Linear Algebra Short Course

Lecture 3

Matthew J. Holland

`matthew-h@is.naist.jp`

Mathematical Informatics Lab
Graduate School of Information Science, NAIST

Some useful references

- Polynomial basics: Axler (1997, Ch. 4)
- Eigenvalue/vector basics: Axler (1997, Ch. 5,10), Horn and Johnson (1985, Ch. 1)
- More advanced results: Axler (1997, Ch. 8-9)

Lecture contents

1. **Invariant spaces and eigenvalues/vectors**
2. **Key questions and foundational results**
3. **More involved topics**

Lecture contents

1. **Invariant spaces and eigenvalues/vectors**
2. **Key questions and foundational results**
3. **More involved topics**

Important idea 1: "powers" of linear operators

Linear operators have important properties not shared by linear maps in general.

For $S \in \mathcal{L}(U, V)$ for $U \neq V$, as

$$S(u) \in V, S(S(u)) \text{ is not defined!}$$

Thus the "product" SS is meaningless.

(*) For operator $T \in L(U)$ however,

$$T^m(u) := \underbrace{(T \cdots T)}_{m\text{-product}}(u)$$

is both defined and indeed $T^m \in \mathcal{L}(U)$.

Important idea 2: invariant sets

Defn. For $T \in \mathcal{L}(U)$, we call subset $E \subset U$ **invariant under** T whenever

$$T(w) \in E, \forall\, w \in E.$$

Of course, for T-invariant $E \subset V$, any $w \in E$ is s.t.

$$T^m(w) \in E,\ m > 0.$$

Thus, we naturally consider the "dynamics" induced by T, namely

$$w_{(k)} := T^k(w),\ k = 0, 1, 2, \ldots$$

with $T^0 := I$.

Comment. These notions are fundamental to *ergodic theory*:

{ergodic theory} $\approx$ {dynamical systems} $\cap$ {measure/prob. theory}.

Some examples

Example. (*) One may readily note for any $T \in \mathcal{L}(U)$,

- U and $\{0\}$ are T-invariant
- $\operatorname{range} T$ and $\operatorname{null} T$ are T-invariant

(*) Let $T \in \mathcal{L}(\mathcal{P}_k(\mathbb{R}))$ be $(Tp)(\cdot) := p'(\cdot)$. Then $\mathcal{P}_l(\mathbb{R})$ is T-invariant for any $l \leq k$.

(*) Say subspace $W \subset U$ of $\dim W = 1$ is T-invariant. Note we may always find a $w_0 \in W$ s.t.

$$T(w_0) = \alpha w_0,$$

an equation that should foreshadow our next topic.

Eigenvalues/vectors of operators

The vectors which are only *scaled* by a given operator (if they exist) can tell us a great deal about the operator (as we'll see).

Assume vector space U on $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$) with $\dim U < \infty$.

Defn. Given $T \in \mathcal{L}(U)$, if $\alpha \in \mathbb{F}$ is s.t.

$$Tu = \alpha u$$

for some $u \neq 0$, then we call α an **eigenvalue** of T.
Call $\sigma(T) := \{\alpha \in \mathbb{F} : \alpha \text{ an eigenvalue of } T\}$ the **spectrum** of T.

For any $\alpha \in \sigma(T)$, if $Tv = \alpha v$, we call v an **eigenvector** associated with α.

The "dynamics" of T initiated at eigenvector u are easy:

$$T^m(u) = \alpha^m u,$$

all we need is to know α.

Basic properties and facts 1

(*) If $\alpha \in \sigma(T)$ with eigenvector u, then βu is also an eigenvector associated with α, for any $\beta \in \mathbb{F}$.

(*) We now can develop our previous remarks further, as

$$\exists\, W \subset U, \dim W = 1, T\text{-invariant} \iff T \text{ has an eigenvalue},$$

a nice characterization.

(*) A useful fact is that

$$\alpha \in \sigma(T) \iff (T - \alpha I) \in \mathcal{L}(U) \text{ is non-invertible},$$

and equivalently non-injective, non-surjective.

