Week 8

???

8.1 Linear Algebra Review and Rational Canonical Form

2/20:

- Nori's change of heart.
  - We've all seen linear algebra; thus, we'll speedrun it and then do exterior algebra and determinants. That's where we'll finish.
- The following is part 1 of a linear algebra course.
- Let F be a field.
- **Vector space**: An F-module.
- **Linearly independent** (subset $S \subset V$): Same definition we're familiar with.
- **Spanning** (subset $S \subset V$): A subset S of V that is a set of generators of V.
- S is a **basis** implies that S generates V and is linearly independent.
- Every linearly independent subset of V can be extended to a basis.
- Every spanning set S contains a basis.
  - Any maximal linearly independent subset of S is a basis.
- S_1, S_2 are bases for V implies that $|S_1| = |S_2|$.
  - The replacement theorem in Dummit and Foote (2004) is a good way to prove this.
- We are now done with part 1; this is part 2 of a linear algebra course.
- Let $T : V \to V$ be a linear transformation.
- Let A be a ring. What is an $A[X]$-module M?
  - It is an abelian group $(M, +)$ and a ring homomorphism $\rho : A[X] \to \mathrm{End}(M, +)$.
  - Since $A \hookrightarrow A[X]$, $\rho|_A$ turns M into an A-module.
  - Since $aX = Xa$, $\rho(a)\rho(X) = \rho(X)\rho(a)$.
  - But since we consider M to be a module, we write $a := \rho(a)$: Thus, $a\rho(X)m = \rho(X)am$ for all $m \in M$.
  - Note that $\rho(X) \in \mathrm{End}_A(M)$ (which is the set of all A-module endomorphisms).
  - Additionally, $\rho(X) : M \to M$ is an A-module homomorphism.

- Put $\rho(X) = T$. Thus, an $A[X]$-module is a pair (M, T), where M is an A-module and $T \in \mathrm{End}_A(M)$.
- Conversely, such (M, T) gives rise to an $A[X]$-module with action

$$\left(\sum_{n=0}^{\ell} a_n X^n\right) m = \sum_{n=0}^{\ell} a_n T^n m$$

- Let F be a field, and consider (V, T) where V is any F-vector space and $T : V \to V$ is a linear transformation.
  - This induces a module over $F[X]$.
- V finite dimensional induces $\rho : F[X] \to \mathrm{End}_F(V) \cong M_n(F)$ defined by $X \mapsto T$.

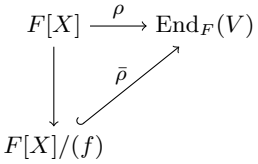

Figure 8.1: $F[X]$-module actions.

  - $\rho(X) = T$ and $\rho(c) = c$ for all $c \in F$.
  - $\ker(\rho) = (f)$ for some monic polynomial f of degree $d \leq n^2$.
  - We have the constraint on the degree of f by the isomorphism from Lecture 3.1.
- **Minimal polynomial** (of T): The polynomial f that generates $\ker(\rho)$.
  - In particular, V is a finitely generated torsion $F[X]$-module.
- **Cyclic vector**: A vector $v \in V$ belonging to (V, T) such that $v, Tv, T^2v, \ldots$ spans V.
- Using cyclic vectors to compute the minimal polynomial.
  - Assume $v, Tv, T^2v, \ldots, T^{k-1}v$ are linearly independent, but $v, Tv, \ldots, T^k v$ are not.
  - Then
$$T^k v = a_0 v + a_1 Tv + \cdots + a_{k-1} T^{k-1} v$$
where all $a_i \in F$ and not all $a_i = 0$.
  - Let $W = \langle v, Tv, \ldots, T^{k-1}v \rangle$. It follows that $T^m k \in W$.
  - Let
$$g(X) = X^k - (a_{k-1}X^{k-1} + \cdots + a_1 X + a_0)$$
Then $g(T)v = 0$. This implies that g is the minimal polynomial of T.
  - It follows that $T^h g(T)v = 0$. Thus, $g(T)T^h v = 0$ for all h.
  - Lastly, it follows that $g(T)w = 0$ for all $w \in W$.
  - Assume v is a cyclic vector. Then $W = V$. It follows that $g(T)v = 0$ for all $v \in V$.
  - The original assumption posits that no polynomial of degree less than or equal to $k - 1$ can annihilate v.
- Consider $V = F[X]/(f)$. Let $\deg(f) = d$, let $T : V \to V$, and let T be the "multiply by X" linear transformation. It follows that if $v_i = \overline{X^{i-1}}$ $(i = 1, \ldots, d)$, then

$$Tv_i = v_{i+1}$$

for $i = 1, \ldots, d-1$ and

$$Tv_d = -(a_0 v_1 + a_1 v_2 + \cdots + a_{d-1} v_d)$$

  - If $d = 3$, then we have
$$M(T) = \begin{pmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{pmatrix}$$
  - The above matrix is called the **companion matrix** of f, for f monic of degree 3.

