Lecture IX: Minimal polynomials and rational canonical form over $\mathbb{R}$

If $T: V \to V$ is a linear transformation of a complex vector space of finite dimension n (or equivalently if $T: \mathbb{C}^n \to \mathbb{C}^n$ is a linear transformation) then we have seen how to put T into a "canonical form", Jordan canonical form by finding basis for V (or $\mathbb{C}^n$) relative to which T has matrix representation consisting of "Jordan blocks" strung together. And we say that this representation was essentially unique in the sense that the eigenvalues, the dimensions of generalized eigenspaces, and the number and sizes of the Jordan blocks for each generalized eigenspace were all uniquely determined by T itself, with no arbitrary basis choices involved.

But if we want to consider real linear transformations — which are of course highly important difficulties arise: they may not have "Jordan form" in any basis; for example, their "eigenvalues" may not be real in the sense that $\det(T - \lambda I) = 0$ may not have all its roots (or indeed any of its roots) real. But it is worthwhile to go as far towards a "canonical form" as one can.

The fundamental idea is to associate a real linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ with a complex one, also denoted T (or if we need to make the distinction $T^{\mathbb{C}}$) from $\mathbb{C}^n$ to $\mathbb{C}^n$ by "complexifying" T: we simply let $T^{\mathbb{C}}$ act with the same matrix A as T relative to the standard basis, thinking of the entries of A as complex numbers — which happen to have 0 imaginary part. Now $T^{\mathbb{C}}$ has of course a Jordan form, but the (generalized) eigenvalues λ_j $j = 1, \ldots n$ (counting multiplicities, some repeated perhaps) are quite possibly not real numbers, e.g. for $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ $\begin{array}{l} a^2+b^2=1 \\ b \neq 0 \end{array}$. (2×2 real orthogonal matrices). This particular example we could already analyze, and did, because it is a normal operator. But it turns out that there is something useful we can do even when $T: \mathbb{R}^n \to \mathbb{R}^n$ is not normal (which "generic" linear transformation T are not! Normality is atypical!)

The construction is facilitated by looking at minimal polynomials. For $T^{\mathbb{C}}$, the "minimal polynomial" [recall: this is the minimal degree monic (= 1st coefficient 1) polynomial P with $\mathbb{C}$ coefficients such that $P(T) = 0$] is exactly and easily seen from the Jordan form, and indeed simply from knowing about the generalized eigenspaces in the following sense: For each eigenvalue λ_j of $T_0 : \mathbb{C}^n \to \mathbb{C}^n$ (whether or not T_0 arose as $T^{\mathbb{C}}$ for some $T : \mathbb{R}^n \to \mathbb{R}^n$) let d_j = the smallest (positive) integer such that $\dim \ker (T_0 - \lambda I)^{d_j+1} = \dim \ker (T_0 - \lambda I)^{d_j}$.
Note that d_j = the size of the largest Jordan block associated to λ_j. If $\lambda_1, \ldots, \lambda_\ell$ are the distinct eigenvalues of T_0 (so ℓ can be $< n$ if the eigenvalues are multiple roots of $\det(T - \lambda I) = 0$ in some cases), then the minimal polynomial of T_0, call it $P(t)$, is

$$P(t) = \prod_j (t - \lambda_j)^{d_j}.$$

This is clear from looking at the Jordan form. Note that this may be quite different from $\det(T_0 - tI)$ as a polynomial in t, the "characteristic polynomial", which is

$$\prod_{i=1}^{n}(t-\lambda_i)$$

where the λ_is are listed multiply if they occur multiply! The obvious example is $\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}$ any $\lambda \neq 0$, which has characteristic polynomial $(t-\lambda)^2$ but minimal polynomial $t-\lambda$. Of course the minimal polynomial divides the characteristic polynomial (since $P(T_0)=0$ if P is the characteristic polynomial: the Cayley Hamilton Theorem. — or one can see this directly by looking at how many time $(t-\lambda_j)$ occurs in the minimal polynomial: d_j is always $\leq$ the multiplicity of λ_j as a root of the characteristic polynomial — check this!)

