G B'_G}(U) M_{B_G}(T) M_{B'_G B_G}(U) = (M_{B'_G B_G}(U))^{-1} M_{B_G}(T) M_{B'_G B_G}(U)$$

which is exactly the conjugate of a matrix.

Definition 1.1. *Two matrices A and B in $M_n(R)$ are **similar** if there exists some invertible $U \in M_n(R)$ s.t. $A = U^{-1}BU$. Two R-linear maps T and $T' : F \to F'$ are **similar** if there exists some isomorphism φ s.t. $T' = \varphi^{-1} T \varphi$.*

Remark 1.3. Similarity is an equivalence relation, with $(A = U^{-1}BU) \wedge (B = V^{-1}CV) \implies A = (VU)^{-1}C(VU)$ for transitivity.

R-linear maps are similar to each other if and only if the corresponding matrix is similar, as on free modules linear maps can be represented by matrices.

Proposition 1.2. *There exists a canonical bijection between:*

$$\{R\text{-linear endomorphisms } F \to F\}/\text{similarity} \simeq M_n(R)/\text{similarity}$$

Proof. Choose B_F to be a basis of F. First verify that the map is bijective: as is formalized above since F is free, with a fixed basis linear transformations could be represented via matrices to indicate how the basis is transformed. Therefore, for any linear transformation there exists one matrix to represent it under B_F and vice versa. Further the choice of M_n is unique as matrices under different basis are conjugate w.r.t. the change of basis matrix. □

Remark 1.4. The bijection will still be valid without the quotient. However this will cease to be canonical as the map differs by the choice of basis on which the matrix conducts the representation.

2 Rational and Smith Normal Form

The main idea of this section is to classify (and represent) linear transformations up to similarity. The construction seeks to embed a specific map into the module structure.

Let k be a field, with V a finite-dimensional k-vector space. Let T be an endomorphism of V. Then (V, T) could be viewed as an $k[x]$-module via specifying that $xu := T(u)$ for all $u \in V$.

Remark 2.1. Since the only difference in module structure introduced by (V, T) from V is the application of T when multiplying by x, submodules are preserved as long as it is closed w.r.t. T. That is, for all $W \subseteq V$ k-vector subspaces, $W \subseteq V$ is a $k[x]$-submodule as long as $T(W) \subseteq W$.

Proposition 2.1. *$(V, T) \simeq (V, T')$ if and only if T and T' are similar.*

Proof. Proceed via showing implication in both directions:

$\Rightarrow$ Suppose that there exists isomorphism φ from (V, T) to (V, T'). Then for $u \in V$ consider

$$\varphi(T(u)) = \varphi(xu) = x\varphi(u) = T'(\varphi(u)) \implies T(u) = \varphi^{-1}(T(\varphi(u)))$$

which implies that T and T' are similar.

$\Leftarrow$ If T and T' are similar, there exists some isomorphism φ s.t. $T' = \varphi^{-1} \circ T \circ \varphi$. Then φ gives an isomorphism from (V, T) to (V, T') with the same process as above.

□

Since $k[x]$ is a PID, applying Structural Theorem gives $V \simeq k[x]^n \oplus k[x]/(f_1) \oplus \cdots \oplus k[x]/(f_r)$. Since $k[x]$ is of infinite dimension if viewed as a k-vector space, $n = 0$, which gives

$$V \simeq k[x]/(f_1) \oplus \cdots \oplus k[x]/(f_r), \quad f_1 \mid \cdots \mid f_r \tag{$*$}$$

To make the representation canonical, fix f_is to be monic, i.e. the leading coefficients for f_is are 1 for all i.

Now consider the matrix representation of applying T on V. It is sufficient to consider the situation under $k[x]/(f_i)$, as V is isomorphic to the direct sum of some copies of this, which only differs in f_is chosen. This gives the following definition:

Definition 2.1 (Companion Matrix)**.** *Let $f = a_0 + a_1 x + \ldots + a_{d-1}x^{d-1} + x^d$. Then multiplication by x (application of T) is represented as*

$$\mathcal{C}_{f_i} = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & \ddots & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{d-1} \end{pmatrix}$$

*which is the **companion matrix** of f.*

Definition 2.2 (Rational Canonical Form)**.** *A matrix $A \in M_n(k)$ is in **rational canonical form** if it is in the form of*

$$\begin{pmatrix} \mathcal{C}_{f_1} & & 0 \\ & \ddots & \\ 0 & & \mathcal{C}_{f_r} \end{pmatrix}, \qquad f_1 \mid \cdots \mid f_r, \quad f_i \text{ monic}, \quad \deg f_i \geq 1 \; \forall \, i$$

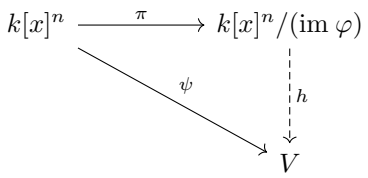
Then choosing basis to be $x^k \in k[x]/(f_i)$ s.t. $k < \deg f_i$ and appending the basis of the summands gives the matrix of T in rational canonical form. Note that this is a basis as a k-vector space, but is not one for V as a $k[x]$-module.

