4-14-1994

The minimal polynomial.

Let V be a finite dimensional F-vector space. A linear transformation $T : V \to V$ is called a **linear operator**.

Let

$$f(x) = a_n x^n + \cdots a_1 x + a_0 \in F[x].$$

You can plug T into $f(x)$ in the obvious way:

$$f(T) = a_n T^n + \cdots a_1 T + a_0 I.$$

For this to make sense, I must insert the identity operator $I : V \to V$ in the last term.

The composite of linear transformations is a linear transformation, the sum of linear transformations is a linear transformation, and a scalar times a linear transformation is a linear transformation. Hence, $f(T)$ is a linear operator $V \to V$.

You can do the same thing with an $n \times n$ matrix $A \in M_n(F)$:

$$f(A) = a_n A^n + \cdots a_1 A + a_0 I.$$

This time, I denotes the $n \times n$ identity matrix.

Lemma. Let $T : V \to V$ be a linear operator on a finite dimensional F-vector space V.

$$\operatorname{Ann} T = \{f \in F[x] \mid f(T) = 0$$

is an ideal in $F[x]$. ($f(T) = 0$ means that $f(T)$ is the zero operator on V.)

$\operatorname{Ann} T$ is called the **annihilator** of T.

Proof. Since $0 \in \operatorname{Ann} T$, $\operatorname{Ann} T \neq \emptyset$.

Suppose $f, g \in \operatorname{Ann} T$, and $h \in F[x]$. Since $f(T) = g(T) = 0$,

$$(f - g)(T) = f(T) - g(T) = 0.$$

Therefore, $f - g \in \operatorname{Ann} T$, and $\operatorname{Ann} T$ is a subgroup of the additive group of $F[x]$.

Moreover,

$$(h \cdot f)(T) = h(T) f(T) = 0.$$

So $h \cdot f \in \operatorname{Ann} T$ and $\operatorname{Ann} T$ is closed under multiplication by ring elements. Hence, $\operatorname{Ann} T$ is an ideal. □

Remark. Essentially the same proof shows that if $A \in M_n(F)$, then

$$\operatorname{Ann} A = \{f \in F[x] \mid f(A) = 0$$

is an ideal in $F[x]$.

If V and W are F-vector spaces, then $\operatorname{Hom}_F(V, W)$ is the F-vector space of linear transformations from V to W.

Proposition. Let V be an n dimensional vector space over F. Then $\operatorname{Hom}_F(V, V) \approx M_n(F)$.

Proof. Choose a basis $\mathcal{B}$ for V. There are maps $\phi : \operatorname{Hom}_F(V, V) \to M_n(F)$, $\psi : M_n(F) \to \operatorname{Hom}_f(V, V)$ given by

$$\phi(T) = [T]_{\mathcal{B},\mathcal{B}}, \qquad \psi(A) = \{v_{\mathcal{B}} \mapsto A v_{\mathcal{B}}\}.$$

The maps are obviously inverses; it is easy to check that ϕ is a linear transformation. □

Proposition. Let V be an n dimensional vector space over F. Then $\dim \operatorname{Hom}_F(V, V) = \dim M_n(F) = n^2$.

Proof. The n^2 matrices E^{ij} given by

$$\left(E^{ij}\right)_{kl} = \delta_{ik}\delta_{jl}$$

form a basis for $M_n(F)$ over F. (E^{ij} has a 1 in the i-j-th position and 0's elsewhere.) □

Lemma. Let $T : V \to V$ be a linear operator on an n dimensional F-vector space V. Then $\operatorname{Ann} T \neq 0$.

Proof. Consider the set

$$\{I, T, T^2, \ldots\} \subset \operatorname{Hom}_F(V, V).$$

Since $\dim \operatorname{Hom}_F(V, V) = n^2$, this set is dependent, and there is a nontrivial finite linear combination of the T^i which equals 0. Consider such a linear combination:

$$a_0 I + a_1 T + \cdots + a_n T^n = 0.$$

Then $a_0 + a_1 x + \cdots a_n x^n \in \operatorname{Ann} T$ is a nonzero element of the annihilator of T. □

Let $T : V \to V$ be a linear operator on an n dimensional F-vector space V. The previous lemma shows that $\operatorname{Ann} T \neq 0$, so since $F[x]$ is a PID, $\operatorname{Ann} T = \langle f(x) \rangle$, where $f(x) \neq 0$.

It is evident by the Division Algorithm that $f(x)$ is a nonzero element of smallest degree in $\operatorname{Ann} T$. Dividing by the leading coefficient if necessary produces a generator $m(x)$ for $\operatorname{Ann} T$ which is **monic** — that is, which has leading coefficient 1.

Definition. A monic polynomial of minimal degree in $\operatorname{Ann} T$ is called a **minimal polynomial** for T.

Observe that being monic implies that the minimal polynomial is nonzero.

