10. Canonical forms of linear transformations and similarity classes of matrices.

Let F be a field, V a vector space over F, $n = \dim V$ and $\mathfrak{gl}(V) = \mathrm{Hom}_F(V, V)$. Note that $\mathfrak{gl}(V) \cong Mat_n(F)$ as F-algebras because of the way matrix multiplication is defined.

To find an explicit isomorphism we choose an (ordered) basis $\Omega = \{e_1, \ldots, e_n\}$ of V. Given $T \in \mathfrak{gl}(V)$ we write $T(e_j) = \sum_{i=1}^{n} a_{ij} e_i$ for $1 \le j \le n$ and set

$$[T]_\Omega = (a_{ij}) \in Mat_n(F), \text{ called the \underline{matrix of T with respect to Ω.}}$$

Then the map $\mathfrak{gl}(V) \to Mat_n(F)$ given by $T \mapsto [T]_\Omega$ is an isomorphism of F-algebras.

If Ω' is another basis of V, the matrices $[T]_\Omega$ and $[T]_{\Omega'}$ are related by the equality

$$[T]_{\Omega'} = P^{-1}[T]_\Omega P$$

where $P \in GL_n(F)$ is the transition matrix [1] from Ω to Ω'. Conversely, given any $P \in GL_n(F)$, there is a basis Ω' such that $[T]_{\Omega'} = P^{-1}[T]_\Omega P$. The general problem we would like to solve is the following:

Problem 10.1. *Given $T \in \mathfrak{gl}(V)$, find a "good" basis Ω (depending on T) such that the matrix $[T]_\Omega$ has certain canonical form.*

We will discuss two types of canonical forms: the rational canonical form (RCF) which exists over arbitrary fields and the Jordan canonical form (JCF) which is only guaranteed to exist over algebraically closed fields. Problem 10.1 can be reformulated entirely in terms of matrices:

Definition. Two matrices $A, B \in Mat_n(F)$ are called similar if there exists $P \in GL_n(F)$ such that $B = P^{-1}AP$. The set of all matrices similar to A is called the similarity class of A.

Problem 10.1'. *Given a matrix $A \in Mat_n(F)$, find a matrix in certain canonical form which is similar to A.*

10.1. Rational canonical form.

Once again fix a field F, a f.d. F-vector space V and $T \in \mathfrak{gl}(V)$. Denote by V_T the $F[x]$-module whose underlying set is V and x acts as T.

[1] This means that if we expand the i^{th} element of Ω' as a linear combination of elements of Ω, then the coefficients of that linear combination form the i^{th} column of P

By classification of modules over PID in IF form there exist nonzero non-constant polynomials $a_1(x) \mid \ldots \mid a_m(x)$ s.t.

$$V_T = V_1 \oplus \ldots \oplus V_m \text{ where } V_i \cong F[x]/(a_i(x)) \text{ as } F[x]\text{-modules.}$$

Note that we cannot have a free summand $F[x]^s$ in this decomposition since $\dim_F F[x] = \infty$ while $\dim_F V < \infty$. Also WOLOG we can assume that polynomials $a_i(x)$ are monic (which makes them unique).

Since each V_i is an $F[x]$-submodule of V_T, we get that V_i is T-invariant, that is, $T(V_i) \subseteq V_i$. Let $T_i = T_{|V_i} \in \mathfrak{gl}(V_i)$ be the restriction of T to V_i. Choose a basis $\Omega_i = \{e_{i_1}, \ldots, e_{i,k_i}\}$ for V_i. Then $\Omega = \Omega_1 \sqcup \ldots \sqcup \Omega_m$ is a basis[2] of V, and it is clear that the matrix $[T]_\Omega$ has the following block-diagonal form:

$$[T]_\Omega = \begin{pmatrix} [T_1]_{\Omega_1} & 0 & \ldots & 0 \\ 0 & [T_2]_{\Omega_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & [T_m]_{\Omega_m} \end{pmatrix}$$

Thus, to get a "satisfactory solution" to Problem 10.1 it is enough to put each T_i into certain canonical form (whatever the latter means). In other words, we are reduced to the following problem:

Let W be a f.d. F-vector space, $T \in \mathfrak{gl}(W)$ and $W_T = F[x]/(a(x))$ as $F[x]$-modules. We want to find a basis Ω of W s.t. $[T]_\Omega$ is in certain canonical form.

Suppose that $a(x) = x^k + a_{k-1}x^{k-1} + \ldots + a_0$. Then clearly $\{\overline{1}, \overline{x^1}, \ldots, \overline{x^{k-1}}\}$ is an F-basis for W where $\overline{x^i}$ is the image of x^i in W_T.

How does T act on this basis? We have

$T(\overline{x^i}) = x \cdot \overline{x^i} = \overline{x^{i+1}}$ – another basis element if $i < k-1$ and
$T(\overline{x^{k-1}}) = -a_{k-1}\overline{x^{k-1}} - \ldots - a_0 \cdot \overline{1}$.

So,

$$[T]_\Omega = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{k-1} \end{pmatrix}$$

This $k \times k$ matrix is called the companion matrix of the polynomial $a(x)$ and will be denoted by $C_{a(x)}$. **Remark:** If $|\Omega| = 1$, then $C_{a(x)}$ is the 1×1 matrix with entry $-a_0$.

[2] We choose the order on Ω such that elements from Ω_1 come first, then elements of Ω_2 etc.

Definition. A matrix $A \in Mat_n(F)$ is said to be in rational canonical form if there exists monic non-constant polynomials $a_1(x) \mid a_2(x) \mid \ldots \mid a_m(x) \in F[x]$ s.t.

$$A = \begin{pmatrix} C_{a_1(x)} & 0 & \ldots & 0 \\ 0 & C_{a_2(x)} & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & C_{a_m(x)} \end{pmatrix} \qquad (***)$$

Theorem 10.2 (Existence and uniquenss of Rational Canonical Form (RCF))**.** *Let V be a f.d. vector space over a field and $T \in \mathfrak{gl}(V)$. Then there is a basis Ω of V and a uniquely defined matrix A in rational canonical form s.t. $[T]_\Omega = A$. The matrix A is called* **the** *rational canonical form of T.*

Proof. The existence part is the result of our earlier considerations. The uniqueness part follows from the uniqueness part in the IF classificaiton theorem since the $F[x]$-module V_T can be recovered from RCF of T – this easily follows from our argument (check details). □

Here is the "matrix form" of Theorem 10.2:

Theorem 10.2'. *Any matrix $A \in Mat_n(F)$ is similar to a unique matrix in rational canonical form.*

Corollary 10.3. *Suppose that F and K are fields with $F \subset K$ and $A, B \in Mat_n(F)$. Then A and B are similar in $Mat_n(F)$ $\iff$ A and B are similar in $Mat_n(K)$.*

Proof. The " $\Leftarrow''$ direction is clear. Now suppose that A and B are not similar in $Mat_n(F)$. Then by Theorem 10.2' they have distinct RCF. Since RCF does not change under field extensions, again by Theorem 10.2' they are still not similar in $Mat_n(K)$. □