- **Rational canonical form**: The form (V, T) given by
$$F[X]/(f_1) \oplus \cdots \oplus F[X]/(f_s)$$
where $f_2 \mid f_1, \ldots, f_s \mid f_{s-1}$ and $\deg(f_s) > 0$.
  - When $V = 0$, then $s = 0$. In this case, f_1 is the minimal polynomial of T.
  - The form consisting of a block diagonal matrix of companion matrices.

- **Jordan canonical form**:
  - Has to do with p-primary components!

- There's one more canonical form, too.

- Since no one knows what canonical forms are and we very much need them for what Nori was planning to do, Nori will change his plans. No tensors in the last week, either.

- p-primary components: When $p = X - a$, $a \in F$.

- (V, T) is ***p*-primary** if there exists an n such that $(T - a)^n v = 0$ for all $v \in V$.

- $1_V : V \to V$ is the identity.

- $a \cdot 1_V = a_V : V \to V$.

- $(T - a_v)^n = 0 \in \mathrm{End}_F(V)$.

- We're now doing generalized eigenspaces ?? lol.

- The p-primary component is as the generalized a-eigenspace.
  - $(T - a)v = 0$, i.e., $Tv = av$ is the a-eigenspace; the eigenspaces are components of the generalized eigenspaces.

- Let $V = F[X]/(X - a)^n$. Let $v_1 = 1$, $v_2 = \overline{X - a}$, $\ldots$, $v_n = \overline{(X - a)^{n-1}}$.
  - We know that $X(X - a)^r = (X - a + a)(X - a)^r = (X - a)^{r+1} + a(X - a)^r$.
  - Nori writes Jordan blocks as
$$\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix}$$
not with 1's in the superdiagonal.
    - Thus, the *last* generalized eigenvector is an eigenvector here, instead of the *first*.

8.2 Office Hours (Nori)

- Midterm: We never covered the universal property of a quotient in class, did we?
  - That's the special lemma from last office hours.
- PSet 7: 7.3 and 7.4 typos.
- Wednesday lecture?
  - Seventh week summary will suffice.
- What do you need us to know about the rational canonical form? Should I still read Dummit and Foote (2004), Section 12.2 or is that no longer necessary?
  - Nori will probably push ahead with 12.2. Thus I should read it. He's not sure what he'll do beyond that, though, since he doesn't want to jam tensors into the last week.
  - I will need tensor products for representation theory, regardless, so if I want to take it, I should self-study it.
  - No chance tensor products will be covered next quarter.
  - Serre is a terrific mathematician whose wife is a super chemist, and that's why he wrote his book on representation theory (and wrote it in a less terse manner than usual).
  - No tensors means no exterior algebra, too.
  - Nori hasn't read any of Dummit and Foote (2004).
  - The transfer theory of groups arises in a later chapter, and that's important for representation theory, though.
- Nori doesn't think any teacher pays attention to what courses are supposed to cover as stated in the course catalog.
  - We will never do modules, multilinear and quadratic forms.
  - p-adic field and Galois theory.
  - Nori thinks the proof of Theorem 12.4 is very difficult to follow for a first-timer.
  - Solvable groups were supposed to be a MATH 25700 topic, but got cut because of 9-week quarters.
  - Syclotomic fields have applications to the representation theory of finite groups; there are theorems of representation theory that you need syclotomic fields to prove.
  - Emil Artin: Galois Theory is worth looking up.
  - Gauss and constructions of 17-gons also needs syclotomic fields.

8.3 Office Hours (Nori)

2/21:

- Lecture 6.1: Proposition proof?
- Lecture 6.1: $(2) \subsetneq \mathbb{Z}$ example?
- Lecture 6.1: The end of the theorem proof.
- Lecture 6.2: Does the first theorem you proved not appear in the book until Chapter 12?
- Lecture 6.2: What is A in the proof?
- Resources for the proofs in Week 6?
- Lecture 7.1: Quotient stuff.

