Advanced Linear Algebra

Week 12

Jamie Gabe

SDU

DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Definition

Let $A \in \mathrm{End}(V)$. For each $\lambda \in \mathcal{F}$ we call

$$V_\lambda := \{x \in V \mid Ax = \lambda x\}$$

the eigenspace corresponding to λ.
If $V_\lambda \neq \{0\}$, λ is an eigenvalue of A.
In this case, all non-zero $x \in V_\lambda$ are called eigenvectors for A (corresponding to the eigenvalue λ).

Consider $A\colon \mathbb{C}^2 \to \mathbb{C}^2$ given by $A\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 3\alpha \\ \beta \end{pmatrix}$.

Question: What can be said about $x = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$?

(1) $x \in V_1$;
(2) $x \in V_2$;
(3) $x \in V_3$;
(4) x is not an eigenvector for A.

SDU DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Definition

Let $A \in \mathrm{End}(V)$. For each $\lambda \in \mathcal{F}$ we call

$$V_\lambda := \{x \in V \mid Ax = \lambda x\}$$

the eigenspace corresponding to λ.
If $V_\lambda \neq \{0\}$, λ is an eigenvalue of A.
In this case, all non-zero $x \in V_\lambda$ are called eigenvectors for A (corresponding to the eigenvalue λ).

Consider $A\colon \mathbb{C}^2 \to \mathbb{C}^2$ given by $A\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 3\alpha \\ \beta \end{pmatrix}$.

Question: What can be said about $x = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$?

$$Ax = \begin{pmatrix} 3 \cdot 2 \\ 0 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 0 \end{pmatrix} = 3x$$

so $x \in V_3$.

Definition

Let $A \in \mathrm{End}(V)$. For each $\lambda \in \mathcal{F}$ we call

$$V_\lambda := \{x \in V \mid Ax = \lambda x\}$$

the eigenspace corresponding to λ.
If $V_\lambda \neq \{0\}$, λ is an eigenvalue of A.
In this case, all non-zero $x \in V_\lambda$ are called eigenvectors for A (corresponding to the eigenvalue λ).

Note: Since $Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$ it follows that $V_\lambda = N(A - \lambda I)$.
Hence V_λ is a subspace of V.
Note: If $x \in V_\lambda$ then $Ax = \lambda x \in V_\lambda$. Hence $A(V_\lambda) \subseteq V_\lambda$.
We say that V_λ is an invariant subspace for A.

Definition

When V is finite-dimensional, we call

$$\sigma(A) := \{\lambda \in \mathcal{F} \mid \lambda \text{ is an eigenvalue of } A\}$$

the spectrum of A.

The (geometric) multiplicity of $\lambda \in \sigma(A)$ is $\dim V_\lambda$.

$A \in \mathrm{End}(V)$ (with $\dim V < \infty$) is diagonalisable if V has a (ordered) basis $\mathcal{B} = (x_1, \dots, x_n)$ consisting of eigenvectors for A. In this case, $Ax_i = \lambda_i x_i$, where λ_i is the corresponding eigenvalue.

Hence the $\mathcal{B}$-coordinates for Ax_i are λ_i in the i'th coordinate, and 0 in all other coordinates.

So ${}_{\mathcal{B}}[A]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Recall that a polynomial over $\mathcal{F}$ is an expression

$$p(X) = \sum_{i=0}^{n} \alpha_i X^i = \alpha_0 + \alpha_1 X + \cdots + \alpha_n X^n$$

where $n \geq 0$, $\alpha_0, \ldots, \alpha_n \in \mathcal{F}$.
The symbol X is called the indeterminate.
The set of all polynomials over $\mathcal{F}$ is denoted $\mathcal{F}[X]$.
$\mathcal{F}[X]$ is an $\mathcal{F}$-vector space with basis $\{1_{\mathcal{F}}, X, X^2, X^3, \ldots\}$.
Question: What is the dimension of $\mathcal{F}[X]$?

(1) 1

(2) n

(3) it depends on $p(X)$

(4) ∞.