(*) Using this, note for any $\alpha \in \sigma(T)$,

$$\{\text{eigenvectors of } T \text{ wrt } \alpha\} \text{ is a subspace of } U.$$

Basic properties and facts 2

Do distinct eigenvalues have some special meaning?

(**) Let $\alpha_1, \ldots, \alpha_m \in \sigma(T)$ be *distinct*, i.e., $\alpha_i \neq \alpha_j$, $i \neq j$. Take any corresponding non-zero eigenvectors $u_1, \ldots, u_m$. Then,

$$\{u_1, \ldots, u_m\} \text{ is linearly independent.}$$

(*) Thus, number of distinct eigenvalues of any $T \in \mathcal{L}(U)$ is controlled by the dimension, namely $|\sigma(T)| \leq \dim U$.

As we work through basic results, a number of natural questions arise...

Lecture contents

1. Invariant spaces and eigenvalues/vectors
2. Key questions and foundational results
3. More involved topics

Lecture contents

1. Invariant spaces and eigenvalues/vectors
2. Key questions and foundational results
3. More involved topics

Questions to motivate our theory

Fix some $T \in \mathcal{L}(U), \dim U < \infty$.

If and when do eigenvalues of T exist?

How many distinct/multiple eigenvalues does T have?

What (useful) information on T does spectrum $\sigma(T)$ encode?

Is spectral information shared between an operator and its matrix representations?

We shall seek (at least partial) answers to all of these questions.

Constructing operator polynomials

Let $T \in \mathcal{L}(U)$. Recalling our definition of $T^m := T \cdots T$ ($m > 0$ times) and $T^0 := I$, for the case of invertible T we add

$$T^{-m} := (T^{-1})^m.$$

(*) We then have for $m, n \geq 0$,

$$T^{m+n} = T^m T^n, \quad (T^m)^n = T^{mn}.$$

Now, take $p \in \mathcal{P}_m(\mathbb{F})$, a function $p : \mathbb{F} \to \mathbb{F}$ taking the form

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m.$$

Defn. Given $T \in \mathcal{L}(U)$, we define for every $p \in \mathcal{P}_m(\mathbb{F})$ the map

$$p(T) := a_0 I + a_1 T + \cdots + a_m T^m.$$

(*) Clearly $p(T) \in \mathcal{L}(U)$.

Properties of operator polynomials

The key utility: we can factor $p(T)$ just as we can factor p!

(*) To see this, verify for $p, q \in \mathcal{P}_m(\mathbb{F})$, we have

$$(pq)(T) = p(T)q(T).$$

(*) From this $q(T)p(T) = p(T)q(T)$ follows.

Example. We know polynomial $p \in \mathcal{P}_m(\mathbb{C})$ factors as

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

for $c \neq 0$ where λ_i are the roots of p (up to multiplicity). Defining $q_i(z) := (z - \lambda_i)$, $i = 1, \ldots, m$ and $q_0(z) := c$, we have that $p = q_0 q_1 \cdots q_m$. Thus,

$$\begin{aligned} a_0 I + a_1 T + \cdots + a_m T^m &= p(T) \\ &= (q_0 q_1 \cdots q_m)(T) \\ &= q_0(T) \cdots q_m(T) \\ &= c(T - \lambda_1 I) \cdots (T - \lambda_m I). \end{aligned}$$

Existence of eigenvalues

We may now (partially) answer our first question.

(**) For finite-dim U on field $\mathbb{C}$, for any $T \in \mathcal{L}(U)$,

$$\sigma(T) \neq \emptyset.$$

i.e., all complex linear operators have an eigenvalue.

(*) This says for operator T on complex vector space U, $\exists$ basis B s.t.

$$M(T;B) = \begin{bmatrix} \lambda & & * & \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

This observation highlights an important theme we look at shortly.

Remark. Note we have made no reference to determinants thus far.

Spectral properties of operators and their matrices

The question of "shared" spectral information is easy to answer.