Now let us specialize these ideas to the transformations $T^{\mathbb{C}}$ which arise from $T: \mathbb{R}^n \to \mathbb{R}^n$ as before. In this situation, the Jordan form of $T^{\mathbb{C}}$ is special in its nature: First of all, it may have some real blocks, if $T^{\mathbb{C}}$ (and hence T) has real eigenvalues. Note that such real-eigenvalue blocks can be taken to be obtained from real basis vectors. They exist for T itself! The reason is essentially that $\ker(T^{\mathbb{C}} - \lambda I)^k \subset \mathbb{C}^n$, λ real, is the "complexification" of $\ker(T - \lambda I)^k (\subset \mathbb{R}^n)$ since T is real.

Now there may also be blocks associated to eigenvalues λ which are not real. These blocks occur in pairs with the blocks for eigenvalues $\bar{\lambda}$. Namely if

$$v, (T-\lambda I)v, \ldots (T-\lambda I)^{d-1}v$$

(with $(T-\lambda I)^d v = 0$) is the basis of a Jordan λ-block, then

$$\bar{v}, (T-\bar{\lambda} I)\bar{v}, \ldots (T-\bar{\lambda} I)^{d-1}\bar{v}$$

is the basis for a $\bar{\lambda}$-Jordan block.

Now let us

(1) save the real Jordan blocks (real eigenvalues, real basis vectors)

and (2) pair up the complex Jordan blocks (nonreal eigenvalues) as indicated.

T operating "on a pair" has minimal polynomial

$$(t-\lambda)^d \, (t-\bar{\lambda})^d$$

: this means that

$$T \,\big|\, \operatorname{span}(v, (T-\lambda I)v, \ldots (T-\lambda I)^{d-1}v, \bar{v}, (T-\bar{\lambda} I)\bar{v}, \ldots (T-\bar{\lambda} I)^{d-1}\bar{v})$$

has this for a minimal polynomial, as one see immediately.

This "pair subspace" span $(v, (T-\lambda I)v, \ldots, (T-\lambda I)^{d-1}v, \bar{v}, (T-\bar{\lambda} I)\bar{v}, \ldots (T-\bar{\lambda} I)^{d-1}\bar{v})$

has a special property: it is "cylic" in the sense that it is generated by the repeated application of T to one element of it.

Namely $v+\bar{v}$, $Tv+T\bar{v}$, $T^2v+T^2\bar{v}$, ..
$T^{2d-1}v + T^{2d-1}\bar{v}$ are linearly independent.
(since these are real vectors, they are independent over $\mathbb{R}$ if and only if they are independent over $\mathbb{C}$).

To see this, suppose there is a dependence relation. (over $\mathbb{R}$)
This corresponds to a polynomial $P(T)$ with real coefficients such that $\deg P < 2d$ and $P(T)\ (v+\bar{v}) = 0$. But then
$P(T)v = 0$ and $P(T)\bar{v} = 0$ since
$P(T)v \in$ span $(v, (T-\lambda I)v, \ldots (T-\lambda I)^{d-1}v)$
(since this span is T-invariant)

and
$P(T)\bar{v} \in$ span $(\bar{v}, (T-\bar{\lambda} I)\bar{v}, \ldots (T-\lambda I)^{d-1}\bar{v})$
while span$(v, \ldots) \cap$ span$(\bar{v} \ldots) = \{\vec{0}\}$
since they are $\mathbb{C}$ generalized eigenspace with different eigenvalues, $\bar{\lambda}$ not being $= \lambda$!
But the minimal polynomial of T on the pair subspace has degree $2d$ and in particular the minimal degree of any polynomial in T that

annihilates both v and $\bar{v}$ is $2d$. So degree $P < 2d$ is impossible. ∎

So on the "pair subspace" span$(v, \ldots, \bar{v}, \ldots)$

T has the form, relative to the real basis vectors (basis over $\mathbb{C}$)

$w_1 = v + \bar{v}$, $w_2 = T(v + \bar{v})$, $w_3 = T^2(v + \bar{v})$, ..