- Lecture 7.2: Why does $\text{Ann}(v) = (p^k)$, why not just $(p^k) \subset \text{Ann}(v)$? Additionally, how does $p^k w' = 0$ imply that $p^k \in \text{Ann}(w)$?
  - R is a PID!
  - $\text{Ann}(w)$ should be $\text{Ann}(w')$ in the centered line.
  - We don't need to know the theorem from the book for a while (second year of graduate school at least).
  - It's good to know the proofs from class just for going forward in math, but we probably will not be asked to reproduce them on an exam.
- Lecture 7.3: RCF proof?
  - It's not $m_i, 1$, it's $m_{i,1}$!
  - Rewrite the proof when I'm awake enough to understand it.

8.4 Noncommutative Polynomial Rings

2/22:

- Adjusted syllabus for those of us who haven't seen block matrices.
  - We're not gonna cover all of the stuff in Sections 12.2-12.3.
  - We'll define determinants via exterior algebras and use them to prove the Cayley-Hamilton theorem.
- Notation for today:
  - We fix R a commutative ring.
  - Any other ring considered is potentially noncommutative.
  - c denotes an element of R.
- ***R*-algebra**: A pair (A, ϕ) where A is a ring and $\phi : R \to A$ is a ring homomorphism such that

$$\phi(c)a = a\phi(c)$$

  for all $a \in A$ and $c \in R$.
  - I.e., we have commutativity with the "integers."
  - Notation: We suppress the letter ϕ.
    - Justification: A can be considered an R-module in a natural manner, so we can use the standard notation $c \cdot a$ in place of $\phi(c) \cdot a$ because we can think of ϕ as the action of R on A.
  - R-algebras are automatically R-modules.
  - $\phi(R)$ is a subring of A (and, specifically, the center of A).
- Now we get into noncommutative polynomial rings.
  - We've defined polynomial rings on one variable, and on multiple variables by induction.
  - There, the variables commuted. However, they need not! We can construct rings such that $XY \neq YX$.
- Consider the ring consisting of n potentially noncommutative variables over the coefficients in R.
  - Essentially, we postulate that $1, X_1, X_2, \ldots, X_n$ are in the ring and consider the set of everything that can be generated from these elements via addition, multiplication, and the action of R.
  - Thus, things like $X_i X_j$ exist. Note that there are n^2 of these. Things like $X_i X_j X_k$ also exist (and there are n^3 of these).

- Polynomials will be of the form

$$c_0 + \sum_{i=1}^{n} c_i X_i + \sum_{\substack{1 \le i \le n \\ 1 \le j \le n}} c_{i,j} X_i X_j + \cdots$$

  - This polynomial has degree at most n.
    - Really?? How does Nori define degree? This would appear to be a break from the convention introduced in Section 9.1.
  - Note that since this is not a **noncommutative power series ring**, we assume that all polynomials have a finite degree.

- **Free R-algebra** (on the set S): The free R-module on the set $\{1\} \sqcup S \sqcup S \times S \sqcup S \times S \times S \sqcup \cdots$. *Denoted by* $\boldsymbol{F_R(S)}$.
  - Elements of the set include 1, X_s for all $s \in S$, $X_s X_t$ for all $s, t \in S$, etc.
  - Multiplication is still associative: $(X_{s_1} X_{s_2})(X_{t_1} X_{t_2} X_{t_3}) = X_{s_1} X_{s_2} X_{t_1} X_{t_2} X_{t_3}$.
  - We still have commutativity for coefficients (i.e., $X_s c = c X_s$ for all $c \in R$ and $s \in S$) since $F_R(S)$ is an R-algebra.
    - Did Nori qualify this?? Did he say that we do not write the c_i's at the front any more?

- ***R*-algebra homomorphism**: A function $h : A \to B$, where A, B are R-algebras, such that...
  (i) h is a ring homomorphism;
  (ii) $h(c) = c$ for all $c \in R$.

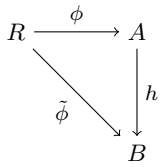

Figure 8.2: Visualizing R-algebra homomorphisms.