Answer: $\{1_{\mathcal{F}}, X, X^2, X^3, \ldots\}$ is a basis with infinitely many elements, so $\dim \mathcal{F}[X] = \infty$.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

We can also multiply polynomials over $\mathcal{F}$: if $p(X) = \sum_{i=0}^{n} \alpha_i X^i$ and $q(X) = \sum_{j=0}^{m} \beta_j X^j$ then

$$pq(X) := \sum_{k=0}^{n+m} \sum_{i+j=k} (\alpha_i \beta_j) X^k.$$

Note: this is just the usual formula for multiplying polynomials

$$(\sum_{i=0}^{n} \alpha_i X^i)(\sum_{j=0}^{m} \beta_j X^j) = \sum_{k=0}^{n+m} \sum_{i+j=k} (\alpha_i \beta_j) X^k.$$

In this way $\mathcal{F}[X]$ is actually an algebra over $\mathcal{F}$: an $\mathcal{F}$-vector space with a multiplication (bilinear map) $\mathcal{F}[X] \times \mathcal{F}[X] \to \mathcal{F}[X]$. Note

$$\underbrace{\overbrace{\text{addition} + \text{scalar multip.}}^{\text{vector space over } \mathcal{F}} + \text{multip.}}_{\text{algebra over } \mathcal{F}}$$

$\mathcal{F}[X]$ is actually an associative and commutative algebra.

Let $p(X) = \sum_{i=0}^{n} \alpha_i X^i \in \mathcal{F}[X]$. There is an induced function $p\colon \mathcal{F} \to \mathcal{F}$ (which we also denote by p) given by

$$p(\gamma) = \sum_{i=0}^{n} \alpha_i \gamma^i.$$

Let $p(X) \in \mathbb{R}[X]$ with induced function $p\colon \mathbb{R} \to \mathbb{R}$.
Question: is the function $p\colon \mathbb{R} \to \mathbb{R}$ linear?

(1) Yes;

(2) No;

(3) It depends on the polynomial $p(X)$.

Answer: Are real polynomial functions linear? Some are (e.g. $p(X) = X$) and some are not (e.g. $p(X) = X^2$). So the answer is (3).

Note that if $p(X) = \sum_{i=0}^{n} \alpha_i X^i$, $q(X) = \sum_{j=0}^{m} \beta_j X^j \in \mathcal{F}[X]$, with induced functions $p, q \colon \mathcal{F} \to \mathcal{F}$, then

$$p(\gamma)q(\gamma) = \sum_i \alpha_i \gamma^i \sum_j \beta_j \gamma^j = \sum_k \sum_{i+j=k} \alpha_i \beta_j \gamma^k = (pq)(\gamma).$$

What is cooler than evaluating polynomials in elements of $\mathcal{F}$? Evaluating in endomorphisms!
E.g. if $p(X) = X + X^2$, and $A \in \mathrm{End}(V)$, I can think of $A + A^2$ as being $p(A)$. More generally:
For $p(X) = \sum_{i=0}^{n} \alpha_i X^i \in \mathcal{F}[X]$ and $A \in \mathrm{End}(V)$ we define

$$p(A) := \sum_{i=0}^{n} \alpha_i A^i \in \mathrm{End}(V),$$

where $A^i = A \circ \cdots \circ A$ ($i \geq 1$ times) and $A^0 = I_V$ (by convention).

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

For $p(X) = \sum_{i=0}^{n} \alpha_i X^i \in \mathcal{F}[X]$ and $A \in \mathrm{End}(V)$ we define

$$p(A) := \sum_{i=0}^{n} \alpha_i A^i \in \mathrm{End}(V)$$

Lemma (6.5)

Let $A \in \mathrm{End}(V)$ be fixed. The map $\mathcal{F}[X] \to \mathrm{End}(V)$ given by $p(X) \mapsto p(A)$ is an algebra homomorphism (i.e. it is linear and preserves multiplication).

This looks more complicated than it is.

Example: if $p(X) = X$ and $q(X) = X^2$, then $(p+q)(X) = X + X^2$. Also $p(A) = A$ and $q(A) = A^2$. Then

$$(p+q)(A) = A + A^2 = p(A) + q(A).$$

The general statement: $(p+q)(A) = p(A) + q(A)$ means that the map in the lemma above is additive.

Similarly, $(\alpha p)(A) = \alpha p(A)$ means it preserves scalar multiplication, and $(pq)(A) = p(A)q(A)$ means it is multiplicative.

For $p(X) = \sum_{i=0}^{n} \alpha_i X^i \in \mathcal{F}[X]$ and $A \in \mathrm{End}(V)$ we define

$$p(A) := \sum_{i=0}^{n} \alpha_i A^i \in \mathrm{End}(V)$$

Lemma (6.5)

Let $A \in \mathrm{End}(V)$ be fixed. The map $\mathcal{F}[X] \to \mathrm{End}(V)$ given by $p(X) \mapsto p(A)$ is an algebra homomorphism (i.e. it is linear and preserves multiplication).

Consider the case $p(X) = X$ and $q(X) = X + X^3$.

Question: What is $(pq)(A)$?

(1) A

(2) $A + A^3$

(3) $A^2 + A^4$.