$w_{2d-1} = T^{2d-1}(v + \bar{v})$:

$w_1 \to w_2 \quad w_2 \to w_3 \quad \ldots, w_{2d-2} \to w_{2d-1}$ and

$w_{d-1} \to \; -(a_0 w_1 + a_1 w_2 + \ldots + a_{2d-1} w_{2d-1})$

where $\quad P_0(t) = t^{2d} + a_{2d-1} t^{2d-1} + \ldots + a_0$

is the minimal polynomial of T on the pair subspace, namely $(t - \lambda)^d (t - \bar{\lambda})^d$.

An example will help to make this clear:

Let $A = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix}$ so there are eigenvalues $\pm i$

for $T_A^{\mathbb{C}}$, with eigenvectors

$$(1, i) \to (i, -1) = i(1, i)$$

and $\quad (1, -i) \to (-i, -1) = -i(1, -i)$

So with $v = (1, i) \quad \bar{v} = (1, -i)$, $\quad v + \bar{v} = (2, 0)$.

$T(2,0) = (0, -2) \quad T^2((2,0)) = T(0, -2)$
$= (-2, 0)$.

So $\quad T^2 + 1 = 0$. The minimal polynomial is $t^2 + 1 = (t - i)(t + i)$ as expected.

Hence $w_1 = (2,0)$ $w_2 = (0,-2)$
and $Tw_1 = v_2$ $Tw_2 = -w_1 = -a_0 w_1$
where $a_0 =$ constant term of the minimal polynomial.
So matrix of $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in this basis.

Definition: If $P(t) = t^\ell + a_{\ell-1}t^{\ell-1} + \dots + a_0$ is a monic polynomial, then the "companion matrix" of P is the $\ell \times \ell$ matrix

$$\begin{pmatrix} 0 & 0 & & -a_0 \\ 1 & 0 & & -a_1 \\ 0 & 1 & \cdots & \vdots \\ \vdots & 0 & & \vdots \\ 0 & 0 & & -a_{\ell-1} \end{pmatrix}$$

In this terminology, the matrix of T on the "pair subspace" with basis $v+\bar{v}, T(v+\bar{v}) \dots, T^{2d-1}(v+\bar{v})$ is the companion matrix of the minimal polynomial $P_0(t) = (t-\lambda)^d (t-\bar{\lambda})^d$.

Now the real Jordan blocks for $T^{\mathbb{C}}$ also can be put in "companion matrix" form. Namely if $v, (T-\lambda I)v, \dots, (T-\lambda I)^{d-1}v$ is a Jordan block basis (so $(T-\lambda I)^d v = 0$), λ real, then the minimal polynomial is $P(t) = (t-\lambda)^d$.

and the vectors $v, Tv, T^2v \dots, T^{d-1}v$

are linearly independent but

$T(T^{d-1}v) = T^d v =$ a linear combination of

$v, Tv \dots T^{d-1}v$

since $(T - \lambda I)^d v = 0$ so $T^d v =$ linear combination of lower powers of T.

Example $\begin{array}{cc} (T-\lambda I)v & v \\ \end{array}$ $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ $\quad w_1 = v, \ (T - \lambda I)v \quad w_2 = Tv$

$(T - \lambda I)^2 = 0$

$Tw_1 = w_2 \qquad T^2 w_1 = T^2 v$

$= 2\lambda Tv - \lambda^2 v$

$= 2\lambda w_2 - \lambda^2 w_1$

So matrix relative to w_1, w_2 is

$$\begin{pmatrix} 0 & -\lambda^2 \\ 1 & 2\lambda \end{pmatrix}$$

Check: Eigenvalues of $\begin{pmatrix} 0 & -\lambda^2 \\ 1 & 2\lambda \end{pmatrix}$ are

t-solutions of $\begin{pmatrix} -t & -\lambda^2 \\ 1 & 2\lambda - t \end{pmatrix} = t^2 - 2\lambda t + \lambda^2$

$= (t - \lambda)^2$ or λ with multiplicity 2.