  - For clarity: If we momentarily drop our convention of suppressing ϕ, (ii) means $h(\phi(c)) = \tilde{\phi}(c)$ for all $c \in R$.
  - Alternate form of (ii): h is an R-module homomorphism.
    - This expresses the idea that the action of elements $c \in R$ on A is preserved under h, i.e., if $c \cdot a = ca$, then $ch(a) = c \cdot h(a) = h(c) \cdot h(a) = h(c \cdot a) = h(ca)$.

- Universal property of $F_R(S)$: Consider $i : S \to F_R(S)$ defined by $i(s) = X_s$. It is a map of sets since S has no additional structure. Given an R-algebra A and function $j : S \to A$, there exists a unique R-algebra homomorphism $h : F_R(S) \to A$ such that the following diagram commutes, i.e., $h(i(s)) = j(s)$ for all $s \in S$.

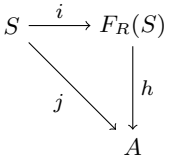

Figure 8.3: Universal property of $F_R(S)$.

- The principle of this universal property is the same as that of the universal property of polynomial rings. Seeing this:
  - Since $i(s) = X_s$, the statement $h(i(s)) = j(s)$ becomes $h(X_s) = j(s)$ for all $s \in S$.
  - This combined with the fact that h is an R-algebra homomorphism forces both existence and uniqueness for a reason symmetric to the proof of the universal property of polynomial rings in Lecture 1.2.
- We now develop a related object that appears in Chapter 10 called a **tensor algebra**.
  - A tensor algebra is very similar to a noncommutative polynomial ring.
- **Tensor algebra** (of M): A pair (A, α), where A an R-algebra and $\alpha : M \to A$ is an R-module homomorphism. *Denoted by* $\boldsymbol{T(M)}$.
- Universal property of the tensor algebra $T(M)$: $T(M)$ is an R-algebra, $u : M \to T(M)$ is an R-module homomorphism such that for all pairs $(A, \alpha : M \to A)$, there exists a unique R-algebra homomorphism $f : T(M) \to A$ such that the following diagram commutes, i.e., $f \circ u = \alpha$.

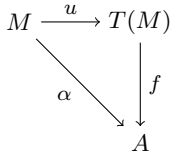

Figure 8.4: Universal property of $T(M)$.

- Example: We now work out one specific case.
- Let's construct $T(M)$ when M is a free R-module with S as the basis.
  - We have
$$M = \left\{ \sum_{s \in S} c_s e_s : c_s \in R,\ S \text{ is finite} \right\}$$
  - There's only finitely many nonzero s??
  - If S is finite, then $M = R^S$??
  - Let $\alpha : M \to A$ with M as above.
  - To specify α, you need only specify its action on a basis of M; thus, there is a bijection between the collection of all α and the set S (??) given by j??
  - This induces a ring homomorphism $h : F_R(S) \to A$ given by $h(X_s) = a_s$.
  - Here, $T(M) = F_R(S)$.
  - How do you know it doesn't depend on the basis chosen? By the universal property (is this what he said??).
  - Any module is the quotient of a free module.
- There will be more questions on the final like questions 1-2 on the midterm!!!
- Certain zeroes exist in $F_R(S)$.
- We now start on Section 10.4 and exterior algebras.
- $F_R(S)/\langle X_s X_t - X_t X_s : s, t \in S \rangle$ is the two-sided ideal generated by the given elements. It is the usual/commutative polynomial ring in the X_s ($s \in S$)??

- We now move on to the Grassmann algebra.
- In calculus, we get things like $v \wedge v = 0$ and $v_1 \wedge v_2 = -v_2 \wedge v_1$.
- M will always be a free R-module on the set S (i.e., with basis S) for us.
- We now define the **exterior algebra** of M.
- **Exterior algebra** (of M): A pair (A, α) where A is an R-algebra and $\alpha : M \to A$ is an R-module homomorphism such that for all $m \in M$, $\alpha(m)^2 = 0$. *Denoted by* $\boldsymbol{\bigwedge(M)}$.
  - We can prove the existence and uniqueness of $\bigwedge(M)$.
- Conditions (is this the universal property??).
  1. Condition (*) holds for the horizontal arrow, i.e., for $\bigwedge(M)$, $u(m)^2 = 0$ for all $m \in M$.
  2. If (*) holds for (A, α), then there exists a unique h such that the diagram below commutes.

