Math 413/513 Chapter 5 (from Friedberg, Insel, & Spence)

David Glickenstein

November 4, 2015

1 Eigenvalues and Eigenvectors

Definition 1 *If V is a finite dimensional vector space, a linear operator $T \in \mathcal{L}(V)$ is* diagonalizable *if there is an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix. A square matrix A is* diagonalizable *if L_A is diagonalizable.*

This maybe requires some basic examples. Notice that

$$T(x, y) = (x + y, 4x + y)$$

does not produce a diagonal matrix in the standard basis

$$[T]_{E_2} = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$$

but if we consider the basis

$$\beta = \{(1, 2), (1, -2)\}$$

we see that

$$T(1, 2) = (3, 6) = 3(1, 2)$$
$$T(1, -2) = (-1, 2) = -(1, -2)$$

so

$$[T]_\beta = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

Hence T is diagonalizable and so is $\begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$.

Notice that the fact that $[T]_\beta$ is diagonal is equivalent to the fact that $T(v_1) = \lambda_1 v_1$ for an ordered basis $\{v_1, \ldots, v_n\}$, where $\lambda_1, \ldots, \lambda_n$ are scalars. The choice of v_i is important here, since if we choose other vectors, T does not have this simple form (in the above example, $T(1, 0) \neq \lambda(1, 0)$ for any choice of scalar λ). This leads to the notion of eigenvalue and eigenvector.

Definition 2 *Let T be a linear operator on a vector space V (not necessarily finite dimensional). A nonzero vector $v \in V$ is called an* eigenvector *of T if there exists a scalar λ such that $T(v) = \lambda v$. The scalar λ is called the* eigenvalue *corresponding to the eigenvector v. Let $A \in F^{n \times n}$. A nonzero vector $v \in F^n$ is called an eigenvector of A if v is an eigenvector of L_A, i.e., $Av = \lambda v$ for some λ. The scalar λ is called the eigenvalue corresponding to the eigenvector v.*

Remark 3 *Some older terms for eigenvalue/eigenvector are characteristic value/characteristic vector and proper value/proper vector.*

Theorem 4 *A vector v is an eigenvector for a linear operator $T \in \mathcal{L}(V)$ corresponding to eigenvalue λ if and only if $v \neq 0$ and $v \in N(T - \lambda I_V)$.*

Theorem 5 *A linear operator T on a finite-dimensional vector space V is diagonalizable if and only if there exists an ordered basis β for V consisting of eigenvectors of T. Furthermore, if T is diagonalizable and β is an ordered basis of eigenvectors of T and $D = [T]_\beta$, then D is a diagonal matrix and the diagonal entries are the corresponding eigenvalues.*

The idea of the proof is essentially the discussion above and is left as an exercise.

Example 6 *Not all matrices have eigenvectors. We can determine rotation by angle θ by left multiplication by the matrix*

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If θ is not an integer multiple of π, then R_θ does not take any nonzero vector to a multiple of itself, and so these matrices have no eigenvectors.

Example 7 *Let C^∞ be the vector space consisting of function $\mathbb{R} \to \mathbb{R}$ with derivatives of all orders (check this is a vector space subspace of the vector space of functions $\mathbb{R} \to \mathbb{R}$). Consider the map $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ given by $T(f) = f'$. For a function to be an eigenvector, we need that $f' = \lambda f$ for some value of λ. There are functions that satisfy this property, namely $f_\lambda(x) = e^{\lambda x}$ is a eigenvector with eigenvalue λ. Note that the constant function is an eigenvector with eigenvalue 0.*

Notice that if v is an eigenvector for a linear operator T with corresponding eigenvalue λ, then $T(v) = \lambda v$ can be written $(T - \lambda I_V)v = 0$ where I_V is the identity transformation on V. Thus there is a nonzero element of the nullspace of the linear operator $T - \lambda I$. For matrices, this allows us to use determinants to find eigenvalues.

Theorem 8 *Let $A \in F^{n \times n}$. The scalar λ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$.*

Proof. If λ is an eigenvalue, its corresponding eigenvector satisfies $(A - \lambda I_n) v = 0$, so there $A - \lambda I_n$ is not invertible, so $\det (A - \lambda I_n) = 0$. Similarly, if $\det (A - \lambda I_n) = 0$, then $A - \lambda I_n$ is not invertible, so there is a nonzero element v in the nullspace of $A - \lambda I_n$, and then $Av = \lambda v$. $\blacksquare$

One can replace λ in the previous theorem to get a polynomial associated with a matrix A.

Definition 9 *Let $A \in F^{n \times n}$. The polynomial $f(t) = \det (A - tI_n)$ is called the* characteristic polynomial *of A.*

Note that the theorem above says that any eigenvalue is a zero of the characteristic polynomial and any zero of the characteristic polynomial is an eigenvalue. We can use this to compute eigenvalues, then eigenvectors, and sometimes even diagonalize the matrix.

Suppose we have a basis of eigenvectors for a square matrix A, then notice that if Q has the elements of the basis as its columns, then $AQ = QD$ where D is the diagonal matrix with the eigenvalues on the diagonal, hence $Q^{-1}AQ = D$.

Example 10 *Consider the matrix*

$$A = \begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}.$$

The characteristic polynomial is

$$(1 - \lambda)(2 - \lambda)(3 - \lambda).$$

Each must correspond to at least one eigenvector, so we need to calculate

$$A - I = \begin{pmatrix} 0 & 4 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 2 \end{pmatrix}$$

and we see that $N(A - I) = \text{span}\{(1, 0, 0)\}$. We calculate

$$A - 2I = \begin{pmatrix} -1 & 4 & 5 \\ 0 & 0 & 6 \\ 0 & 0 & 1 \end{pmatrix}$$

and so $N(A - 2I) = \text{span}\{(4, 1, 0)\}$. We calculate

$$A - 3I = \begin{pmatrix} -2 & 4 & 5 \\ 0 & -1 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

and so

$$N(A - 3I) = \text{span}\{(29, 12, 2)\}.$$

Let's double check that these are, in fact, eigenvectors

$$\begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 29 \\ 12 \\ 2 \end{pmatrix} = \begin{pmatrix} 87 \\ 36 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 29 \\ 12 \\ 2 \end{pmatrix}.$$

One can check that

$$Q = \begin{pmatrix} 1 & 4 & 29 \\ 0 & 1 & 12 \\ 0 & 0 & 2 \end{pmatrix}$$

has inverse

$$Q^{-1} = \begin{pmatrix} 1 & -4 & \frac{19}{2} \\ 0 & 1 & -6 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & -4 & \frac{19}{2} \\ 0 & 1 & -6 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 & 29 \\ 0 & 1 & 12 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

We can see how similar matrices affect the characteristic polynomial.

Proposition 11 *If A is similar to B, that is, there exists an invertible matrix Q such that $A = Q^{-1}BQ$, then the characteristic polynomials of A and B are the same.*

Proof. Exercise. ∎

Recall that if β, β' are ordered bases for a finite-dimensional vector space V, then there exists and invertible matrix Q such that for any $T \in \mathcal{L}(V)$ we have $[T]_\beta = Q^{-1}[T]_{\beta'}Q$. Thus, we can define the characteristic polynomial of a linear transformation.

Definition 12 *Let V be a finite-dimensional vector space and let $T \in \mathcal{L}(V)$. We define the characteristic polynomial of T to be the characteristic polynomial of $[T]_\beta$ for any basis β of V, i.e., the characteristic polynomial is $f(t) = \det\left([T]_\beta - \lambda I_n\right)$ where $n = \dim V$.*

Remark 13 *Note that $[T]_\beta - \lambda I_n = [T - \lambda I_V]_\beta$ so we could have defined it thus.*

4

It should be fairly clear from the expansion in cofactors formula that the characteristic polynomial is, in fact, a polynomial. We can be a bit more precise.

Theorem 14 *Let $A \in F^{n \times n}$. The characteristic polynomial of A is a polynomial of degree n with leading coefficient $(-1)^n$. Also, A has at most n distinct eigenvalues.*

Proof. Exercise. ∎

2 Problems

- Be sure to read more examples in the book on diagonalizing matrices by finding zeroes of the characteristic polynomial and then finding eigenvectors.

- FIS Section 5.1 exercises 2, 3, 7, 8ab, 9 12, 14, 15, 16, 19-22

3 Diagonalizability

Theorem 15 *Let T be a linear operator on a vector space V and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of T. If $v_1, \ldots, v_k$ are eigenvectors of T such that $T(v_i) = \lambda_i v_i$ for $i = 1, \ldots, k$, then $\{v_1, \ldots, v_k\}$ is linearly independent.*

Proof. Consider a linear combination

$$\sum_{i=1}^{k} a_i v_i = \vec{0}.$$

It follows by applying $T - \lambda_k I_V$ that

$$\sum_{i=1}^{k-1} a_i (\lambda_i - \lambda_k) v_i = \vec{0}.$$

We can now use this to induct on k. If $k = 1$, then since $v_1 \neq \vec{0}$ (since eigenvectors are nonzero), $\{v_1\}$ is linearly independent. Now suppose any set of eigenvectors of distinct eigenvalues that has fewer than k elements is linearly independent. We see that if

$$\sum_{i=1}^{k} a_i v_i = \vec{0}$$

then we can apply $T - \lambda_k I_V$ to both sides to find that

$$\sum_{i=1}^{k-1} a_i (\lambda_i - \lambda_k) v_i = \vec{0}.$$

Since $\{v_1, \ldots, v_{k-1}\}$ are linearly independent and $\lambda_i \neq \lambda_k$, we must have that $a_1 = \cdots = a_{k-1} = 0$. It then follows that $a_k = 0$ as well. ∎

Corollary 16 *Let V be an n-dimensional vector space and $T \in \mathcal{L}(V)$. If T has n distinct eigenvalues, then T is diagonalizable.*

Proof. T has at least one eigenvector for each eigenvalue, yielding a set $\{v_1, \ldots, v_n\}$. By the theorem, this set is linearly independent, and thus a basis. By a previous theorem, if V has a basis of eigenvectors of T, then T is diagonalizable. ∎

Note that the converse of the theorem is false; we can have a linear operator that is diagonalizable that has repeated eigenvalues (for instance, the identity map).

Definition 17 *A polynomial $f(t)$ in $P(F)$ splits over F if there are scalars $c, a_1, \ldots, a_n$ (not necessarily distinct) in F such that*

$$f(t) = c(t - a_1)(t - a_2) \cdots (t - a_n).$$

So basically a polynomial splits if it can be factored with all linear factors. Note that the field is important. For instance, $t^2 + 1$ does not split over $\mathbb{R}$ or $\mathbb{Q}$ but it does split over $\mathbb{C}$ since $t^2 + 1 = (t + i)(t - i)$. Also note that we did not preclude $(t - 1)^2$, which also splits.

Theorem 18 *The characteristic polynomial of any diagonalizable linear operator splits.*

Proof. It is clear that if D is a diagonal matrix, then its characteristic polynomial splits. If T is diagonalizable, then there is a basis β such that $[T]_\beta$ is diagonal, so the theorem follows. ∎

If we think a bit harder about the proof, we see that it may be possible to have an eigenvalue repeated when written in diagonal form, and this leads to a repeated zero in the characteristic polynomial.

Definition 19 *Let λ be an eigenvalue of a linear operator or matrix with characteristic polynomial $f(t)$. The (algebraic) multiplicity of λ is the largest positive integer k such that $(t - \lambda)^k$ is a factor of $f(t)$.*

While we have shown that if a transformation is diagonalizable and a diagonal representation has a repeated eigenvalue, then their must be a repeated zero in the characteristic polynomial. It is natural to ask if a transformation whose characteristic polynomial splits with a repeated eigenvalue must be diagonalizable, but this is false.

Example 20 *Consider the matrix*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Its characteristic polynomial is λ^2, so it splits. We see that $N(A) = \operatorname{span}\{(1,0)\}$, which is one-dimensional. It is therefore impossible to have a basis of eigenvectors and so A is not diagonalizable.

The problem in the last example is that we are unable to find a basis of eigenvectors since there is only one eigenvalue and the nullspace of $A - \lambda I_V$ (where $\lambda = 0$) is only one-dimensional. This leads to the following definition.