(*) Take $T \in \mathcal{L}(U)$ and any basis B. Then,

$$\alpha \in \sigma(T) \iff \alpha \in \sigma(M(T;B)),$$

that is eigenvalues of T coincide with those of *any* matrix representation of T, thus spectral info encoded.

Simple matrix representations and invariant sets 1

We've seen $M(T;B)$ encodes information on T.
Matrices with a simple structure are "easier to decode" than others.

Goal: choose B so that $M(T;B)$ is "simple."

What properties of T determine whether this can be done or not?

Defn. We call $A \in \mathbb{F}^{n\times n}$ **upper-triangular** when $a_{ij} = 0$ for all $i > j$.

(*) Let $T \in \mathcal{L}(U)$, with basis $B = \{u_1, \ldots, u_n\}$. The following are equivalent:

- A $M(T;B)$ is upper-triangular.
- B $T(u_j) \in [\{u_1, \ldots, u_j\}]$, for $j = 1, \ldots, n$.
- C $[\{u_1, \ldots, u_j\}]$ is T-invariant, for $j = 1, \ldots, n$.

(*) Clearly, if T has upper-tri representation, $\sigma(T)$ is non-empty.

Simple matrix representations and invariant sets 2

Do we always have an upper-tri representation?
If U is on $\mathbb{C}$, then yes.

(**) For any $T \in \mathcal{L}(U)$, $\exists\, B$ s.t.

$$M(T; B) \text{ is upper-triangular.}$$

The proof is straightforward, using an induction argument on dimension of U, and fact that $\sigma(T)$ non-empty.

(*) For $T \in \mathcal{L}(U)$, let $A := M(T; B)$ be upper-tri. Then,

$$T \text{ is invertible} \iff a_{ii} \neq 0, i = 1, \ldots, n.$$

Thus we may trivially read off A whether or not T is invertible.
Proof best done using contraposition (both ways).

Simple matrix representations and invariant sets 3

Operator eigenvalues also can be read off upper-tri representations.

(*) For $T \in \mathcal{L}(U)$, $\dim U = n$, let B be such that $A := M(T;B)$ is upper-tri. Then,

$$\sigma(T) = \{a_{11}, \ldots, a_{nn}\}.$$

Just investigate $M(T - \alpha I; B)$ for any $\alpha \in \sigma(T)$.

Defn. An even simpler special case is that of **diagonal** matrix representations, where $a_{ij} = 0$, all $i \neq j$.

(*) We may readily note for $T \in \mathcal{L}(U)$ and basis $B = \{u_1, \ldots, u_n\}$,

$$M(T;B) \text{ is diagonal } \iff u_i \text{ are eigenvectors of } T.$$

Simple matrix representations and invariant sets 4

Existence of diagonal representations is much stronger than that of upper-triangularity:

Example. (*) Consider $T : \mathbb{C}^2 \to \mathbb{C}^2$ defined $T(z) := (z_2, 0)$. Show 0 is the only valid eigenvalue of $T \in \mathcal{L}(\mathbb{C}^2)$, and that

$$\{z \in \mathbb{C}^2 : z_2 = 0\}$$

is precisely the set of eigenvectors of T, to conclude T has no diagonal representation.

Thus, even on $\mathbb{C}$, cannot always diagonalize.

On $\mathbb{R}$, *even* upper-tri representations need not exist. . .

Characterizing diagonalizability

(*) From our previous results, we may easily note for $T \in \mathcal{L}(U)$ and $\dim U = n$ that

$$|\sigma(T)| = n \implies \exists\, B \text{ s.t. } M(T; B) \text{ is diagonal.}$$

This is not a necessary condition, however (stronger than needed).

(**) Let $\dim U = n$, and let $\alpha_1, \ldots, \alpha_m \in \sigma(T)$ be distinct, $0 \leq m \leq n$. The following are equivalent:

- A $\exists\, B$ s.t. $M(T; B)$ is diagonal.
- B $\exists\, B = (u_1, \ldots, u_n)$ s.t. all u_i are eigenvecs of T.
- C Exists subspaces $U_1, \ldots, U_n$, all T-invariant and $\dim U_i = 1$, where $U = U_1 \oplus \cdots \oplus U_n$.
- D $U = \operatorname{null}(T - \alpha_1 I) \oplus \cdots \oplus \operatorname{null}(T - \alpha_m I)$
- E $\dim U = \dim \operatorname{null}(T - \alpha_1 I) + \cdots + \dim \operatorname{null}(T - \alpha_m I)$

Clearly, *need not have* all distinct eigenvalues.