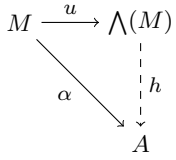

Figure 8.5: Existence and uniqueness of the exterior algebra.

- h is an R-algebra homomorphism such that $(h \circ u)(m) = \alpha(m)$ for all $m \in M$.
- Rather than write out the proof, just ask, "why not take $F_R(S) = \bigwedge(M)$?"

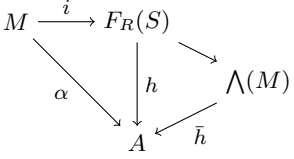

Figure 8.6: The free R-algebra on the set S and the exterior algebra.

  - Consider the above commutative diagram.
  - Then replace $F_R(S)$ by $F_R(S)/\left\langle i(m)^2 : m \in M \right\rangle = \bigwedge(M)$ and extend the commutative diagram.
- Let $m = \sum_{s \in S} c_s e_s \mapsto \sum_{s \in S} c_s X_s$. We have

$$\left(\sum_s c_s X_s \right)^2 = \sum_{s \in S} c_s^2 X_s^2 + \sum_{\substack{\{s,t\} \\ s \neq t}} c_s c_t (X_s X_t + X_t X_2)$$

- For the moment, the R-module spanned modules X_s^2 for all $s \in S$, including $X_s X_t + X_t X_s$ and $(X_s + X_t)^2 = X_s^2 + X_t^2 + (X_s X_t + X_t X_s)$.
- Something on the difference between $v^2 = 0$ and $v_1 v_2 = -v_2 v_1$ in normal calculus and what we've done today...??

8.5 Office Hours (Callum)

- Lecture 7.3: Why did you use the direct product instead of the direct sum in your proof of the RCF theorem? Also, do we still need the N condition on the $m_{i,j}$, or did the way I phrase it suffice? How do you know that there's only finitely many distinct primes?
  - If you have a finitely generated PID R and an R-module M, then is any submodule of M finitely generated?
- Problem 7.3?
  - It is a straight-up ring problem; we shouldn't be doing anything fancy with modules.

8.6 Office Hours (Nori)

2/23:

- Lecture 6.1: Proposition proof?
- Lecture 6.1: $(2) \subsetneq \mathbb{Z}$ example?
- Lecture 6.1: The end of the theorem proof.
- Lecture 6.2: Does the first theorem you proved not appear in the book until Chapter 12?
- Lecture 6.2: What is A in the proof?
- Resources for the proofs in Week 6?
- Lecture 7.1: Quotient stuff.
- Is my proof for Q7.1(ii) sufficiently not hand-wavey?
- Question 7.7?
  - Note: The four given submodules satisfy the $T(N) \subset N$ condition even if M does not have the particular form given! It just so happens that in this case, these four are the only allowable ones.
- Solving Question 7.7?
  - WTS: $N = pM$ or $\ker(p_M)$.
  - We know that

$$pM = pR/(p^2) \oplus 0 \qquad \ker(p_M) = pR/(p^2) \oplus R/(p)$$

  - $\ker(p_M)$ is a 2D vector space over $R/(p)$.
  - WTS: $N \cap \ker(p_M) \neq 0$ iff $N \neq 0$.
    - We know that $pN \subset N$ by the definition of N as a submodule.
    - Let $n \in N$ be nonzero. Suppose $n \notin \ker(p_M)$. Then $pn \in \ker(p_M)$. We know that $pn \in N$ as well. Thus, $pn \in N \cap \ker(p_M)$.
  - $N \cap \ker(p_M) \subset \ker(p_M)$. Thus, $N \cap \ker(p_M)$ is either a 1D or a 2D vector space over $R/(p)$. We WTS that if it's 2D, then it equals $\ker(p_M)$, and if it's 1D, then it equals pM.
  - 2D case.
    - We know that $N \cap \ker(p_M) \subset \ker(p_M)$.
    - 2D implies that $N \not\subseteq \ker(p_M)$.
    - Thus, either $N = \ker(p_M)$ or $N \supsetneq \ker(p_M)$.
    - In the first case, we are done.
    - In the second case, we can show that this implies that $N = M$.

