4.3.7 Proposition Suppose L is diagonalizable w/ pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_n$. Then

$$m_L(x) = \prod_{i=1}^{n} (x-\lambda_i).$$

I.e., the minimal polynomial has no repeated roots.

Proof. By Thm 4.2.15, write

$$cp_L(x) = \prod_{i=1}^{n} (x-\lambda_i)^{d_i}$$

for $d_i = \dim E_{\lambda_i}$. By Thm 4.3.2, $m_L(x)$ has the same roots as $cp_L(x)$. So $p(x) := \prod_{i=1}^{n} (x-\lambda_i)$ is the candidate for $m_L(x)$ of smallest degree. Let $v \in V$. Then $v = v_1 + \cdots + v_n$ where $v_i \in E_{\lambda_i}$. Hence,

$$(p(L))(v) = \sum_{i=1}^{n} (p(L))(v_i)$$

$$= 0$$

Since

$$p(L)(v_i) = \left(\prod_{j \neq i} (L-\lambda_j) \right) (L-\lambda_i)(v_i)$$

$$= \left(\prod_{j \neq i} (L-\lambda_j) \right) \underbrace{(Lv_i - \lambda_i v_i)}_{=0}$$

$$= 0$$

Thus, $p(x) = m_L(x)$. □

Remark: Let $L_1 : V_1 \longrightarrow V_1$ and $L_2 : V_2 \longrightarrow V_2$ be linear maps. Using the Universal Property of $V_1 \oplus V_2$ (see Weekly 2.3), there is a unique linear map

$$L_1 \oplus L_2 : V_1 \oplus V_2 \longrightarrow V_1 \oplus V_2$$

s.t. $(L_1 \oplus L_2)(v_1, v_2) = (L_1(v_1), L_2(v_2))$. With respect to appropriate bases

$$[L_1 \oplus L_2] = \begin{pmatrix} [L_1] & 0 \\ 0 & [L_2] \end{pmatrix} = [L_1] \oplus [L_2].$$

So,

$$\begin{aligned} cp_{L_1 \oplus L_2}(x) &= \det\left(xI - \begin{pmatrix} [L_1] & \\ & [L_2] \end{pmatrix} \right) \\ &= \det \begin{pmatrix} xI - [L_1] & \\ & xI - [L_2] \end{pmatrix} \\ &= \det(xI - [L_1]) \det(xI - [L_2]) \\ &= cp_{L_1}(x)\, cp_{L_2}(x). \end{aligned}$$

How about $m_{L_1 \oplus L_2}(x)$? Let $p(x) \in F[x]$. Then

$$p([L_1 \oplus L_2]) = p\begin{pmatrix} [L_1] & 0 \\ 0 & [L_2] \end{pmatrix} = \begin{pmatrix} p([L_1]) & \\ & p([L_2]) \end{pmatrix}.$$

So $p(L_1 \oplus L_2) = 0$ if and only if $p(L_1) = 0$ and $p(L_2) = 0$.

Thus,

$$\begin{aligned} m_{L_1 \oplus L_2}(x) \mid p(x) &\iff p(L_1 \oplus L_2) = 0 \\ &\iff p(L_1) = 0 \text{ and } p(L_2) = 0 \\ &\iff m_{L_1}(x) \mid p(x) \text{ and } m_{L_2}(x) \mid p(x). \end{aligned}$$

So $m_{L_1 \oplus L_2}(x)$ is the least common multiple (lcm) of $m_L(x)$ and $m_{L_2}(x)$.

This provides another proof of 4.3.7.

Consider $\lambda I_n = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix}$

$$cp_{\lambda I_n}(x) = (x-\lambda)^n$$

$$m_{\lambda I_n}(x) = (x-\lambda).$$

Consider a diagonal matrix $A = \begin{pmatrix} \lambda_1 Id_1 & & \\ & \ddots & \\ & & \lambda_n Id_n \end{pmatrix}$

Then $cp_A(x) = \prod_{i=1}^{n} (x-\lambda_i)^{d_i}$. