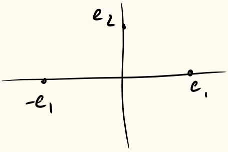
Remark: Since a polynomial (over any field) of degree n has at most n roots, Thm 4.2.6 implies that an operator $L: V \longrightarrow V$ has at most $n := \dim V$ eigenvalues. However, an operator could have no eigenvalues at all.

4.2.7 Example

Let $L: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be rotation by $\pi/2$ degrees. W/ $E = (e_1, e_2)$,



$$[L]_E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

So

$$cp_L(x) = \det \begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix}$$
$$= x^2 + 1$$

which has no real roots, hence no eigenvalues. Geometrically, it makes sense - rotation does not expand or contract in any direction.

□

4.2.8 Example

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

$$cp_A(x) = \det \begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix}$$

$$= (x-3) \begin{vmatrix} x-2 & 1 \\ -2 & x \end{vmatrix} + \begin{vmatrix} -2 & 1 \\ -2 & x \end{vmatrix} + \begin{vmatrix} -2 & x-2 \\ -2 & -2 \end{vmatrix}$$

$$= (x-3)(x^2 - 2x + 2) + (-2x + 2) + (4 + 2x - 4)$$

$$= \dots$$

$$= x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

So A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$.

Eigenspace $E_1 = \ker(I-A)$

$$I-A = \begin{pmatrix} -2 & -1 & 1 \\ -2 & -1 & 1 \\ -2 & -2 & 1 \end{pmatrix}$$

(row reduce)

$$\sim \begin{pmatrix} -2 & -1 & 1 \\ -2 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} -2 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So $\mathrm{rk}(I-A) = 2 \Rightarrow \dim E_1 = 1$.

Specifically, $E_1 = \mathrm{span}\{(1,0,2)\}$

Eigenspace $E_2 = \ker(2I-A)$

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

$$\sim \begin{pmatrix} -1 & -1 & 1 \\ -2 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} -1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

So $\mathrm{rk}(2I-A) = 2 \Rightarrow \dim E_2 = 1$.

Specificall, $E_2 = \mathrm{span}\{(1,1,2)\}$.

□

4.2.9 Example

$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{Z}_2^{3\times 3}$

$$c_{pA}(\lambda) = \det\begin{pmatrix} x+1 & 0 & 1 \\ 0 & x+1 & 1 \\ 1 & 1 & x+1 \end{pmatrix} = (x+1)\begin{vmatrix} x+1 & 1 \\ 1 & x+1 \end{vmatrix} + \begin{vmatrix} 0 & x+1 \\ 1 & 1 \end{vmatrix}$$

$$= (x+1)\left((x+1)^2 + 1\right) + (x+1)$$

$$= (\lambda+1)^3$$

So A has one eigenvalue $\lambda_1 = 1$. To find E_1,

$$I-A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

So $\mathrm{rank}(I-A) = 2 \Longrightarrow \dim E_1 = 1$. Specifically,

$E_1 = \mathrm{span}\{(1,1,0)\}$.

□

4.2.10 Lemma If $v_1, \dots, v_k$ are eigenvectors for L associated to pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_k$, then $v_1, \dots, v_n$ are independent.

Proof. Induction on k. The case $k=1$ is clear. Suppose $k \geq 1$. Assume $\sum_{i=1}^{k+1} \alpha_i v_i = 0$. Apply $L - \lambda_{k+1} id$ to the eq:

$$
\begin{aligned}
0 = (L - \lambda_{k+1} id)(0) &= (L - \lambda_{k+1} id)\left(\sum_{i=1}^{k+1} \alpha_i v_i\right) \\
&= \sum_{i=1}^{k+1} \alpha_i (L v_i - \lambda_{k+1} v_i) \\
&= \sum_{i=1}^{k} \alpha_i (L v_i - \lambda_{k+1} v_i) + L v_{k+1} - \lambda_{k+1} v_{k+1} \\
&= \sum_{i=1}^{k} \alpha_i (\lambda_i v_i - \lambda_{k+1} v_i) + 0 \\
&= \sum_{i=1}^{k} \alpha_i (\lambda_i - \lambda_{k+1}) v_i .
\end{aligned}
$$

