4. Canonical Forms

Canonical Forms: Preliminaries

Set-up

Let F be a field, let V be a nontrivial finite dimensional vector space over F, and let $T : V \to V$ be a linear transformation.

Then V is an $F[x]$-module.

Lemma: V is a finitely generated torsion module over $F[x]$.

Proof: V is generated by a basis, which is finite by assumption. If V were not torsion, there would be a $v \in V$ so that the map $F[x] \to V$ given by $f(x) \mapsto f(T)v$ is injective; but that contradicts the fact that V is finite dimensional.

Eigenvalues and the Characteristic Polynomial.

Definition: An element $\lambda \in F$ is an eigevanlue of T if there is a nonzero $v \in V$ with $Tv = \lambda v$. The vector v is an eigenvector for this eigenvalue.

Definition: Let $c(x) = \det(xI - T)$. Then $c(x)$ is a monic polynomial of degree n called the *characteristic polynomial* of T.

Lemma: λ is an eigenvalue of T if and only if $c(\lambda) = 0$.

Proof: If $c(\lambda) = 0$, then $\det(\lambda I - T) = 0$ so $\lambda I - T$ has a nontrivial kernel; an element of the kernel is an eigenvector with eigenvalue λ. Conversely, if λ is an eigenvalue, then $\lambda I - T$ has nontrivial kernel, so its determinant is zero.

The Minimal polynomial

Since V is a torsion $F[x]$ module, its annihilator

$$\mathrm{Ann}(V) = \{f(x) \in F[x] : f(T)v = 0 \forall v \in V\}$$

is a nonzero ideal in $F[x]$, hence a principal ideal generated by a unique monic polynomial $m(x)$.

Definition: The unique monic generator $m(x)$ of $\mathrm{Ann}(V)$ is called the minimal polynomial of T.

Lemma: The minimal polynomial $m(x)$ is the monic polynomial of minimal degree such that $m(T)v = 0$ for all $v \in V$.

Proof: By definition $m(T)v = 0$ for all $v \in V$. The collection of polynomials $h(x)$ such that $h(T)v = 0$ is exactly $\text{Ann}(V)$, and the generator of the ideal $\text{Ann}(V)$ in the Euclidean ring $F[x]$ is its monic polynomial element of minimal degree.

The structure of V

By the fundamental theorem of finitely generated modules over PID's, we have two different ways to represent V as an $F[x]$ module.

Invariant Factors: There are monic polynomials $f_1(x)|f_2(x)|\cdots|f_k(x)$ such that

$$V = F[x]/(f_1(x)) \oplus F[x]/(f_2(x)) \oplus \cdots \oplus F[x]/(f_k(x)).$$

Elementary Divisors: There are irreducible polynomials $f_1(x), \ldots, f_k(x)$ and nonnegative integers $e_1, \ldots, e_k$ such that

$$V = F[x]/(f_1(x)^{e_1}) \oplus F[x]/(f_2(x)^{e_2}) \oplus \cdots \oplus F[x]/(f_k(x)^{e_k})$$

The invariant factors are uniquely determined; and the elementary divisors are uniquely determined up to order.

Invariant factors and the minimal polynomial

Lemma: The minimal polynomial $m(x)$ of T is the last (the "largest") invariant factor of T acting on V; all the invariant factors divide $m(x)$.

The Rational Normal Form

The term **"canonical normal form"** is also used to refer to the rational normal form. ### The cyclic case

Let's focus our attention for the moment on a cyclic $F[x]$ module of the form $M = F[x]/(f(x))$ where

$$f(x) = x^d + a_{d-1}x^{d-1} + \ldots + a_0$$

- M is a finite dimensional vector space of degree equal to the degree d of $f(x)$.
- $1, x, x^2, \ldots, x^{d-1}$ is a basis for this module.
- x is an F-linear transformation $M \to M$.

Matrix of x

The linear map $[x]$ given by multiplication by x has the following matrix form in the basis $1, x, x^2, \ldots, x^{d-1}$:

$$[x] = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}$$

Definition: The matrix for $[x]$ above is called the companion matrix for the polynomial $f(x)$.

Characteristic and minimal polynomials of the companion matrix

Lemma: The characteristic and minimal polynomials of this linear transformation are both $f(x)$.

Proof: The fact that the characteristic polynomial of $[x]$ is $f(x)$ is a computation. The fact that $f(x)$ is the minimal polynomial follows from the fact that x clearly satisfies $f(x)$, and, since $1, x, \ldots, x^{d-1}$ are linearly independent, there is no relation $\sum a_i x^i = 0$ of degree less than d.

The general case

If M is a finitely generated torsion $F[x]$ module, then it is a direct sum of cyclic modules:

$$M = F[x]/(f_1(x)) \oplus \cdots \oplus F[x]/(f_k(x))$$

where $f_1(x)|f_2(x)|\cdots|f_k(x)$.

