MATH2040A Week 7 Tutorial Notes

1 Eigenvalues and Eigenvectors

Let V be a vector space, and $T \in L(V)$. A nonzero vector $v \in V$ is an eigenvector / characteristic vector if $Tv = \lambda v$ for some $\lambda \in F$. λ is the associated eigenvalue.

Easy to see that

- $\lambda \in F$ is an eigenvalue if and only if $T - \lambda \mathrm{Id}$ is not injective. If V is finite dimensional, this is equivalent to $T - \lambda \mathrm{Id}$ not being invertible.
- if v is an eigenvector, then v spans a one-dimensional invariant subspace $S = \mathrm{Span}(\ \{ v \}\ )$: $T(S) \subseteq S$

The *spectrum* $\sigma(T) \subseteq F$ of a linear map $T \in L(V)$ is the set of scalars $\lambda \in F$ such that $T - \lambda I$ is not invertible. If V is finite dimensional, it is exactly the set of all eigenvalues of T.

In lecture, the following theorem is proven:

Theorem 1.1. *If V is a finite dimensional vector space of dimension n, and $\lambda \in F$, $T \in L(V)$, then the following are equivalent:*

- *λ is an eigenvalue of T*
- *λ is an eigenvalue of the matrix $[T]_\beta$ in some/every ordered basis β of V*
- *λ is a root of the* characteristic polynomial $f_T(t) = \det([T]_\beta - tI_n)$ *for some/every ordered basis β of V*

The characteristic polynomial $f_T(t) = \det([T]_\beta - tI_n)$ is a polynomial of degree n and leading coefficient $(-1)^n$. Using fundamental theorem of algebra on it implies that a linear map on a finite dimensional vector space over *complex* numbers has exactly n eigenvalues (counted with multiplicity).

To find all eigenvalues and eigenvectors of a linear map on a finite dimensional space, typically you need to

1. compute the characteristic polynomial $f_T(t) = \det([T]_\alpha - tI_n)$ in e.g. standard ordered basis, then solve for all its roots, typically by factorizing it
2. for each eigenvalue λ, find a basis for $\mathsf{N}(\ T - \lambda \mathrm{Id}\ )$, which can be done via working on the matrix representation $[T]_\alpha - \lambda I_n$ in e.g. standard ordered basis

As you can see, finding all eigenvalues and eigenvectors involves only basic (yet cumbersome) operations of

- computing determinant
- factorizing a polynomial
- finding a basis of null space of a matrix (e.g. RREF)

It is quite common to make mistakes in these steps, so my suggestion is to check if your results are correct after each step:

- does the characteristic polynomial have the right degree and leading coefficient?
- are the eigenvalues you found actually roots of the characteristic polynomial?
- is the null space nontrivial?
- is the basis you found consisting of eigenvectors of the original map, associated with the eigenvalue you have chosen to solve for?

If not, there is a mistake somewhere in your computation.

2 Diagonalizability

If there is an (ordered) basis β of V that every element is an eigenvector of T, then T is *diagonalizable.* If V is finite dimensional, it is easy to see that $[T]_\beta$ is a diagonal matrix (with diagonal entries being the eigenvalues), a *canonical form* that is easier to handle than a generic (dense) matrix.

The following theorem has been shown in lecture:

Theorem 2.1. *If $S \subseteq V$ is a (finite) set of eigenvectors each associated to a distinct eigenvalue, then S is linearly independent.*

So if a linear map has $\dim(V)$ distinct eigenvalues, it is diagonalizable.

We will discuss more about diagonalizability in the next tutorial sessions.

3 Algebra of Linear Maps, and Polynomials

We know that $L(V)$ is a vector space with naturally defined addition and scalar multiplication:

- $T + U$ is defined pointwise as $(T + U)(v) = T(v) + U(v)$
- aT is defined pointwise as $(aT)(v) = a \cdot T(v)$

We also know that we can define a (noncommutative but associative) multiplication on $L(V)$ via composition: TU is defined pointwise as $(TU)(v) = T(U(v))$. With product, we can define powers with nonnegative exponents:

- $T^0 = \text{Id}$
- $T^{n+1} = T^n T$

If a linear map $T \in L(V)$ is invertible, we can also define powers with negative exponents: T^{-1} is the inverse of T, and $T^{-n} = (T^{-1})^n$.