Existence of eigenvalues in the real case 1

The $\mathbb{F} = \mathbb{R}$ case is less friendly.
(*) For U on $\mathbb{R}$ and $T \in \mathcal{L}(U)$, $\sigma(T)$ may be empty. If so,

$$\implies \text{ exists no } T\text{-invariant } W \subset U \text{ with } \dim W = 1.$$

This trickiness is closely related to the existence of roots of real polynomials:

(*) Recall that $p(x) := x^2 + ax + b$, for $a, b \in \mathbb{R}$, is such that

$$p(x) = (x - \alpha_1)(x - \alpha_2), \alpha_1, \alpha_2 \in \mathbb{R} \iff a^2 \geq 4b.$$

(**) As well, for non-constant $p \in \mathcal{P}(\mathbb{R})$, have unique factorization

$$p(x) = c(x - \alpha_1) \cdots (x - \alpha_m)(x^2 + a_1 x + b_1) \cdots (x^2 + a_M x + b_M),$$

for $m, M \geq 0$ where $c, \alpha_i \in \mathbb{R}$, $(a_i, b_i) \in \mathbb{R}^2$, and $a_i^2 < 4b_i$.

Existence of eigenvalues in the real case 2

We use these basic facts to consider the real case.

In moving from $\mathbb{C}$ to $\mathbb{R}$, existence statements for invariant subspaces must be weakened:

(**) Take U over $\mathbb{R}$, $1 \leq \dim U = n$. Then, $\exists\, W \subset U$ s.t.

$$W \text{ is } T\text{-invariant and } 1 \leq \dim W \leq 2.$$

The problem of course:
dim-2 invariant subspaces need not imply an eigenvalue.

Existence of eigenvalues in the real case 3

While not ideal, the previous result allows us to prove a neat fact:

(**) Take U on $\mathbb{R}$. Then,

$$\dim U \text{ is odd} \implies \forall\, T \in \mathcal{L}(U),\ \sigma(T) \neq \emptyset.$$

The base case of $\dim U = 1$ is trivial; an induction argument on the dimension of U proves this.

Lecture contents

1. Invariant spaces and eigenvalues/vectors
2. Key questions and foundational results
3. More involved topics

Lecture contents

1. Invariant spaces and eigenvalues/vectors
2. Key questions and foundational results
3. **More involved topics**

Answering some more subtle questions

Recall that our previous question,

How many distinct/multiple eigenvalues does T have?

has not been answered yet. As well, related questions such as

Is there a "middle ground" between upper-triangular and diagonal representations?

Can we define "invariants" of T using intrinsic information?

still require answers.

With some effort, we can answer these questions.

Generalized eigenvectors 1

Take $T \in \mathcal{L}(U)$, $\dim U = n$.
Let $\alpha_1, \ldots, \alpha_m$ be distinct eigenvalues of T. Even if $m < n$, recall

T "has enough eigenvectors"
(= exists eigenvecs $u_1, \ldots, u_n$ s.t. $\{u_1, \ldots, u_n\}$ a basis of U)

$\implies$ T is "as nice as possible."

That is, we require that the bases of the $m \leq n$ "eigenspaces"

$$\text{null}(T - \alpha_i I) = \{u \in U : Tu = \alpha_i u\}, \quad i = 1, \ldots, m$$

furnish a basis of U.

In general (non-diagonal cases), we don't have enough linearly independent eigenvectors. Thus "ideal case" structural results fail.

Generalized eigenvectors 2

For more general cases, the following generalization is key.

Defn. Fix $\alpha \in \mathbb{F}$. Say $u \in U$ is a **generalized eigenvector** of T if

$$(T - \alpha I)^k(u) = 0,$$

for some integer $k > 0$.