By inductive hypothesis, $v_1, \dots, v_k$ are indep, so

$$\alpha_i (\lambda_i - \lambda_{k+1}) = 0$$

for all $i = 1, \dots, k$. Note that $(\lambda_i - \lambda_{k+1}) \neq 0$ since $\lambda_1, \dots, \lambda_{k+1}$ are pairwise distinct. Thus, $\alpha_i = 0$ for $i = 1, \dots, k$. In particular, $\lambda_{k+1} v_{k+1} = 0$ which implies $\lambda_{k+1} = 0$ since $v_{k+1} \neq 0$ by def of eigenvector. $\square$

4.2.11 Corollary Let $\lambda_1, \dots, \lambda_k$ be pairwise distinct eigenvalues of L. Then

$$E_{\lambda_1} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Thus,

$$\dim\left(\sum_{i=1}^{k} E_{\lambda_i}\right) = \sum_{i=1}^{k} \dim(E_{\lambda_i})$$

and if $B_1, \ldots, B_k$ are ordered bases for $E_{\lambda_1}, \ldots, E_{\lambda_k}$, then $B := (B_1, \ldots, B_k)$ is an ordered basis for $\sum_{i=1} E_{\lambda_i}$.

Proof. Suffices to show that

$$E_{\lambda_i} \cap \left(\sum_{j \neq i} E_{\lambda_j}\right) = \{0\} \quad \forall i = 1, \ldots, k$$

Everything else follows from previous results. If

$v_i \in E_{\lambda_i}$ and $v_i \in \sum_{j \neq i} E_{\lambda_j}$, then

$$v_i = \sum_{j \neq i} v_j$$

for some $v_j \in E_{\lambda_j}$, $j \neq i$. This contradicts independence of eigenvectors from distinct eigenspaces in 4.2.10. $\square$

4.2.12 Definition A linear operator $L: V \to V$ is diagonalizable if there exists an ordered basis for V such that $[L]_B$ is a diagonal matrix. $\square$

Remark: Suppose $L: V \to V$ is diagonalizable. Then there is a basis B such that

$$[L]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & 0 & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & 0 & & & & & \ddots & \\ & & & & & & \lambda_n & \\ & & & & & & & \ddots \\ & & & & & & & & \lambda_n \end{pmatrix}$$

where $\lambda_1, \dots, \lambda_k$ are pairwise distinct and λ_i appears w/ multiplicity d_i. That is,

$$[L]_{\mathcal{B}} = \lambda_1 I_{d_1} \oplus \lambda_2 I_{d_2} \oplus \cdots \oplus \lambda_n I_{d_n}$$

$$= \begin{pmatrix} \lambda_1 I_{d_1} & & & \\ & \lambda_2 I_{d_2} & & \\ & & \ddots & \\ & & & \lambda_n I_{d_n} \end{pmatrix}.$$

Then $cp_L(x) = \prod_{i=1}^{n} (x - x_i)^{d_i}$. Thus, if L is diagonalizable, the characteristic polynomial splits into a product of linear polynomials.

Note: Over an algebraically closed field like $\mathbb{C}$, every polynomial splits into a product of linear factors. But for a general field, this is saying something special about L. (e.g. $x^2+1 = (x-i)(x+i)$ over $\mathbb{C}$, but over $\mathbb{R}$ x^2+1 has no roots).

Further, the diagonal entries are the eigenvalues of L and $\dim E_{\lambda_i} = d_i$:

$$E_{\lambda_i} = \ker\left(\lambda_i I - [L]_B\right)$$

$$= \ker\begin{pmatrix} (\lambda_i - \lambda_1) I_{d_1} & & \\ & \ddots & \\ & & (\lambda_i - \lambda_n) I_{d_n} \end{pmatrix}$$

and this matrix has precisely d_i zeros on the diagonal. This motivates the following definition.