Lemma. The minimal polynomial is unique.

Proof. Suppose

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$
$$g(x) = x^n + a_{n-1}x^{n-1} + \cdots + b_1 x + b_0,$$

are minimal polynomials for T.

Now $(f - g)(T) = f(T) - g(T) = 0$, and $\deg(f - g) \leq n - 1$. This is a polynomial of degree $< n$ which kills T, and I can make it monic if necessary by dividing it by its leading coefficient. This contradicts minimality. Hence, $f - g = 0$ and $f = g$. □

By the lemma, I'm justified in speaking of *the* minimal polynomial for an operator T.

Everything I just did works for matrices as well. Thus, the **minimal polynomial** of a matrix $A \in M_n(F)$ is the monic generator of $\operatorname{Ann} T$.

Example. Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}.$$

You can check by direct computation that $A^2 - 2A - 3I = 0$. If A was killed by a polynomial of the form $x - k$, then $A - kI = 0$ would imply that A was a multiple of I. Since this isn't the case, the minimal polynomial of A is $x^2 - 2x - 3$. □

Remark. If you know *one* polynomial $f(x)$ which kills A, then the minimal polynomial for A must divide $f(x)$. This problem is easy modulo factoring polynomials. The **Cayley-Hamilton theorem** produces a polynomial which kills A — namely, the characteristic polynomial of A. So finding the minimal polynomial is no harder (or easier) than plugging A into divisors of the characteristic polynomial.

Lemma. Similar matrices have the same minimal polynomial.

Proof. If $A \in M_n(F)$ and $P \in M_n(F)$ is invertible, then $\operatorname{Ann} A = \operatorname{Ann} PAP^{-1}$. For

$$a_n A^n + \cdots + a_1 A + a_0 I = 0 \quad \Longleftrightarrow \quad a_n (PAP^{-1})^n + \cdots + a_1 (PAP^{-1}) + a_0 I = 0.$$

Since the annihilators are equal, they obviously have the same monic generators. □

I could have proceeded by defining minimal polynomials for matrices, *then* defining the minimal polynomial of an operator to be the minimal polynomial of a matrix which represents the operator. The preceding lemma shows that this definition would have made sense.

Notice that there is a *logical difference* between defining the minimal polynomial of an operator in this way as opposed to defining it as I did (to be the minimal monic generator of $\operatorname{Ann} T$). Nevertheless, one would hope that the two definitions agree. The next result asserts that they do.

Lemma. The minimal polynomial of a linear operator $T : V \to V$ is the same as the minimal polynomial of $[T]_{\mathcal{B},\mathcal{B}}$, where $\mathcal{B}$ is a basis for V.

Proof. Observe that

$$[T]_{\mathcal{B},\mathcal{B}} v_{\mathcal{B}} = [Tv]_{\mathcal{B}}.$$

Hence,

$$[T]^2_{\mathcal{B},\mathcal{B}} v_{\mathcal{B}} = [T]_{\mathcal{B},\mathcal{B}} [Tv]_{\mathcal{B}} = [T^2 v]_{\mathcal{B}}.$$

Inductively, it follows that

$$[T]^n_{\mathcal{B},\mathcal{B}} v_{\mathcal{B}} = [T^n v]_{\mathcal{B}}.$$

In words, the matrix of T^n is the matrix of T to the n-th.

The lemma now follows from the equivalence of the following statements:

$$\begin{aligned}
T^n + \cdots + a_1 T + a_0 I = 0 &\iff \\
T^n v + \cdots + a_1 Tv + a_0 v = 0 &\quad \text{for all v} \iff \\
[T^n v]_{\mathcal{B}} + \cdots + a_1 [Tv]_{\mathcal{B}} + a_0 v_{\mathcal{B}} = 0 &\quad \text{for all v} \iff \\
[T]^n_{\mathcal{B},\mathcal{B}} v_{\mathcal{B}} + \cdots + a_1 [T]_{\mathcal{B},\mathcal{B}} v_{\mathcal{B}} + a_0 v_{\mathcal{B}} = 0 &\quad \text{for all v} \iff \\
\left([T]^n_{\mathcal{B},\mathcal{B}} + \cdots + a_1 [T]_{\mathcal{B},\mathcal{B}} + a_0\right) v_{\mathcal{B}} = 0 &\quad \text{for all v} \iff \\
[T]^n_{\mathcal{B},\mathcal{B}} + \cdots + a_1 [T]_{\mathcal{B},\mathcal{B}} + a_0 = 0. \;\square &
\end{aligned}$$

Example. Let $\dim V = n$, and let $T : V \to V$ be a linear operator. Suppose the minimal polynomial for T is $m(x) = x^n$.

$T^{n-1} \neq 0$, else x^{n-1} would be a polynomial of smaller degree than $m(x)$ killing T. Therefore, there is a vector $v \in V$ such that $T^{n-1} v \neq 0$. Obviously, this implies that $v, Tv, T^2 v, \ldots, T^{n-2} v \neq 0$ as well.

