
In this lecture we look at a property possessed by some linear transformations which makes their behaviour very easy to understand, and obtain a useful characterization of the transformations having this property. We begin by obtaining a suitable definition of the minimal polynomial of a matrix.

Let $A \in \mathbb{M}_n(\mathbb{F})$. Via matrix multiplication, A is a linear transformation of the row space $\mathbb{F}^n$. Let the MINIMAL POLYNOMIAL OF A, denoted $m_A(x)$, be the minimal polynomial of that transformation. Note that, if $\mathcal{B}_e$ is the elementary basis of $\mathbb{F}^n$, then we are really defining $m_A(x)$ to be $m_{T(A)}(x)$, where $T(A)$ is the transformation of $\mathbb{F}^n$ such that ${}_{\mathcal{B}_e}A_{T(A)} = A$. Note further that if we choose a different basis relative to which to represent T, we obtain a different matrix; what should be the minimal polynomial of this new matrix? In order to be a useful definition, the minimal polynomials of the two matrices should be equal. Our first result confirms that this is, indeed, the case.

(Diag1) Lemma. *If $A, A' \in \mathbb{M}_n(\mathbb{F})$ are similar, then $m_A(x) = m_{A'}(x)$.*

Proof. Suppose that $A \sim A'$. Then there exists an invertible matrix P such that $A' = PAP^{-1}$. For any $f(x) \in \mathbb{F}[x]$ observe that $f(PAP^{-1}) = Pf(A)P^{-1}$. We have $v.m_A(x) = vm_A(A) = 0$ for all $v \in \mathbb{F}^n$. Fix $v \in V$ and consider

$$vm_A(A') \;=\; vm_A(PAP^{-1}) \;=\; vPm_A(A)P^{-1} \;=\; ((vP)m_A(A))P^{-1} \;=\; 0P^{-1} \;=\; 0.$$

Thus $m_A | m_{A'}$. The result now follows by symmetry. □

You might want think about whether or not the converse holds: is it true that matrices having the same minimal polynomial similar? We can now restate **(Dec3)** in terms of matrices.

(Diag2) Theorem. *Let $m_A(x) = p_1(x)^{n_1} \dots p_k(x)^{n_k}$ be the unique factorization of the minimal polynomial of $A \in \mathbb{M}_n(\mathbb{F})$. Then, for each $1 \le i \le k$, there exists a unique sequence*

$$n_i = n_{i1} \ge n_{i2} \ge \dots \ge n_{im_i} \ge 1$$

of natural numbers and a set $A_{i1}, \dots, A_{im_i}$ of square matrices $A_{ij} \in \mathbb{M}_{n_{ij}}(\mathbb{F})$ such that, for some invertible matrix P,

$$PAP^{-1} \;=\; \mathrm{diag}(A_{11}, \dots, A_{1m_1}, \dots, A_{k1}, \dots, A_{km_k}),$$

and the minimal polynomial of A_{ij} is $p_i(x)^{n_{ij}}$. $\square$

There is an analogous matrix formulation of **(Dec4)** which I leave for you to write down. What would be the simplest possible form for the matrix PAP^{-1} above? A rather vague question, but a pleasing possibility is that the matrices A_{ij} are all 1×1 matrices, in which case $PAP^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ($\lambda_i \in \mathbb{F}$) is DIAGONAL. We shouldn't expect this to occur very often but it is worth a little effort to figure out exactly when it does. We call a matrix A DIAGONALIZABLE if there exists an invertible matrix P such that PAP^{-1} is a diagonal matrix. Equivalently we will call a linear transformation T diagonalizable if there exists a basis $\mathcal{B}$ such that ${}_{\mathcal{B}}A_T$ is diagonal.

Examples.

1. If $T \in \mathrm{End}_{\mathbb{F}}(V)$, where V is an n-dimensional $\mathbb{F}$-vector space, and $m_T(x) = x - \lambda$ is linear, then T is diagonalizable. Indeed, if $\mathcal{B}$ is *any* basis of V, then

$$ {}_{\mathcal{B}}A_T \;=\; \mathrm{diag}(\lambda, \ldots, \lambda) = \lambda I_n, $$

where I_n is the $n \times n$ identity matrix. For, if $v \in V$ is any vector, then $0 = v.m_T(x) = v.(x - \lambda) = v.x - \lambda v = vT - \lambda v$, whence $vT = \lambda v$.

2. Let's revisit an example we looked at in the previous lecture. Let

$$ B \;=\; \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. $$

Then we saw that $\mathbb{F}^3$ has a B-cyclic decomposition $\mathbb{F}^3 = \mathrm{sp}(e_2) \oplus \mathrm{sp}(e_1, e_3)$. Furthermore, we have $e_2B = e_2$ and $e_3B = 2e_3$; if we could find $v \in \mathrm{sp}(e_1, e_3) \setminus \mathrm{sp}(e_1)$ and $\lambda \in \mathbb{F}$ with $vB = \lambda v$, then we will have shown that B is diagonalizable. Put $v := e_1 + \alpha e_3$ and compute $vT = (e_1 + \alpha e_3)B = e_1 + 2e_3 + 2\alpha e_3 = e_1 + (2 + 2\alpha)e_3$. In order that $vT = \lambda v$, we must have $\lambda = 1$. In this case, we must also have $\alpha e_3 = (2 + 2\alpha)e_3$, so that $\alpha = -2$. We have shown that $(e_1 - 2e_2)B = e_1 - 2e_2$, so that B is, indeed, diagonalizable. In fact, putting

$$ P \;:=\; \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, $$

we have $PAP^{-1} = \mathrm{diag}(1, 1, 2)$.