Definition 21 *Let T be a linear operator on a vector space V and let λ be an eigenvalue of T. Define $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$. The set E_λ is called the* eigenspace *of T corresponding to the eigenvalue λ. Analogously, the eigenspace of a matrix A is the eigenspace of L_A.*

Theorem 22 *Let T be a linear operator on a finite-dimensional vector space V and let λ be an eigenvalue of T with multiplicity m. Then $1 \leq \dim E_\lambda \leq m$.*

Proof. Choose an ordered basis $\{v_1, \ldots, v_p\}$ for E_λ, then extend to a basis β for V. Then $[T]_\beta$ will have the block form

$$\begin{pmatrix} \lambda I_p & A \\ 0 & B \end{pmatrix}.$$

It is clear that the characteristic polynomial will look like $f(t) = (\lambda - t)^p \det(B - tI)$. It follows that the multiplicity m is at least as big as $p = \dim E_\lambda$. Since λ is an eigenvalue, it has at least one eigenvector so $\dim E_\lambda \geq 1$. ∎

Theorem 23 *Let T be a linear operator on a vector space V and let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T. For each $i = 1, 2, \ldots, k$ let S_i be a finite linearly independent subset of E_{λ_i}. Then $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent subset of V. [Note: the union should really be more of a concatenation since it needs to preserve the ordering of the S_i.]*

Proof. We need some notation to keep track of the vectors in S_i and S. Suppose $S_i = \{v_{i1}, \ldots, v_{in_i}\}$ for each i and that $S = \{v_{11}, \ldots, v_{1n_1}, v_{21}, \ldots, v_{kn_k}\}$. Consider scalars a_{ij} such that

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} v_{ij} = 0.$$

For each i, let $w_i = \sum_{j=1}^{n_i} a_{ij} v_{ij}$. Note that $w_i \in S_i$ and $\sum_{i=1}^{k} w_i = 0$. If we could show that this implies $w_i = 0$ for all $i = 1, \ldots, k$, then it will follow that $a_{ij} = 0$ for all i, j (since $\{v_{i1}, \ldots, v_{in_i}\}$ is a basis for each i). This is the content of the following lemma. ∎

Lemma 24 *Let T be a linear operator and let $\lambda_1, \ldots, \lambda_p$ be distinct eigenvalues of T. For each $i = 1, \ldots, p$, let $v_i \in E_{\lambda_i}$. If*

$$v_1 + \cdots + v_p = \vec{0}$$

then $v_i = 0$ for all i.

Proof. By Theorem 15 , $N = \left\{ v_i : v_i \neq \vec{0} \right\}$ is a linearly independent set. Thus, if N is nonempty, then $v_1 + \cdots + v_p \neq \vec{0}$. ∎

Theorem 25 *Let T be a linear operator on a finite-dimensional vector space V and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. then*

1. *T is diagonalizable if and only if the characteristic polynomial of T splits and the multiplicity of λ_i is equal to $\dim E_{\lambda_i}$ for all i.*

2. *If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T.*

Proof. If T is diagonalizable, then there is a basis β such that $[T]_\beta$ takes the block form

$$
[T]_\beta = \begin{pmatrix} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k I_{n_k} \end{pmatrix}.
$$

(Technically, there is a basis such that the matrix is diagonal, but then we can reorder that basis so that equal eigenvalues are adjacent.) It follows that the characteristic polynomial of T is $(\lambda_1 - t)^{n_1} \cdots (\lambda_k - t)^{n_k}$ where $n_1, \ldots, n_k$ are the dimensions of $E_{\lambda_1}, \ldots, E_{\lambda_k}$.

Now suppose the characteristic polynomial of T splits is equal to $(\lambda_1 - t)^{n_1} \cdots (\lambda_k - t)^{n_k}$ where $n_1, \ldots, n_k$ are the dimensions of $E_{\lambda_1}, \ldots, E_{\lambda_k}$. Then by Theorem 23 we can take bases β_i of E_{λ_i} and the set $\beta = \beta_1 \cup \cdots \cup \beta_k$ is linearly independent. Since it has size equal to $n_1 + \cdots + n_k = \dim V$ (since this is the degree of the characteristic polynomial), β must be a basis and in that basis, $[T]_\beta$ has the form above and is thus diagonalizable. Note that this proves the second statement as well. ∎

4   Problems

- Be sure to read the examples in the book on diagonalizing.

- FIS Section 5.2 exercises 2, 3, 6, 7 (diagonalize first: see Example 7), 8, 12, 13, 18, 19.