Remark 2.2. $T, T' \in \mathrm{End}_K V$ are similar if and only if they can be described in some basis in the rational canonical form:

Proof. By Remark 2.1 this implies that $(V, T) \simeq (V, T')$. Using Eq. $(*)$ gives that they have the same f_is in $(*)$, i.e. they have the same rational canonical form. □

Proposition 2.2. *Let V be a finite-dimensional k-vector space with basis $B_V = (u_1, \ldots, u_n)$. Let (V, T) be the extension into $k[x]$-module with $T \in \mathrm{End}_K(V)$, and A be the matrix representation of T using basis B_V. Consider $\varphi : k[x]^n \to k[x]^n$ s.t. it is represented by $(x\mathrm{Id}_n - A)$ in the canonical basis. Then $V \simeq k[x]^n/(\mathrm{im}\ \varphi)$; and φ is injective.*

Proof. Consider the following commutative diagram:



Specify $\psi : e_i \mapsto u_i \; \forall \, i$. To prove that $V \simeq k[x]^n/(\operatorname{im}\varphi)$ it suffices to verify that $\ker\psi = \operatorname{im}\varphi$. Proceed via showing inclusion in both directions:

$\supseteq$: It suffices to verify that for all j, $\psi(\varphi(e_i)) = 0$. Applying the definition of the maps gives

$$\psi(\varphi(e_i)) = \psi(xe_i - \sum_{j=1}^{n} a_{ij}e_j) = x\psi(e_i) - \sum_{j=1}^{n} a_{ij}\psi(e_j) = T(u_i) - T(u_i) = 0$$

$\subseteq$: Specify $h : \bar{e}_i \mapsto u_i$. $\ker\psi \supseteq \operatorname{im}\varphi$ gives that h is surjective. Notice that h is an isomorphism between $k[x]$-modules if and only if h is an isomorphism of k-vector spaces that are closed w.r.t. multiplication by x (or application of T). Therefore, to show that h is an isomorphism it suffices to show that the k-linear span of $\bar{e}_i$s is $k[x]^n/(\operatorname{im}\varphi) =: U$.

It suffices to verify that $x^m\bar{e}_i \in U$. Proceed via induction:

- *Base case.* e_i by definition is in the span of e_is for all i.
- *Inductive step.* Suppose that $x^k\bar{e}_i \in U$. Since φ is a map of free $k[x]$-modules, $U \simeq \ker\varphi$, which gives

$$\varphi(e_i) = (xI_n - A)e_i = xe_i - \sum_{j=1}^{n} a_{ij}e_j = x^m e_i - \sum_{j=1}^{n} x^{m-1}a_{ij}e_j = 0$$

i.e. $x^m e_i$ is spanned by $(x^{m-1}a_{ij}e_j)$s, which by inductive hypothesis are in k-linear span of e_is.

It remains to show that φ is injective. Since $\operatorname{rank}_{k[x]}(V) = 0$, by $V \simeq k[x]^n/(\operatorname{im}\varphi)$, $\operatorname{rank}(\operatorname{im}\varphi) = \operatorname{rank}(k[x]^n) = n$. But notice $k[x]^n/\ker\varphi \simeq \operatorname{im}\varphi$, i.e. $\operatorname{rankker}\varphi = \operatorname{rank}k[x]^n - \operatorname{rankim}\varphi = 0$. But $\ker\varphi$ as a submodule of $k[x]$ is free, which implies that $\ker\varphi = \{0\}$, i.e. φ is injective. □

The map $(x\mathrm{Id}_n - A)$ is important as it annihilates the torsion module of modules on $k[x]$, it reveals the information of f_is. Specifically, for each summand in the direct sum of $\bigoplus_i k[x]/(f_i)$, f_i is the minimal annihilator of the whole module.

Definition 2.3 (Smith Normal Form). *Let (V, T) be a $k[x]$-module where V is a finite-dimensional k-vector space, and $T \in \operatorname{End}_k(V)$. Let the matrix representation of T in a certain basis to be A. The **Smith Normal Form** of $(x\mathrm{Id}_n - A) =: M$, is a matrix in the form of*

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & f_1 & \\ & & & & \ddots \\ & & & & & f_r \end{pmatrix} \qquad f_1 \mid \cdots \mid f_r, \quad f_i \text{ monic}, \quad \deg f_i \geq 1 \; \forall \, i$$

s.t. it can be transformed from M via the following transformations:

1. *Swap the rows/columns of M.*
2. *Multiply a row/column of M by $\lambda \in k \setminus \{0\}$.*
3. *Add one row/column of M multiplied by $f \in k[x]$ to another row/column.*

The f_is are called elementary divisors, or invariant factors.