3. Consider the matrix

$$C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Here we have a C-cyclic decomposition $\mathbb{F}^3 = \mathrm{sp}(e_1, e_2) \oplus \mathrm{sp}(e_3)$ such that $e_2C = e_2$ and $e_3C = 2e_3$. Suppose we play the same game and try to find $v = e_1 + \alpha e_2 \in \mathrm{sp}(e_1, e_2)$ and $\lambda \in \mathbb{F}$ with $vC = \lambda v$. Then $(e_1 + \alpha e_2)C = e_1 + 2e_2 + \alpha e_2 = e_1 + (2 + \alpha)e_2 = \lambda(e_1 + \alpha e_2)$. Once again we must have $\lambda = 1$, but now we have $2 + \alpha = \alpha$, which is absurd. It turns out that C is not diagonalizable. Look closely at the matrices B and C and try to distinguish the <u>essential</u> difference between them.

Let's have a look at the minimal polynomials in the three examples above. In example 1 we observed in general that, if $m_T(x)$ is linear, then T is diagonalizable. For the matrix B in example 2, we have calculated earlier that $m_B(x) = (x - 1)(x - 2)$. A similiar computation with the matrix C in example 3 reveals that $m_B(x) = (x-1)^2(x-2)$. Examples 1 and 2 provided examples of diagonalizable transformations; example 3 did not. What is the common thread?

(Diag3) Theorem. *A matrix $A \in \mathbb{M}_n(\mathbb{F})$ is diagonalizable if and only if its minimal polynomial $m_A(x)$ factors as a product of distinct linear factors*

$$m_A(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k).$$

Proof. ($\Rightarrow$) Suppose that A is diagonalizable, and let P be an invertible matrix such that $A' = PAP^{-1} = \mathrm{diag}(\lambda_1 I_{m_1}, \lambda_2 I_{m_2}, \ldots, \lambda_k I_{m_k})$, where the λ_i are distinct scalars and the m_i are integers. A simple induction confirms that

$$(A' - \lambda_1 I_n)(A' - \lambda_2 I_n) \ldots (A' - \lambda_k I_n) = 0.$$

It follows that the element $f(x) = (x - \lambda) \ldots (x - \lambda_k) \in \mathbb{F}[x]$ is divisible by $m_{A'}(x) = m_A(x)$ and hence that $m_A(x)$ factors in $\mathbb{F}[x]$ as the product of distinct linear polynomials..

($\Leftarrow$) Suppose that $m_A(x) = (x - \lambda_1) \ldots (x - \lambda_k)$ where the λ_i are distinct. Then, by **(Diag2)**, for each $1 \leq i \leq k$, there exists a unique sequence

$$1 = n_{i1} = n_{i2} = \ldots = n_{im_i} = 1$$

and a list $\alpha_{i1}, \ldots, \alpha_{im_i}$ of scalars (1×1 matrices) such that

$$PAP^{-1} = \mathrm{diag}(\alpha_{11}, \ldots, \alpha_{1m_1}, \ldots, \alpha_{k1}, \ldots, \alpha_{km_k})$$

for some invertible P. Hence A is diagonalisable. Furthermore, since $p_i(x) = (x - \lambda_i)$ is the minimal polynomial of each 1×1 matrix $[[\alpha_{ij}]]$, it follows that $\lambda_i = \alpha_{i1} = \ldots = \alpha_{im_i}$. □

Exercises.

1. Find the number of similarity classes in $\mathbb{M}_5(\mathbb{Q})$ having minimal polynomial $(x-1)(x-2)$. What about $\mathbb{M}_6(\mathbb{Q})$? Can you find some sort of formula for the number in $\mathbb{M}_n(\mathbb{Q})$?
2. Show that the matrix
$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$
is diagonalizable in $\mathbb{M}_3(\mathbb{C})$ but not in $\mathbb{M}_3(\mathbb{R})$.
3. Show that the formal derivative operator $D \colon W \to W$, defined in **(Dec)**, Example 1, is not diagonalizable.
4. If G is a group, then an *involution* in G is any element $g \in G$ of order 2 (i.e. $g^2 = \mathrm{id}_G$). Show that each involution in $\mathrm{GL}_n(\mathbb{F})$ (the group of invertible $n \times n$ matrices) is diagonalizable.
5. Equip $\mathbb{R}^n$ with the inner product $(v, w) = v \cdot w = \sum_{i=1}^n v_i w_i$. We say that a basis $e_1, \ldots, e_n$ is ORTHONORMAL if $(e_i, e_j) = \delta_{ij}$ for $1 \le i, j \le n$. Let $T \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$ be diagonalizable. Show that there is an orthonormal basis $\mathcal{B}$ of $\mathbb{R}^n$ such that ${}_{\mathcal{B}}A_T$ is a lower triangular matrix.