(*) If holds for $u \neq 0$, then note that any such α is $\alpha \in \sigma(T)$.

Example. (*) Let $T \in \mathcal{L}(\mathbb{C}^3)$ be $T(z) := (z_1, 0, z_3)$. Show $0, 1 \in \sigma(T)$ and that letting $\alpha_1 := 0, \alpha_2 := 1$, have

$$\mathbb{C}^3 = \mathrm{null}(T - \alpha_1 I)^2 \oplus \mathrm{null}(T - \alpha_2 I)$$

This example foreshadows an important general result (shortly).

Nilpotent operators

Defn. We call $S \in \mathcal{L}(U)$ **nilpotent** if $S^k = 0$ for finite $k > 0$.

(*) Say S is nilpotent on U. Then, have

$$\{0\} = \operatorname{null} S^0 \subset \operatorname{null} S \subset \operatorname{null} S^2 \subset \cdots ,$$

and if have $\operatorname{null} S^m = \operatorname{null} S^{m+1}$ for some m, then

$$\implies \operatorname{null} S^m = \operatorname{null} S^l, \; \forall\, l \geq m.$$

(*) Fortunately we always reach such a limit, as

$$\operatorname{null} S^{\dim U} = \operatorname{null} S^{\dim U + 1} = \cdots .$$

(*) In fact, for nilpotent $S \in \mathcal{L}(U)$ we always have

$$S^{\dim U} = 0.$$

Nilpotent operators and gen'd eigenvectors

(*) From these basic results it is not hard to verify for $\alpha \in \sigma(T)$ that

$$\mathrm{null}(T - \alpha I)^{\dim U} = \{\text{all gen'd eigenvectors of } T \text{ wrt } \alpha\}.$$

Using some of these ideas, we move forward to a very important topic.

Multiplicities 1

For $T \in \mathcal{L}(U)$, $\dim U = n$, let $A := M(T; B)$ for any basis B s.t. A is upper-tri. Then, we know

$$\sigma(T) = \{a_{11}, \ldots, a_{nn}\},$$

So when $|\sigma(T)| < n$ we must have "multiples" on diagonal of A.

Do the duplicates on diagonal of $M(T; B)$ depend on choice of B?

Can we characterize how many times a given $\alpha \in \sigma(T)$ appears?

An idea: number of multiples of α = $\dim \operatorname{null}(T - \alpha I)$?

Multiplicities 2

Example. (*) Define $T \in \mathcal{L}(\mathbb{C}^2)$ s.t.

$$M(T) = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$$

taken WRT the standard basis. Note eigenvalue 5 appears twice.

However, $\dim \operatorname{null}(T - 5I) = 1$, so the idea fails.

Defn. The dimension of the eigenspace of α is often called the **geometric multiplicity** of $\alpha \in \sigma(T)$.

Multiplicities 3

A very important result with a slightly involved proof gives us the “right” answer for characterization.

(**) Let $T \in \mathcal{L}(U)$ on $\mathbb{F}$, and $\alpha \in \sigma(T)$. For any basis B where $M(T;B)$ is upper-tri,

$$\alpha \text{ repeats on diagonal } \dim \operatorname{null}(T - \alpha I)^{\dim U} \text{ times.}$$

Defn. We'll call the dimension of the “generalized eigenspace” $\operatorname{null}(T - \alpha I)^{\dim U}$ the **(algebraic) multiplicity** of α.

(*) Take U over $\mathbb{C}$ and $T \in \mathcal{L}(U)$ with distinct eigenvalues $\alpha_1, \ldots, \alpha_m$ with multiplicities $d_1, \ldots, d_m$ we pleasantly have

$$d_1 + \cdots + d_m = \dim U.$$

This certainly need not hold for the geometric multiplicities.

An intuitive characteristic polynomial 1

All this talk of multiplicities relates to a key concept: the “characteristic polynomial” of a linear operator.