4.2.13 Definition Let $L: V \to V$ be a linear operator on a finite-dimensional vector space V and let λ be an eigenvalue of L. Then the **geometric multiplicity** of λ is $\dim E_\lambda$ and the **algebraic multiplicity** of λ is the multiplicity of λ as a root of $cp_L(x)$. □

The above remark showed that if L is diagonalizable, then for each eigenvalue λ, the geometric and algebraic multiplicities of λ agree.

4.2.14 Proposition The geometric multiplicity of λ is $\leq$ the algebraic multiplicity of λ.

Proof. Write $k := \dim E_\lambda$. There is a basis $v_1, \dots, v_k$ for E_λ. Extend to a basis $v_1, \dots, v_k, v_{k+1}, \dots, v_n$ for V. Then for this basis (call it B),

$$[L]_B = \left(\begin{array}{c|c} \lambda I_k & B \\ \hline 0 & C \end{array}\right)$$

for some matrices B and C. Then

$$cp_L(x) = \det\left(\begin{array}{c|c} (x-\lambda)I_k & -B \\ \hline & xI_{n-k}-C \end{array}\right)$$

$$= \det((x-\lambda)I_k)\det(xI_{n-k}-C)$$

$$= (x-\lambda)^k\, cp_C(x).$$

Thus, k is $\leq$ the algebraic multiplicity of λ. $\square$

4.2.15 Theorem Let L be a linear operator on a finite-dimensional space V. The following are equivalent:

(a) L is diagonalizable

(b) There is a basis for V consisting of eigenvectors of L.

(c) The eigenvectors of L span V

(d) The characteristic polynomial splits into a product of linear polynomials and the geometric multiplicity of each eigenvalue coincides with its algebraic multiplicity

(i.e, $cp_L(x) = \prod_{i=1}^{n}(x-\lambda_i)^{d_i}$ where $\lambda_1, \dots, \lambda_n$ are pairwise distinct eigenvalues of L and $d_i = \dim E_{\lambda_i}$)

Proof. The remark above showed (a) $\Rightarrow$ (d).

(d) $\Rightarrow$ (c) If $cp_L(x)$ splits as above, then

$$\dim V = \deg cp_L(x) = \sum_{i=1}^{n} d_i$$

$$= \sum_{i=1}^{n} \dim E_{\lambda_i}.$$

Thus, $V = \sum_{i=1}^{n} E_{\lambda_i}$.

(c) ⇒ (b) $V = \sum_{i=1}^{n} E_{\lambda_i} = \bigoplus_{i=1}^{n} E_{\lambda_i}$.

(b) ⇒ (a) Obvious. $\square$

Remark The operators in Examples 4.2.7, 4.2.8, and 4.2.9 were not diagonalizable.

Method for diagonalizing a matrix Let A be a square matrix.

(1) Compute $cp_A(x)$ and determine its roots. If it does not factor as a product of linear polynomials in $F[x]$, then A is not diagonalizable.

(2) If $cp_A(x) = \prod_{i=1}^{n} (x-\lambda_i)^{d_i}$, then determine the dimension of each eigenspace $E_{\lambda_1}, \ldots, E_{\lambda_n}$. For each i, compute the kernel of $\lambda_i I - A$. If $\dim E_{\lambda_i} < d_i$ for some i, A is not diagonalizable.

(3) If A is diagonalizable, then we just need to find an invertible matrix P such that $P^{-1}AP$ is diagonal. Compute a basis B_i for each E_{λ_i}. Then $B = (B_1, \ldots, B_n)$ is a basis for V consisting of eigenvectors. Set $P = T^{E}_{B}$ where E is the standard basis for F^n. Then

[$A\cdot$ denotes left mult by A]

$$P^{-1}AP = T^{B}_{E} [A\cdot]^{E}_{E} T^{E}_{B}$$
$$= [A\cdot]^{B}_{B}$$

$$= \begin{pmatrix} \lambda_1 \cdots \lambda_1 & & & \\ & \lambda_2 \cdots \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \cdots \lambda_n \end{pmatrix}$$

☐

Remark: An operator L is diagonalizable iff V decomposes as the direct sum of the eigenspaces for L. Write $V = \bigoplus_{i=1}^{n} E_{\lambda_i}$. For each i, notice that $L(E_{\lambda_i}) \subseteq E_{\lambda_i}$. Thus, L restricts to a linear operator

$$L|_{E_{\lambda_i}} : E_{\lambda_i} \longrightarrow E_{\lambda_i}.$$

In fact, $L|_{E_{\lambda_i}} = \lambda Id_{E_{\lambda_i}}$ and thus $L = \bigoplus_{i=1}^{n} \lambda_i Id_{E_{\lambda_i}}$.

Remark: We have seen that not every operator L is diagonalizable, i.e., V cannot always be decomposed as a direct sum of eigenspaces for L. Ideally, we would like to find a simple matrix representation for every operator. The preceding remark suggests that we look for direct sum decompositions which are invariant under L.

4.2.16 Definition

Let $L: V \longrightarrow V$ be a linear operator. A subspace W of V is L-invariant if $L(W) \subseteq W$.

☐

Remark: • If W is L-invariant, the L restricts to an operator on W:

$$L|_W : W \longrightarrow W.$$

- Every subspace W of V has a complement W' so that $V = W \oplus W'$. However, there is no reason that any complement of W will be L-invariant. Choose bases B, B' for W and W'. Then $B \cup B'$ is a basis for V. Say $\dim W = k$, $\dim V = n$. Then the matrix for L has the form

$$[L]_{B\cup B'} = \begin{array}{cc} \begin{array}{cc} \overbrace{\quad\;}^{k} & \overbrace{\quad\;}^{n-k} \end{array} & \\ \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] & \begin{array}{l} \}\,k \\ \}\,n-k \end{array} \end{array}$$

where $A \in F^{k\times k}$, $B \in F^{k\times(n-k)}$, $C \in F^{(n-k)\times k}$, $D \in F^{(n-k)\times(n-k)}$. Then W is L-invariant iff $C = 0$ and W' is L-invariant iff $B = 0$. So a direct sum decomp of V into L-invariant subspaces yields a matrix in block diagonal form.

Note: the above argument turned up in Prop 4.2.14. Since E_λ was L-invariant, the matrix had the form

$$[L] = \left(\begin{array}{c|c} \lambda I & B \\ \hline 0 & D \end{array}\right).$$

However, it is possible that $B \neq 0$ since E_λ need not have an L-invariant complement.

Remark: The remarks provide a strategy for understanding an operator L. Find a decomposition

$$V = V_1 \oplus \cdots \oplus V_n$$

into L-invariant subspaces. Then study the individual operators.

$$L|_{V_i} : V_i \longrightarrow V_i .$$

One can then understand L using the fact that $L = \hat{\bigoplus_{i=1}} L|_{V_i}$. For a diagonalizable operator, we can take V_i to be the eigenspaces.

4.3 The Minimal Polynomial

4.3.1 Theorem Let $L: V \to V$ be a linear operator on a finite-dimensional vectorspace V. There exists a monic polynomial $m_L(x) \in F[x]$ that is unique with the following property: $m_L(L) = 0$ and if $p(x) \in F[x]$ also satisfies $p(L) = 0$, then $m_L(x)$ divides $p(x)$. The polynomial $m_L(x)$ is called the **minimal polynomial** of L.