- 1D case.
  - We know that $pM \cap \ker(p_M) = p_M$.
  - Assume $N \neq pM$.
  - Then $N \cap \ker(p_M) = \langle (pa, 1) \rangle$.
  - But T exists, where $T : M \to M$ sends $T(1,0) = (1,0)$ and $T(0,1) = (p,1)$.
  - Therefore we must have $N \cap \ker(p_M) = pM$.
- Suppose that $N \supsetneq pM$.
  - $N/pM \subset M/pM$. Then use $T(1,0) = (1,1)$ and $T(0,1) = (0,1)$.

8.7 Chapter 12: Modules over Principal Ideal Domains

From Dummit and Foote (2004).

Section 12.2: The Rational Canonical Form

2/21:

- As stated previously, we apply the results of Section 12.1 to $F[X]$-modules herein.
- Let V be a finite dimensional vector space over F of dimension N. Let (V, T) be an $F[X]$-module.
- Since V is finite dimensional, it is finitely generated as an F-module and hence also as an $F[X]$-module.
- If V were free, it would be isomorphic to a direct sum of copies of $F[X]$ (by Theorem 12.5(1)) and hence be infinite dimensional.
  - Thus, V is a torsion $F[X]$-module.
  - Theorem 12.5(3): V is isomorphic to the direct sum of cyclic, torsion $F[X]$-modules.
  - This decomposition will allow us to choose a basis for V with respect to which the matrix representation for the linear transformation T is in a specific simple form.
- **Rational canonical form** (of a matrix): The form obtained when we use the invariant factor decomposition of the relevant vector space.
- **Jordan canonical form** (of a matrix): The form obtained when we use the elementary divisor decomposition (and when F contains all the eigenvalues of T).
- Theorem 12.9 ensures that the RCF and JCF are unique, justifying the labeling of them as *canonical*.
- An application of canonical forms: Classifying distinct linear transformations.
  - Two matrices that represent the same linear transformation (hence are similar) have the same RCF and JCF.
  - This is another instance of the structure of the space being acted upon (e.g., the invariant factor decomposition of V) providing information on the algebraic objects (e.g., linear transformations) which are acting.
- **Representation Theory of Groups**: The special case of algebraic objects acting on spaces concerning groups acting on vector spaces.
- **Eigenvalues**, **eigenvectors**, **eigenspaces**, and the **determinant** are defined for linear transformations and analogously for matrices.
- Properties of eigenvalues.

  Proposition 12.12. TFAE.

  1. λ is an eigenvalue of T.

2. $\lambda I - T$ is a singular linear transformation of V.
3. $\det(\lambda I - T) = 0$.

Proof. Given. □

- **Characteristic polynomial** (of a linear transformation): The polynomial defined as follows, where T is the linear transformation in question. *Denoted by* $\boldsymbol{c_T(X)}$. *Given by*

$$c_T(X) = \det(XI - T)$$

  - Defined similarly for matrices A.
  - A monic polynomial of degree $\dim V$.
  - The eigenvalues are the roots.

- **Minimal polynomial** (of of a linear transformation): The unique monic polynomial which generates the ideal $\mathrm{Ann}(V)$ in $F[X]$. *Denoted by* $\boldsymbol{m_T(X)}$.
  - Defined similarly for matrices A.
  - We know that such a polynomial exists by Theorem 12.5(3).
  - Exercise 12.2.5: The degree of the minimal polynomial is at most n^2.

- **Cayley-Hamilton Theorem**: The minimal polynomial for T is a divisor of the characteristic polynomial for T.
  - Thus, the degree of the minimal polynomial is at most n.

- We now build up to the **rational canonical form**.

- Introduction.
  - Theorem 12.5: There exists an isomorphism

$$V \cong F[X]/(a_1(X)) \oplus \cdots \oplus F[X]/(a_m(X)) \tag{12.1}$$

  - The invariant factors a_i are only determined up to units, but since $F[X]^\times = F - \{0\}$, we can make the a_i unique by requiring them to be monic.
  - Theorem 12.5(3) asserts that $(a_m(X)) = \mathrm{Ann}(V)$.

- The minimal polynomial and the invariant factors.

  Proposition 12.13. The minimal polynomial $m_T(X)$ is the largest invariant factor of V. All of the invariant factors of V divide $m_T(X)$.

- We now build up to calculating the minimal polynomial of T and the other invariant factors.