It is easy to verify that all these operations combined satisfy *most* of the usual laws you would expect.

This means that given a polynomial $p(X) = \sum a_i X^i \in \mathsf{P}(F) = F[X]$, we can define a linear map $p(T) = \sum a_i T^i$. Informally, this is "evaluating the polynomial at the linear map".

With simple computations, it is easy to see the following properties:

- if $Tv = \lambda v$, then $T^n v = \lambda^n v$ for $n \geq 0$. If T is invertible, this also holds for $n < 0$ if $v \neq 0$
- if p, q are polynomials, then $(p + q)(T) = p(T) + q(T)$ and $(pq)(T) = p(T)q(T)$
- if A is a matrix, then $p(L_A) = L_{p(A)}$ (with $p(A) = \sum a_i A^i$)

Combining these properties, it is easy to see the following result: if v is an eigenvector of T with associated eigenvalue λ, then $p(T)v = p(\lambda)v$, so $p(\lambda)$ is an eigenvalue of $p(T)$. (See also exercise Q4.)

4 Exercises

1. (Textbook Sec. 5.1 Q23)

   Let V be a finite dimensional vector space and $T \in L(V)$ be diagonalizable with characteristic polynomial $p(t)$. Show that $p(T) = 0$.

   Solution: Since T is diagonalizable, there exists an eigenbasis β of V.
   Let $v \in \beta$, and the associated eigenvalue be λ.
   Then λ is a root of p, so $p(\lambda) = 0$.
   By property of eigenvalue, we also have $p(T)v = p(\lambda)v = 0$.
   Since $v \in \beta$ is arbitrary, $p(T) = 0$ on a basis β of V, thus $p(T) = 0$.

Note

You may already be aware of the *Cayley–Hamilton theorem* which is the focus of upcoming / current lectures (the relevant lecture notes are posted online already). This question is a specialized version of the theorem that does not need many setups, at the cost of a strong assumption (diagonalizability).

2. Let V be a finite dimensional vector space, and $T, U \in L(V)$. Show that $\sigma(TU) = \sigma(UT)$.

Solution: By symmetry, it suffices to show that $\sigma(TU) \subseteq \sigma(UT)$.

Let $\lambda \in \sigma(TU)$ be nonzero. Then there exists nonzero eigenvector $v \in V$ such that $TUv = \lambda v$.
Since $v \neq 0$ and $\lambda \neq 0$, $TUv = \lambda v \neq 0$. In particular, $Uv \neq 0$.
As $UTUv = U(\lambda v) = \lambda Uv$ and $Uv \neq 0$, Uv is an eigenvector of UT with eigenvalue λ, so $\lambda \in \sigma(UT)$.

Suppose now $0 \in \sigma(TU)$. Then TU is not invertible.
Take an ordered basis β of V. Then $[TU]_\beta$ is not an invertible matrix.
By the property of determinant (from MATH1030), $0 = \det([TU]_\beta) = \det([T]_\beta[U]_\beta) = \det([U]_\beta[T]_\beta) = \det([UT]_\beta)$.
This implies that $[UT]_\beta$ is also not invertible, and so UT is not invertible, $0 \in \sigma(UT)$.

Thus $\sigma(TU) \subseteq \sigma(UT)$.

3. Let $A \in \mathbb{C}^{n\times n}$. For each j define $R_j = \sum_{k\neq j} | A_{jk} |$, and $D_j = \{ z \in \mathbb{C} : | z - A_{jj} | \leq R_j \}$.
Let $\lambda \in \mathbb{C}$ be an eigenvalue of A. Show that $\lambda \in D_j$ for some j.

Solution: Let $x \in \mathbb{C}^n$ be an eigenvector associated to the eigenvalue λ.
Then $x \neq 0$ and $Ax = \lambda x$, or $\lambda x_j = (Ax)_j = \sum_k A_{jk}x_k = A_{jj}x_j + \sum_{k\neq j} A_{jk}x_k$ for each j.
Rearranging and using triangle inequality, we have $| \lambda - A_{jj} || x_j | \leq \sum_{k\neq j} | A_{jk} || x_k |$ for each j.
Let $j_0 \in \{ 1, \dots, n \}$ be the index such that $| x_{j_0} |$ is maximal. Then $| x_k | \leq | x_{j_0} |$ for all k.
Since $x \neq 0$, $| x_{j_0} | \neq 0$, for otherwise $x_k = 0$ for all k and so $x = 0$, a contradiction.
Then $| \lambda - A_{j_0,j_0} || x_{j_0} | \leq \sum_{k\neq j_0} | A_{j_0,k} || x_k | \leq (\sum_{k\neq j_0} | A_{j_0,k} |) | x_{j_0} | = R_{j_0} | x_{j_0} |$, so $| \lambda - A_{j_0,j_0} | \leq R_{j_0}$.
This implies that $\lambda \in D_{j_0}$.