Proof. The most elegant proof uses ring theory. You can skip this if you haven't taken 111B. The ring $F[x]$ is a principal ideal domain - every ideal is generated by a single element. The set $I = \{p(x) \in F[x] \mid p(L) = 0\} \subseteq F[x]$ is an ideal in $F[x]$. It suffices to show that $I \neq 0$, then there is a unique monic polynomial which generates I, and this is equivalent to the theorem statement. Consider the operators $1, L, L^2, \dots, L^{n^2-1}, L^{n^2}$ where $n = \dim V$. Since there are n^2+1 operators, they must be dependent.

There exists $a_0, \dots, a_{n^2} \in F$, not all zero, such that

$$\sum_{i=0}^{n^2} a_i L^i = 0.$$

Thus, $p(x) := \sum_{i=1}^{n^2} a_i x^i$ is a nonzero element of I.

☐

4.3.2 Theorem Let $L: V \to V$ be a linear operator on a finite-dimensional vector space V. Then $m_L(x)$ and $cp_L(x)$ have the same roots (with possibly different multiplicities).

Proof. Suppose $m(\lambda)=0$. Then $m(x) = (x-\lambda)p(x)$ for some $p(x) \in F[x]$. Since $\deg m(x) > \deg p(x)$, $p(L) \neq 0$ by minimality of m. So, there exists $v \in V$ such that $w := p(L)(v) \neq 0$. Then,

composition of functions!

$$0 = m(L) = (L - \lambda \mathrm{Id}_V) p(L),$$

Evaluating at v yields

$$0 = m(L)(v) = (L - \lambda \mathrm{Id}_V) w = Lw - \lambda w.$$

But $w \neq 0$, thus w is an eigenvector and λ is an eigenvalue, that is, a root of $cp_L(x)$.

Conversely, suppose λ is an eigenvalue of L. Then $Lv = \lambda v$ for some $v \neq 0$.

Exercise: If $Lv = \lambda v$, then $p(L)(v) = p(\lambda) v$ for any polynomial $p(x) \in F[x]$

Using the exercise together w/ $m(L) = 0$, we get

$$0 = m(L)(v) = m(\lambda)v.$$

But $v \neq 0$, so $m(\lambda) = 0$. $\square$

4.3.3 Theorem (Cayley-Hamilton) Let L be a linear operator on a finite-dimensional vector space. Then $cp_L(L) = 0$.

Proof. Homework. $\square$

4.3.4 Corollary The minimal polynomial $m_L(x)$ divides the characteristic polynomial $cp_L(x)$. $\square$

4.3.5 Corollary The minimal polynomial has degree at most $n = \dim(V)$. $\square$

4.3.6 Example Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 4 \\ 0 & 1 & 0 \end{pmatrix} \in (\mathbb{Z}_5)^{3\times 3}$.

We compute $m_A(x)$. First, find $cp_A(x)$:

$$\begin{aligned} cp_A(x) &= \det\begin{pmatrix} x+4 & 3 & 0 \\ 0 & x+3 & 1 \\ 0 & 4 & x \end{pmatrix} \\ &= (x+4)\big((x+3)x - 4\big) \\ &= (x+4)(x^2+3x+1) \\ &= (x+4)^3 \end{aligned}$$

$p(x) := x^2 + 3x + 1$

$p(1) = 0$ so $x+4$ divides $p(x)$

$\Rightarrow p(x) = (x+4)(x+4)$

Since $m_A(x)$ divides $cp_A(x)$, there are three possibilities: $m_A(x) \in \{x+4, (x+4)^2, (x+4)^3\}$. It suffices to check whether $(A+4I)^2 = 0$. We have

$$(A+4I)^2 = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 4 \end{pmatrix}^2$$

$$= \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, $m_A(x) = (x+4)^3$. Notice that in fact $(A+4I)^3 = 0$, as claimed by Cayley-Hamilton. To determine if A is diagonalizable, we compute E_λ for the eigenvalue $\lambda = 1$.

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

So $\operatorname{rank}(I-A) = 2$ and $\dim E_\lambda = 1$. Thus, A is not diagonalizable. Later, we will see that A is not diagonalizable because m_A has a repeated root. $\square$