You may recall from undergraduate LA that fixing $A \in \mathbb{F}^{n \times n}$,

$$\det(\alpha I - A)$$

viewed as a function of $\alpha \in \mathbb{F}$ is often of interest.

This works for square matrices, but there are some issues:

- We want a CP for general $T \in \mathcal{L}(U)$
- We haven’t defined $\det T$ for general operators yet
- **We want a CP defined using intrinsic properties of T**

An intuitive characteristic polynomial 2

We thus define a natural polynomial q_T which encodes all the spectral information of a given T.

Defn. Assume U on $\mathbb{C}$. Let $T \in \mathcal{L}(U)$ with distinct eigenvalues α_i having multiplicities d_i, $i = 1, \ldots, m$. We define the **characteristic polynomial** q_T of T by

$$q_T(z) := (z - \alpha_1)^{d_1} \cdots (z - \alpha_m)^{d_m}.$$

Important: this elegant definition is only for $\mathbb{C}$.

(*) For any $T \in \mathcal{L}(U)$, degree of q_T is $\dim U$. Roots of q_T are precisely the distinct eigenvalues of T.

(**) A straightforward argument shows

$$q_T(T) := (T - \alpha_1 I)^{d_1} \cdots (T - \alpha_m I)^{d_m} = 0,$$

which is the Cayley-Hamilton theorem on $\mathbb{C}$.

What about the case of U on $\mathbb{R}$?

While still doable, working on $\mathbb{R}$ is more troublesome.

The results are all analogous, just less pliable than the $\mathbb{C}$ case.

First, instead of being able to upper-triangularize, we can "upper-block-triangularize":

(**) Let $T \in \mathcal{L}(U)$, U over $\mathbb{R}$. Then exists basis B s.t.

$$M(T;B) = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$

where A_i are *at most* 2×2 real matrices, and if 2×2, then have no eigenvalues.

The "no eigenvalues" quality will be utilized shortly.

Multiplicities in the $\mathbb{R}$ case 1

Recall for $T \in \mathcal{L}(U)$ on $\mathbb{C}$ we had the nice fact that

$$\text{sum of multiplicities of } T\text{'s eigenvals} = \dim U.$$

On $\mathbb{R}$, this cannot possibly hold in general ($\sigma(T) = \emptyset$ possible).

In fact, even if have some eigenvalues, need not hold:

Example. (*) Let $T \in \mathcal{L}(\mathbb{R}^3)$ where (WRT std. basis)

$$M(T) = \begin{bmatrix} 3 & -1 & 2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{bmatrix}.$$

Clearly $1 \in \sigma(T)$, eigenvec $(1, 0, 1)$. With effort, can check that $\sigma(T) = \{1\}$, just one eigenval. However, verify that

$$\dim \operatorname{null}(T - I)^3 = 1 < \dim \mathbb{R}^3.$$

Clearly something is missing.

Multiplicities in the $\mathbb{R}$ case 2

We need a new eigenvalue-like object to fill in the gap. Recall

$$M(T;B) = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

the “block-upper-tri” form is always feasible.

We define a new object using the 2×2 blocks.

How should we do this?
To motivate: let’s extend our CP q_T definition to the $\mathbb{R}$ case.

Extending q_T to the $\mathbb{R}$ case 1

In the $\mathbb{C}$ case, for $T \in \mathcal{L}(U)$ we can always get

$$M(T;B) = \begin{bmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \end{bmatrix},$$

and using the algebraic multiplicities of the α_i, we built q_T as

$$q_T := q_T^{(1)} \cdots q_T^{(n)}$$

using $q_T^{(i)}(z) := (z - \alpha_i)$ as "building blocks."

On $\mathbb{R}$, we need to settle for

$$M(T;B) = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}.$$

If $m = n$ and $A_i = [\alpha_i]$, $i = 1, \ldots, n$ then fine; original q_T works.

Extending q_T to the $\mathbb{R}$ case 2

In general, will have $m \leq n$ and 2×2 block matrices A_i present.