The "companion matrix" form $\begin{pmatrix} 0 & -\lambda^2 \\ 1 & 2\lambda \end{pmatrix}$

looks a bit stranger than the

Jordan block form $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ but they

are "similar" matrices — same transformation, different bases.

Now we can see what the situation is for a general linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ with matrix A (relative to the standard basis). The equation $\det(A - \lambda I_n) = 0$ will have, say, real solutions (distinct) $\mu_1, \dots \mu_p$ and distinct complex nonreal solutions in conjugate pairs $\lambda_1, \bar{\lambda}_1, \lambda_2, \bar{\lambda}_2, \dots \lambda_q, \bar{\lambda}_q$.
The minimal polynomial of T has real coefficients of course (min. poly. of $T: \mathbb{R}^n \to \mathbb{R}^n$ is real by nature, by definition), but this polynomial can be expressed in terms of the possibly not real situation.

Namely it is

$$P_0(t) = (t-\mu_1)^{D_1} \dots (t-\mu_p)^{D_p} \cdot [(t-\lambda_1)(t-\bar{\lambda}_1)]^{d_1} \dots [(t-\lambda_q)(t-\bar{\lambda}_q)]^{d_q}$$

where D_j = the size of the largest Jordan block for $T^{\mathbb{C}}$ with eigenvalue μ_j while d_k = the size of the largest Jordan block for eigenvalue λ_j (= size of largest Jordan block for eigenvalue $\bar{\lambda}_j$).
Note that $(t-\lambda_j)(t-\bar{\lambda}_j)$ has real coefficients, so this representation of $P_0(t)$ corresponds to the factorization of $P_0(t)$ as a real polynomial into powers of irreducible (prime power) factors which are relatively prime pairwise.

As a real matrix, A (= matrix of T) now has a "canonical form" which is a real matrix of a specific sort. Namely there are blocks down the diagonal, all of which are "companion matrices" either of $[(t-\lambda_j)(t-\bar{\lambda}_j)]^\ell$, $\ell \le d_j$ $j=1,\dots q$ or of $(t-\mu_j)^\ell$, $\ell \le D_j$ $j=1,\dots p$.

And for each $j=1,\dots q$ at least one block is the companion matrix of $[(t-\lambda_j)(t-\bar{\lambda}_j)]^{d_j}$ and for each $j=1,\dots p$, at least one block is the companion matrix of $(t-\mu_j)^{D_j}$.

Two matrices are "similar" iff and only if this so called "rational canonical form" is the same for each, up to obvious permutations.

Example: $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ $a^2+b^2=1$, $b \neq 0$.
minimal polynomial $\lambda^2 - 2a\lambda + 1$, companion matrix $\begin{pmatrix} 0 & -1 \\ 1 & 2a \end{pmatrix}$. Rational canonical form basis

$w_1 = (1,0)$ $w_2 = (a,-b)$ $Tw_1 = w_2$ $Tw_2 = (a^2-b^2, -2ab)$

So $T^2 w_1 = (a^2-b^2, -2ab) = -w_1 + 2a\, w_2$

(Check $w_1 - 2a\,w_2 = -(a^2+b^2, 0) + 2a(a,-b) = (a^2-b^2, -2ab)$)

Note that $\begin{pmatrix} 0 & -1 \\ 1 & 2a \end{pmatrix}$ does indeed have minimal polynomial $-\lambda(2a-\lambda)+1 = \lambda^2 - 2a + 1$, just as the orginal matrix did. It works!

Some similar ideas work for $T: V \to V$
V finite dimensional over a (general) field F, or $T: F^n \to F^n$
To get Jordan form, one needs to pass to
the "splitting field" E of the characteristic
polynomial of T ($= \det(A - \lambda I_n)$ if A
= the matrix of T relative to the standard basis
of F^n). This is a finite (algebraic) extension
of F. Then one can go from the Jordan form
over E back to rational canonical form
over F by combining Jordan blocks associated
by the action of the Galois group of E over
F (in the same way we paired complex items with
their complex conjugates). We omit the details
of this process but trust that the essential
idea is clear.