Each of the factor modules has a basis consisting of powers of x; so M has a basis obtained by stringing together the powers of x corresponding to each factor. Therefore the basis of the linear map $[x]$ is made up of blocks, where each block is the companion matrix for the polynomial $f_i(x)$ for $i = 1, \ldots, k$.

The rational normal form of a matrix

Now let V be a finite dimensional vector space over F and $T : V \to V$ be a linear map. Viewing V as a finitely generated torsion $F[x]$ module, we can find a basis for V so that the matrix of T is in block form, where each block is the companion matrix to a monic polynomial $f_i(x)$, and where $f_1(x)|\cdots|f_k(x)$.

This is called the *rational normal form* or *rational canonical form* for the linear map T.

- The polynomials that occur in the companion matrices are unique.
- Every matrix is conjugate to a unique matrix in rational normal form.

The minimal and characteristic polynomials

Proposition: If T is in rational normal form, then the minimal polynomial of T is the "last" polynomial $f_k(x)$ and the characteristic polynomial is the product $f_1(x) \cdots f_k(x)$.

Proof: The result about the minimal polynomial follows because $f_k(x)$ is zero but x cannot satisfy a lower degree polynomial because $1, x, \ldots, x^{d-1}$ are linearly independent in the "last" factor. The result about the characteristic polynomial follows because the determinant of a block matrix is the product of the determinant of the blocks.

Corollary: (The Cayley-Hamilton Theorem) The minimal polynomial of T divides its characteristic polynomial, and the characteristic polynomial divides some power of the minimal polynomial. In particular they have the same roots. T satisfies its characteristic polynomial.

Fields and subfields

If A is a matrix in $M_n(F)$, and $F \subset K$ is a subfield of K, then the rational normal form of A is the same whether computed over F or K. This is because if $f_1 | \cdots | f_k$ are the invariant factors of the module obtained from A over F, they satisfy the uniqueness properties over K as well.

If A and B are similar over K, then they are similar over F; because they have the same rational normal form over K, and therefore also over F.

Jordan normal form

Elementary divisors version

The rational normal form comes from the invariant factor version of the fundamental theorem, while the Jordan normal form comes from the elementary divisors version. In other words, we write our module

$$M = \oplus F[x]/(f_1(x)^{e_1}) \oplus F[x]/(f_2(x)^{e_2}) \oplus \cdots \oplus F[x]/(f_k(x)^{e_k})$$

where the $f_i(x)$ are all *monic irreducible* polynomials. This decomposition is unique up to reordering of the factors.

Characteristic and minimal polynomials

Lemma: The characteristic polynomial of $[x]$ is the product of the $f_i^{e_i}$ and the minimal polynomial is the least common multiple of the $f_i^{e_i}$.

Proof: In this version, one can use the powers of x in each factor and find a block diagonal representation of the linear map $[x]$; each block is the companion matrix

of $f_i^{e_i}$ and so the characteristic polynomial is the product of the elementary divisors. The minimal polynomial is the polynomial of smallest degree divisible by all of the $f_i^{e_i}$ which is their least common multiple.

Jordan normal form

We *assume that the roots of the characteristic/minimal polynomial of T all belong to the field F*. In this case, the f_i are all of degree one and

$$M = \oplus F[x]/((x-\lambda_1)^{e_1}) \oplus \cdots \oplus F[x]/((x-\lambda_k)^{e_k})$$

where the λ_i include all of the eigenvalues of T, although they aren't necessarily distinct. The action of $[x]$ breaks up into blocks according to how x acts on each factor.

Jordan normal form continued

We take one of the factors $F[x]/((x-\lambda)^e)$ and use $v_i = (x-\lambda)^i$ for $i = 0, \ldots, e-1$ as our basis. Then

$$xv_i = x(x-\lambda)^i = (x-\lambda)^{i+1} + \lambda(x-\lambda)^i$$

remembering that $(x-\lambda)^e = 0$.

Jordan normal form continued

The matrix of $[x]$ in this basis looks like:

$$[x] = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 \cdots & \lambda & 1 \\ 0 & 0 & 0 \cdots & 0 & \lambda \end{pmatrix}$$

This is called a *Jordan Block* and every linear map has a basis in which it is built out of Jordan blocks on the diagonal. Similarly every matrix is conjugate to a matrix in Jordan block form, assuming its eigenvalues lie in the field F.

Uniqueness of Jordan block

The invariants of the Jordan block form are the eigenvalues λ and the sizes of the blocks e. The form is unique up to rearrangement of these blocks.

Diagonalizability

A matrix A is diagonalizable over a field F if and only if its minimal polynomial has no repeated roots.

View as slides