- Choosing a basis for each of the summands in Equation 12.1.
  - Recall that the action of T on V is equivalent to the action of X on each summand.
  - Recall also (from the Example following Proposition 11.1) that $1, \bar{X}, \bar{X}^2, \ldots, \bar{X}^{k-1}$ gives a basis of $F[X]/(a(X))$, where $a(X) = X^k + b_{k-1}X^{k-1} + \cdots + b_0$.
  - With respect to this basis, the linear transformation $T = l_X$ acts via

$$\begin{aligned}
1 &\mapsto \bar{X} \\
\bar{X} &\mapsto \bar{X}^2 \\
\bar{X}^2 &\mapsto \bar{X}^3 \\
&\vdots \\
\bar{X}^{k-2} &\mapsto \bar{X}^{k-1} \\
\bar{X}^{k-1} &\mapsto \bar{X}^k = -b_0 - b_1\bar{X} - \cdots - b_{k-1}\bar{X}^{k-1}
\end{aligned}$$

- The last equality holds since $a(\bar{X}) = 0$ in $F[X]/(a(X))$.

– With respect to this basis, the matrix for multiplication by X is called the **companion matrix** of $a(X)$.

– Applying this procedure to each of the cyclic modules on the right side of Equation 12.1 under an appropriate basis yields the **direct sum** of the companion matrices for the invariant factors as the matrix of T.

– Note that this matrix is uniquely determined by the invariant factors of the $F[X]$-module V. These invariant factors, in turn, uniquely determine V up to isomorphism by Theorem 12.9.

- **First subdiagonal**: The set of entries in a matrix which lie directly below a diagonal entry. *Also known as* **subdiagonal**.

- **Companion matrix** (of a polynomial): The $k \times k$ matrix, pertaining to the polynomial $a(X) = X^k + b_{k-1}X^{k-1} + \cdots + b_0$, which consists of 1's down the first subdiagonal, $-b_0, \ldots, -b_{k-1}$ down the last column, and zeros elsewhere. *Denoted by* $\boldsymbol{\mathcal{C}_{a(X)}}$. *Given by*

$$\mathcal{C}_{a(X)} = \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & -b_0 \\ 1 & 0 & \cdots & \cdots & \cdots & -b_1 \\ 0 & 1 & \cdots & \cdots & \cdots & -b_2 \\ 0 & 0 & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & -b_{k-1} \end{pmatrix}$$

- **Direct sum** (of matrices): The block diagonal matrix consisting of the component matrices.

  – See the RCF example below.

- **Rational canonical form** (of a matrix): A matrix that is the direct sum of companion matrices for monic polynomials $a_1(X), \ldots, a_M(X)$ of degree at least one with $a_1(X) \mid a_2(X) \mid \cdots \mid a_m(X)$. *Also known as* **RCF**. *Given by*

$$\begin{pmatrix} \mathcal{C}_{a_1(X)} & & & \\ & \mathcal{C}_{a_2(X)} & & \\ & & \ddots & \\ & & & \mathcal{C}_{a_m(X)} \end{pmatrix}$$

- **Invariant factors** (of the RCF): The polynomials a_i in the above definition.

- Definition of a **block diagonal** matrix.

- **Rational canonical form** (of a linear transformation): The matrix representing T which is in rational canonical form.

- Dummit and Foote (2004) proves that the rational canonical form is unique by means of running the generation process in reverse.

  Theorem 12.14 (Rational Canonical Form for Linear Transformations)**.** Let V be a finite dimensional vector space over the field F, and let T be a linear transformation of V.

  1. There is a basis for V with respect to which the matrix for T is in rational canonical form, i.e., is a block diagonal matrix whose diagonal blocks are the companion matrices for monic polynomials $a_1(X), \ldots, a_m(X)$ of degree at least one with $a_1(X) \mid a_2(X) \mid \cdots \mid a_m(X)$.
  2. The rational canonical from for T is unique.

- Why the *rational* canonical form?

- "Rational" refers to the fact that this canonical form is calculated entirely within the field F and exists for any linear transformation T.
- This is not the case for the JCF, which only exists if the field F contains the eigenvalues for T.

- Similar matrices, modules, and the RCF.

Theorem 12.15. Let S and T be linear transformations of V. Then TFAE.

1. S and T are similar linear transformations.
2. The $F[X]$-modules obtained from V via S and via T are isomorphic $F[X]$-modules.
3. S and T have the same rational canonical form.