Remark 2.3. The valid operations allowed in making the transformation of M are the same as applying invertible transformation, with column/row operations corresponding to manipulation of basis in the source/image.

Multiplication of one row is only allowed up to non-zero elements in k as $k[x]$ is not a domain, where multiplication is not invertible.

Remark 2.4. It could be read off from the matrix that $\mathrm{Ann}_{k[x]}(V) = \{g \in k[x] : f_i \mid g \ \forall\, i\}$, as there are no 1 elements on the diagonal of the matrix.

3 Minimal and Characteristic Polynomials

Return to the case where V is a finite-dimensional k-vector space; and morphisms are considered in the context of $k[x]$-modules. Recall that the structure of V, given as a $k[x]$-module, is

$$V \simeq k[x]/(f_1) \oplus \cdots \oplus k[x]/(f_r), \quad f_1 \mid \cdots \mid f_r$$

Definition 3.1 (Minimal Polynomial)**.** *The **minimal polynomial** of $T \in \mathrm{End}_k(V)$ is the monic generator of $\mathrm{Ann}_{k[x]}(V)$, denoted to be $m_T(x)$. Expressing V in the form of Eq. $(*)$, or using Smith Normal Form, $m_T(x) = f_r$.*

Definition 3.2 (Characteristic Polynomial)**.** *Let $A \in M_n(k)$. The **characteristic polynomial** of A is given as $c_A(x) = \det(x\mathrm{Id}_n - A) \in k[x]$.*

Remark 3.1. Inheriting the notations in Smith Normal Form of V, $c_T(x) = f_1 \dots f_r$.

Remark 3.2. If $A \sim A'$, then $m_A(x) = m_{A'}(x)$, and $c_A(x) = c_{A'}(x)$; but the converse does not necessarily hold.

Remark 3.3. Since the minimal/characteristic polynomial is defined on the structure or invariant factors in Smith Normal Form, they are invariant w.r.t. change of basis.

Definition 3.3 (Eigenvalue)**.** *Given $T \in \mathrm{End}_k(V)$, λ is an **eigenvalue** of T if the following equivalent conditions are satisfied:*

- *There exists some $v \in V \setminus \{0\}$ s.t. $\lambda v = Tv$, i.e. $(\lambda\mathrm{Id} - T)v = 0$.*
- *$\det(\lambda\mathrm{Id} - T)v = 0$, i.e. $c_T(\lambda) = 0$.*

Proposition 3.1. *If $(V, T) \simeq \bigoplus_{i=1}^r k[x]/(f_i)$ where f_is are the invariant factors of T as a morphism of $k[x]$-modules. Then $c_T(x) = \Pi_{i=1}^r f_i$.*

Proof. There exists some basis in which V is in the rational canonical form. Then $c_T(x) = \Pi_{i=1}^r \det \mathcal{C}_{f_i}$. It then suffices to show that $\det \mathcal{C}_{f_i} = f_i$ for all i. Notice

$$\det \begin{pmatrix} x & & 0 & a_0 \\ -1 & \ddots & & a_1 \\ & \ddots & x & \vdots \\ 0 & & -1 & x + a_{d-1} \end{pmatrix} = x \begin{pmatrix} x & & 0 & a_1 \\ -1 & \ddots & & a_2 \\ & \ddots & x & \vdots \\ 0 & & -1 & x + a_{d-1} \end{pmatrix} + \underbrace{a_0 \cdot (-1)^{n+1} \cdot (-1)^{n-1}}_{a_0}$$

where, performing finitely many (d) steps recovers the full polynomial. □

Remark 3.4. This proposition is also a direct result of the existence of Smith Normal Form, as after finitely many elementary row/column operations which do not change the determinant, $(x\mathrm{Id}_n - A)$ has f_is and 1s on the diagonal; and the result is given via simply multiplying all of them.

The conclusion is that for $V \simeq k[x]/(f_1) \oplus \cdots \oplus k[x]/(f_r)$, $\quad f_1 \mid \cdots \mid f_r$ where $f_1 \mid \cdots \mid f_r$, $m_T = f_r$; and $c_T = f_1 \dots f_r$. This gives

$$m_T \mid c_T, \qquad c_T \mid m_T^r$$

Further, since f_r is the monic generator of $\mathrm{Ann}_{k[x]}(V)$, $m_T(T)(u) = f_r \cdot u = 0$ for all $u \in V$, i.e. $m_T(T) = 0$. Since $m_T \mid c_T$, $c_T(T) = 0$, which is the result from Cayley-Hamilton.

4 Jordan Normal Form