Let's "aim for Cayley-Hamilton."
Start with $\dim U = 1$ case. Trivially, $M(T; B) = [\alpha_1]$, and

$$q_T(T) := q_T^{(1)}(T) := (T - \alpha_1 I) = 0.$$

(*) Next, $\dim U = 2$ case. Let $B = (u_1, u_2)$. Say

$$M(T; B) = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Might naturally check $(z - a)(z - d)$ as a first try. Note that

$$(T - aI)(T - dI)(u_i) = bcu_i, \quad i = 1, 2.$$

Fortunately $(T - aI)(T - dI) - bcI = 0$, yielding a good suggestion.

Extending q_T to the $\mathbb{R}$ case 3

An idea for an extended characteristic polynomial is thus born.
Let A_i, $i = 1, \dots, m$ be upper-tri blocks of $M(T; B)$ for U on $\mathbb{R}$. Let

$$q_T^{(i)}(z) := \begin{cases} (z - a) & \text{if } A_i = [a] \\ (z - a_{11})(z - a_{22}) - a_{21}a_{12} & \text{if } A_i \text{ is } 2 \times 2 \end{cases}$$

and then set $q_T := q_T^{(1)} \cdots q_T^{(m)}$.

Note however, that

The A_i are basis-dependent; is this new q_T well-defined?

Actually, *yes*. We'll see this now.

New object to fill in the multiplicity gap 1

Note that each piece $q_T^{(i)}(T)$ (in operator form) of our improved q_T is

$$\text{either } (T - \alpha I) \text{ or } (T^2 + aT + b).$$

(**) Take any basis B of U such that $M(T;B)$ is block-upper-tri, with blocks $A_1, \ldots, A_m$. Then, for any $\alpha \in \mathbb{R}$,

$$|\{j : A_j = [\alpha]\}| = \dim \operatorname{null}(T - \alpha I)^{\dim U},$$

and for any $a, b \in \mathbb{R}$ where $a^2 < 4b$,

$$|\{j : q_T^{(j)}(z) = z^2 + az + b\}| = \frac{\dim \operatorname{null}(T^2 + aT + bI)^{\dim U}}{2}.$$

It is precisely these special "**eigenpairs**" (a, b) that close the gap...

New object to fill in the multiplicity gap 2

Admittedly, the previous result is a bit subtle.
Let B be basis s.t. $M(T;B)$ in block-upper-tri form (blocks $A_1, \ldots, A_m$).

Regarding "eigenpairs":
(*) First note, if 2×2 matrix A has no eigenvalues, then exists $a, b \in \mathbb{R}$ such that $a^2 < 4b$ and

$$x^2 + ax + b = (x - a_{11})(x - a_{22}) - a_{12}a_{21}.$$

(*) Also, if $a^2 < 4b$, then

$$T^2 + aT + bI \text{ not injective} \iff \dim \operatorname{null}(T^2 + aT + bI)^{\dim U} > 0.$$

Define **eigenpair** to be any (a, b) satisfying left-hand side.

- If (a, b) an eigenpair, then A_j such that $q_T^{(j)}(z) = z^2 + az + b$ will always appear.
- Conversely, for each of the 2×2 matrices A_j, have that $\dim \operatorname{null}(q_T^{(j)}(T))^{\dim U} \geq 2$.

New object to fill in the multiplicity gap 3

Regarding eigenvalues:

- If $\alpha \in \sigma(T)$, then $A_j = [\alpha]$ will always appear for some j.
- Conversely, if $A_j = [\alpha]$, then necessarily $\alpha \in \sigma(T)$.

Important conclusion:

Regardless of what B we take to define q_T, it is the same.

This of course uses the fact that

$$q_T^{(1)} \cdots q_T^{(m)} = q_T^{(\pi_1)} \cdots q_T^{(\pi_m)}$$

under any permutation π.

Thus, our new q_T extended to $\mathbb{R}$ case is valid.

New object to fill in the multiplicity gap 4

Recall extending q_T was a means to a separate end, namely, "filling in the multiplicity gap" on $\mathbb{R}$.