Answer: $(pq)(X) = p(X)q(X) = X(X + X^3) = X^2 + X^4$. So $(pq)(A) = A^2 + A^4$.

For $p(X) = \sum_{i=0}^{n} \alpha_i X^i \in \mathcal{F}[X]$ and $A \in \mathrm{End}(V)$ we define

$$p(A) := \sum_{i=0}^{n} \alpha_i A^i \in \mathrm{End}(V)$$

Lemma (6.5)

Let $A \in \mathrm{End}(V)$ be fixed. The map $\mathcal{F}[X] \to \mathrm{End}(V)$ given by $p(X) \mapsto p(A)$ is an algebra homomorphism (i.e. it is linear and preserves multiplication).

Proof.

If $p(X) = \sum_i \alpha_i X^i, q(X) = \sum_j \beta_j X^j$ then $p(A) + q(A) = \sum_i \alpha_i A^i + \sum_j \beta_j A^j = \sum_i (\alpha_i + \beta_i) A^i = (p+q)(A)$. Hence the map is additive. Similarly it preserves scalar multiplication.
$p(A)q(A) = (\sum_i \alpha_i A^i)(\sum_j \beta_j A^j) = \sum_k \sum_{i+j=k} \alpha_i \beta_j A^k = (pq)(A)$
so the map is multiplicative.

SDU
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Recall, (given $A \in \mathrm{End}(V)$) that $V_\lambda = \{x \in V : Ax = \lambda x\}$.

Lemma

Suppose $x \in V_\lambda$ and $p(X) \in \mathcal{F}[X]$. Then $p(A)x = p(\lambda)x$. So if x is an eigenvector for A corresponding to λ, then x is an eigenvector for $p(A)$ corresponding to $p(\lambda)$.

Proof.

We first show that $A^n x = \lambda^n x$ for all $n \geq 0$. First, $A^0 x = Ix = \lambda^0 x$. So it holds for $n = 0$. It also holds for $n = 1$ by definition, and by induction

$$A^n x = A(A^{n-1}x) = A(\lambda^{n-1}x) = \lambda^{n-1}Ax = \lambda^{n-1}(\lambda x) = \lambda^n x.$$

Now, if $p(X) = \sum_i \alpha_i X^i$, then

$$p(A)x = \sum_i \alpha_i A^i x = \sum_i \alpha_i \lambda^i x = p(\lambda)x. \qquad \square$$

SDU
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE

Suppose $p(X) \in \mathcal{F}[X]$. An element $\gamma \in \mathcal{F}$ is a root of p if $p(\gamma) = 0$.
Let $R(p) \subseteq \mathcal{F}$ be the set of all roots of p.
Recall: $\sigma(A) \subseteq \mathcal{F}$ is the set of eigenvalues of A.

Lemma (6.6)

Let $A \in \mathrm{End}(V)$ and $p(X) \in \mathcal{F}[X]$ such that $p(A) = 0$. Then $\sigma(A) \subseteq R(p)$.

Proof.

Let $\lambda \in \sigma(A)$ and let $x \in V_\lambda$ be non-zero. By the previous lemma, we have

$$p(\lambda)x = p(A)x = 0x = 0.$$

As $x \neq 0$ it follows that $p(\lambda) = 0$. $\square$

Lemma (6.6)

Let $A \in \mathrm{End}(V)$ and $p(X) \in \mathcal{F}[X]$ such that $p(A) = 0$. Then $\sigma(A) \subseteq R(p)$.

We work over $\mathcal{F} = \mathbb{R}$. Consider $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{End}(\mathbb{R}^2)$.

One easily checks that $A^2 = -I$, so if $p(X) = 1 + X^2$ then $p(A) = I + A^2 = I - I = 0$.

Question: What can we conclude?

(1) $i \in \sigma(A)$;

(2) $i \in R(p)$;

(3) $1 \in R(p)$;

(4) $\sigma(A) = \emptyset$.

Answer: $p(X) = 1 + X^2$ has no real roots, so $R(p) = \emptyset$. By Lemma 6.6 we have $\sigma(A) \subseteq R(p) = \emptyset$, so $\sigma(A) = \emptyset$.

Definition

Let $\dim V < \infty$ and $A \in \mathrm{End}(V)$. A minimal polynomial for A is a non-zero polynomial $p(X) \in \mathcal{F}[X]$ such that $p(A) = 0$, and which has minimal degree among all such polynomials.

Note: for constant polynomials $p(X) = \alpha$ we have $p(A) = \alpha I_V$. Hence (unless $V = \{0\}$) a minimal polynomial is never constant.