Proof. Given. □

- Observation: Any $n \times n$ matrix A with entries in F arises as the matrix for some linear transformation T of an n-dimensional vector space.
- This observation allows us to restate Theorems 12.14-12.15 in the language of matrices.

Theorem 12.16 (Rational Canonical Form for Matrices)**.** Let A be an $n \times n$ matrix over the field F.

1. The matrix A is similar to a matrix in rational canonical form, i.e., there is an invertible $n \times n$ matrix P over F such that $P^{-1}AP$ is a block diagonal matrix whose diagonal blocks are the companion matrices for monic polynomials $a_1(X), \ldots, a_m(X)$ of degree at least one with $a_1(X) \mid a_2(X) \mid \cdots \mid a_m(X)$.
2. The rational canonical from for A is unique.

Theorem 12.17. Let A, B be $n \times n$ matrices over the field F. Then A, B are similar iff A, B have the same RCF.

- **Invariant factors** (of a matrix): The invariant factors of the matrix's RCF.
- RCF and similarity questions for A do not depend on which field contains the entries of A.

Corollary 12.18. Let A, B be two $n \times n$ matrices over a field F, and suppose F is a subfield of the field K.

1. The rational canonical form of A is the same whether it is computed over K or over F. The minimal and characteristic polynomials and the invariant factors of A are the same whether A is considered as a matrix over F or as a matrix over K.
2. The matrices A, B are similar over K iff they are similar over F, i.e., there exists an invertible $n \times n$ matrix P with entries from K such that $B = P^{-1}AP$ iff there exists an (in general different) invertible $n \times n$ matrix Q with entries from F such that $B = Q^{-1}AQ$.

Proof. Given. □

- Takeaways from Corollary 12.18.
  - The RCF for A is an $n \times n$ matrix with entries in the smallest field containing the entries of A.
  - Further explanation of the word *rational*: The RCF is the same matrix even if we allow conjugation of A by nonsingular matrices whose entries come from larger fields.
- Characteristic polynomials and invariant factors.

Lemma 12.19. Let $a(X) \in F[X]$ be any monic polynomial.

1. The characteristic polynomial of the companion matrix of $a(X)$ is $a(X)$.

2. If M is the block diagonal matrix

$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix}$$

given by the direct sum of matrices $A_1, \ldots, A_k$, then the characteristic polynomial of M is the product of the characteristic polynomials of $A_1, \ldots, A_k$.

Proof. See the exercises. □

Proposition 12.20. Let A be an $n \times n$ matrix over the field F.

1. The characteristic polynomial of A is the product of all the invariant factors of A.
2. (The Cayley-Hamilton Theorem) The minimal polynomial of A divides the characteristic polynomial of A.
3. The characteristic polynomial of A divides some power of the minimal polynomial of A. In particular, these polynomials have the same roots, not counting multiplicities.

Proof. Given. □

- The relations in Proposition 12.20 are frequently useful in determining the invariant factors of A, particularly for $\deg(A)$ small.
- **Elementary row and column operations**: The following three operations, where A is an $n \times n$ matrix over the field F and $XI - A$ is an $n \times n$ matrix with entries in $F[X]$. *Given by*

  (i) Interchanging two rows or columns.
  (ii) Adding a multiple (in $F[X]$) of one row or column to another.
  (iii) Multiplying any row or column by a unit in $F[X]$, i.e., by a nonzero element in F.

- **Smith Normal Form** (of a matrix): The following form of the $n \times n$ matrix $XI - A$ with entries from $F[X]$, where $a_1, \ldots, a_m$ are the invariant factors of A. *Given by*

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & a_1(X) & & & & \\ & & & & a_2(X) & & & \\ & & & & & \ddots & & \\ & & & & & & a_m(X) \end{pmatrix}$$

- Computing the invariant factors in general.

  Theorem 12.21. Let A be an $n \times n$ matrix over the field F. Using the three elementary row and column operations above, the $n \times n$ matrix $XI - A$ is with entries in $F[X]$ can be put into Smith Normal Form.

- Dummit and Foote (2004) provides algorithms for computing the invariant factor decomposition and the RCF. Return to later.

Exercises

5. Prove directly from the fact that the collection of all linear transformations of an n-dimensional vector space V over F to itself form a vector space over F of dimension n^2 that the minimal polynomial of a linear transformation T has degree at most n^2.