I claim that $\{v, Tv, T^2 v, \ldots, T^{n-1} v\}$ is independent. Suppose that

$$c_0 v + c_1 Tv + \cdots + c_{n-1} T^{n-1} v = 0.$$

Apply T^{n-1} to both sides. Since $T^n = 0$, all the terms but the first die, leaving $c_0 T^{n-1} v = 0$. Now $T^{n-1} v \neq 0$, so $c_0 = 0$.

By applying $T^{n-2}, T^{n-3}, \ldots$ to the equation, it follows in similar fashion that $c_i = 0$ for all i. Therefore, $\{v, Tv, T^2 v, \ldots, T^{n-1} v\}$ is independent.

Since $\dim V = n$, $\mathcal{B} = \{v, Tv, T^2 v, \ldots, T^{n-1} v\}$ is in fact a basis. What is the matrix of T relative to $\mathcal{B}$? Observe that

$$Tv = Tv, \quad T(Tv) = T^2 v, \ldots.$$

So

$$[T]_{\mathcal{B},\mathcal{B}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \quad \square$$

This is a special case of the following construction.

Definition. Let

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a monic polynomial of degree n. The **companion matrix** of p is

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}.$$

Example. The companion matrix of $x^3 - 5x^2 + 4x - 7$ is

$$\begin{bmatrix} 0 & 0 & 7 \\ 1 & 0 & -4 \\ 0 & 1 & 5 \end{bmatrix}. \quad \square$$

Companion matrices are the building blocks of the **rational canonical form**. A rational form matrix is a matrix constructed by "stringing" companion matrices down the main diagonal. Important fact: Every (square) matrix is similar to a unique rational form matrix.

Proposition. Let

$$m(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a monic polynomial, and let A be the companion matrix of m. Then $m(x)$ is the minimal polynomial of A, and the characteristic polynomial of A is $(-1)^n m(x)$.

Proof. To show m is the minimal polynomial for T, I'll first show that $m(T) = 0$.

Let $T : F^n \to F^n$ be the linear operator induced by A:

$$Tv = Av.$$

Let e_i denote the i-th standard basis vector. Then

$$\begin{aligned} Te_1 &= e_2, \\ Te_2 &= e_3, \\ &\vdots \\ Te_{n-1} &= e_n, \\ Te_n &= -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n. \end{aligned}$$

In particular,

$$\{e_1, Te_1, T^2e_1, \ldots, T^{n-1}e_n\} = \{e_1, e_2, e_3, \ldots, e_n\}.$$

Since $\{e_1, Te_1, T^2e_1, \ldots, T^{n-1}e_n\}$ is a basis — in fact, the standard basis — I may write any $v \in F^n$ as

$$v = c_0e_1 + c_1Te_1 + \cdots + c_{n-1}T^{n-1}e_1.$$

Let $f(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$. Then

$$v = f(T)e_1.$$

However,

$$Te_n = -a_0e_1 - a_1e_2 - \cdots - a_{n-1}e_n,$$

$$T^ne_1 = -a_0e_1 - a_1Te_1 - \cdots - a_{n-1}T^{n-1}e_1,$$

$$0 = T^ne_1 + a_{n-1}T^{n-1}e_1 + \cdots + +a_1Te_1 + a_0e_1,$$

$$0 = m(T)e_1.$$

Consequently, if $v \in F^n$, then

$$m(T)v = m(T)f(T)e_1 = f(T)m(T)e_1 = 0.$$

Since $m(T)$ kills everything in V, it's the zero operator: $m(T) = 0$.

Next, I must show that no polynomial of degree less than the degree of m kills T. Suppose $\deg g = n-1$ and $g(T) = 0$. Let

$$g(x) = g_{n-1}x^{n-1} + \cdots + g_1x + g_0.$$

Since $g(T)e_1 = 0$,

$$g_{n-1}T^{n-1}e_1 + \cdots + g_1Te_1 + g_0e_1 = 0.$$

This contradicts the independence of $\{e_1, Te_1, T^2e_1, \ldots, T^{n-1}e_n\}$. Therefore, there is no such g, and m is indeed the minimal polynomial of T.

To compute the characteristic polynomial of A, expand $|A - \lambda I|$ by cofactors of the n-th column. The cofactor corresponding to $-a_i$ is

$$(-1)^{n+i+1}\begin{bmatrix} -\lambda & 0 & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix},$$

where the matrix has i $-\lambda$'s on the main diagonal followed by $n-i$ 1's.

Multiplying by $-a_i$ and expanding, I get

$$(-1)^{n+i+1}(-a_i)(-\lambda)^i = (-1)^n a_1 \lambda^i.$$

The sum of all the terms is exactly $(-1)^n m(x)$. □