Note

This is Gershgorin circle theorem, which gives (computationally) quick estimates of locations of eigenvalues. There are a few generalizations, including Brauer oval theorem in which the regions are (Cassini) ovals. See N. Higham's blog post (and the references therein) for more detail.

4. Let V be a finite dimensional vector space over *complex numbers*, $T \in L(V)$, $p(X) = \sum_{j=0}^{n} c_j X^j \in \mathsf{P}(\mathbb{C})$ be a nonzero complex polynomial of degree $n \geq 1$.
Let $p(\sigma(T)) = \{ p(\lambda) \mid \lambda \in \sigma(T) \}$. Show that $p(\sigma(T)) = \sigma(p(T))$.

(Hint: use the fundamental theorem of algebra)

Solution: Let $\lambda \in \sigma(T)$. Then there exists a nonzero vector $v \in V$ such that $Tv = \lambda v$.
So $p(T)v = (\sum c_j T^j)(v) = \sum c_j T^j(v) = \sum c_j \lambda^j v = p(\lambda)v$. This implies $p(\lambda)$ is an eigenvalue of $p(T)$, so $p(\lambda) \in \sigma(p(T))$. As λ is arbitrary, $p(\sigma(T)) \subseteq \sigma(p(T))$.

Let $\lambda \in \sigma(p(T))$. Then there exists a nonzero vector $v \in V$ such that $p(T)v = \lambda v$, so $0 = (p(T) - \lambda \mathrm{Id})v = q(T)v$ on $q(X) = p(X) - \lambda$.

By the fundamental theorem of algebra, there exist $c \in \mathbb{C} \setminus \{0\}$ and $z_1, \ldots, z_n \in C$ such that $q(X) = c(X - z_n) \ldots (X - z_1)$, so $q(T) = c(T - z_n\mathrm{Id}) \ldots (T - z_1\mathrm{Id})$.
Since $q(T)v = 0$ with $v \neq 0$, $q(T)$ is not injective.
As composition of injective linear maps is injective, this means that (at least) one of $T - z_j\mathrm{Id}$ is not injective. Let one such linear map be $T - z_k\mathrm{Id}$. This implies that z_k is an eigenvalue of T.
By assumption, z_k is also a root of $q(X) = p(X) - \lambda$, so $\lambda = p(z_k) \in p(\sigma(T))$.
As λ is arbitrary, $\sigma(p(T)) \subseteq p(\sigma(T))$.

Therefore, $p(\sigma(T)) = \sigma(p(T))$.

Note

Instead of considering composition of injective linear maps, we can also consider the vectors $v_0 = v$, $v_j = (T - z_j\mathrm{Id})v_{j-1}$ for $j \in \{ 1, \ldots, n \}$. Since $v_n = c^{-1}q(T)v = 0$ and $v_0 \neq 0$, there must be one minimal $k \in \{ 1, \ldots, n \}$ such that $v_{k-1} \neq 0$ but $(T - z_k\mathrm{Id})v_{k-1} = 0$, which also implies that z_k is an eigenvalue of T.

This is (a simple version of) spectral mapping theorem (for polynomials), an interesting theorem in functional calculus. The use of fundamental theorem of algebra is critical: consider the *real* vector space $\mathbb{R}^2$ and $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, $p(X) = X^2 + 1 \in \mathsf{P}(\mathbb{R})$. It is easy to see that

- L_A has no (real) eigenvalue, so the (real) spectrum is $\emptyset$
- $p(L_A) = L_{p(A)} = 0_{2\times 2}$ and so has (real) spectrum $\{0\}$

which implies that $p(\sigma(L_A)) = \emptyset \subsetneq \{0\} = \sigma(p(L_A))$