Defn. We analogously define the **multiplicity** of eigenpair (a, b) by

$$\widetilde{d} := \frac{\dim \operatorname{null}(T^2 + aT + bI)^{\dim U}}{2}.$$

(*) Thus for $T \in \mathcal{L}(U)$ with eigenvals $\alpha_1, \ldots, \alpha_m$ and eigenpairs $(a_1, b_1), \ldots, (a_M, b_M)$,

$$\dim U = \sum_{i=1}^{m} d_i + \sum_{j=1}^{M} 2\widetilde{d}_j.$$

Of course $0 \leq m, M \leq \dim U$, but $\max\{m, M\} > 0$.

Return of invariant subspaces: key structural result

To wrap this discussion up, a great result. Using notation from previous slide

$$U_i := \text{null}(T - \alpha_i I)^{\dim U},\ i = 1, \ldots, m$$
$$\widetilde{U}_j := \text{null}(T^2 + a_j T + b_j I)^{\dim U},\ j = 1, \ldots, M.$$

(**) We have for $T \in \mathcal{L}(U)$, U on $\mathbb{R}$, that

$$U = U_1 \oplus \cdots \oplus U_m \oplus \widetilde{U}_1 \oplus \cdots \oplus \widetilde{U}_M$$

and all $U_i, \widetilde{U}_j$ are indeed T-invariant subspaces.

Defining "invariants" of T intrinsically 1

One may note for $T \in \mathcal{L}(U)$ on $\mathbb{F}$, $\dim U = n$, expanding

$$q_T(z) = z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0,$$

we have that in particular c_{n-1} and c_0 take simple forms:

$$\begin{aligned} c_0 &= (-1)^n \alpha_1 \cdots \alpha_m b_1 \cdots b_M \\ c_{n-1} &= (-1)(\alpha_1 + \cdots + \alpha_m - a_1 - \cdots - a_M) \end{aligned}$$

We have in fact found definitions of the **trace** and **determinant** of operator T,

$$\begin{aligned} \operatorname{trace} T &:= \alpha_1 + \cdots + \alpha_m - a_1 - \cdots - a_M \\ \det T &:= \alpha_1 \cdots \alpha_m b_1 \cdots b_M. \end{aligned}$$

It turns out these quantities indeed coincide with their standard matrix counterparts.

Defining "invariants" of T intrinsically 2

(**) It is quite remarkable that for *any* basis B,

$$\operatorname{trace} T = \operatorname{trace} M(T; B)$$
$$\det T = \det M(T; B)$$

which is reassuring, and indeed our characteristic polynomial also is equivalent with the usual definition,

$$q_T(z) = \det(zI - T).$$

Proving these facts isn't difficult. See Axler (1997, Ch. 10).

Finding sparse representations

Recall our final remaining subtle question:

Is there a "middle ground" between upper-triangular and diagonal representations?

Yes, and it's about as sparse as a non-diagonal (but still upper-tri) representation can be.

This is the "Jordan form" of operator T, defined for every T in the complex case.

We have the tools need to prove this, but omit due to time constraints. Nice proofs in Axler (1997, Ch. 8), Horn and Johnson (1985, Ch. 3).

Jordan form

(**) For $T \in \mathcal{L}(U)$, U on $\mathbb{C}$, there exists basis B_J s.t.

$$M(T; B_J) = \begin{bmatrix} J_1 & & * \\ & \ddots & \\ 0 & & J_m \end{bmatrix}$$

where each block matrix J ($k \times k$ for $1 \leq k$) is of the form

$$J = \begin{bmatrix} \alpha_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \alpha_k \end{bmatrix}.$$

Defn. Call such a basis a **Jordan basis** of T, and $M(T; B_J)$ a **Jordan form** of T.

Equivalently, every complex matrix is similar to a Jordan matrix.

Lecture contents

1. **Invariant spaces and eigenvalues/vectors**
2. **Key questions and foundational results**
3. **More involved topics**

Lecture contents

1. **Invariant spaces and eigenvalues/vectors**
2. **Key questions and foundational results**
3. **More involved topics**

References

Axler, S. (1997). *Linear Algebra Done Right*. Springer, 2nd edition.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press, 1st edition.