Consider $A = \left(\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right) \in \mathrm{End}(\mathbb{R}^2)$, and $p(X) = 1 + X^2$ so that $p(A) = 0$. If $q(X) = \alpha + \beta X$, then

$$q(A) = \alpha I + \beta A = \left(\begin{array}{cc} \alpha & \beta \\ -\beta & \alpha \end{array} \right).$$

Question: Is $p(X)$ a minimal polynomial for A?

(1) Yes
(2) No
(3) There is not enough information to determine this.

Definition

Let $\dim V < \infty$ and $A \in \mathrm{End}(V)$. A minimal polynomial for A is a non-zero polynomial $p(X) \in \mathcal{F}[X]$ such that $p(A) = 0$, and which has minimal degree among all such polynomials.

Consider $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{End}(\mathbb{R}^2)$, and $p(X) = 1 + X^2$ so that $p(A) = 0$. If $q(X) = \alpha + \beta X$, then

$$q(A) = \alpha I + \beta A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Question: Is $p(X)$ a minimal polynomial for A?
Answer: $p(X)$ has degree 2. Any polynomial of degree ≤ 1 $q(X) = \alpha + \beta X$ satisfies $q(A) = 0$ only when $\alpha = \beta = 0$. Hence $p(X)$ is minimal.

Lemma (6.8)

Let $\dim V < \infty$ *and* $A \in \mathrm{End}(V)$*. There exists a minimal polynomial for* A*, and if* $p(X)$ *is a minimal polynomial, then* $\sigma(A) = R(p)$*.*

Proof.

Existence: Let $n := \dim V$.

It is enough to prove that $p(A) = 0$ for a non-zero $p(X)$. Since $\dim \mathrm{End}(V) = n^2$, the $n^2 + 1$ vectors $A^0, A^1, \ldots, A^{n^2}$ cannot be linearly independent. Hence $\sum_{i=0}^{n^2} \alpha_i A^i = 0$ for some non-trivial linear combination. Hence $p(A) = 0$ for $p(X) = \sum_{i=0}^{n^2} \alpha_i X^i$. $\square$

Lemma (6.8)

Let $\dim V < \infty$ *and* $A \in \mathrm{End}(V)$. *There exists a minimal polynomial for* A, *and if* p *is a minimal polynomial, then* $\sigma(A) = R(p)$.

Proof.

Now suppose $p(X)$ is a minimal polynomial for A. We already know $\sigma(A) \subseteq R(p)$.

Let $\lambda \in R(p)$. By the division algorithm for polynomials (Algebra 1), $p(X) = (X - \lambda)q(X)$ for some polynomial $q(X)$ of degree less than $p(X)$. By minimality of the degree of $p(X)$, we know that $q(A) \neq 0$. Let $x \in V$ such that $y := q(A)x \neq 0$. Then

$$(A - \lambda I)y = (A - \lambda I)q(A)x = p(A)x = 0.$$

Hence y is an eigenvector for A with eigenvalue λ, and thus $\lambda \in \sigma(A)$. Hence $R(p) \subseteq \sigma(A)$.

A field $\mathcal{F}$ is called algebraically closed if every non-constant polynomial $p(X) \in \mathcal{F}[X]$ has at least one root, i.e. there exists $\lambda \in \mathcal{F}$ such that $p(\lambda) = 0$.

The fundamental theorem of algebra states that $\mathbb{C}$ is algebraically closed.

Suppose $\lambda \in \mathcal{F}$ is a root of $p(X)$. By polynomial division there is a polynomial $q(X) \in \mathcal{F}[X]$ with degree one less than $p(X)$, such that

$$p(X) = (X - \lambda)q(X).$$

By induction on the degree it follows that if $\mathcal{F}$ is algebraically closed then every polynomial of degree $k \geq 1$ can be factored as

$$p(X) = \alpha(X - \lambda_1)(X - \lambda_2)\cdots(X - \lambda_k).$$

Moreover, α and the scalars $\lambda_1, \ldots, \lambda_k$ are unique (up to permutation of the indices). Each λ_i can occur multiple times, this number is called the multiplicity of the root.

Theorem (6.9)

Assume that $\mathcal{F}$ is algebraically closed and $0 < \dim V < \infty$. Let $A \in \mathrm{End}(V)$. Then $\sigma(A) \neq \emptyset$. In particular, there exists an eigenvector $x \in V$ for A.

Proof.

By Lemma 6.8, there exists a minimal polynomial $p(X)$ for A, and $\sigma(A) = R(p)$. As $p(X)$ is non-constant and as $\mathcal{F}$ is algebraically closed $p(X)$ has a root. In other words, $\sigma(A) = R(p) \neq \emptyset$. □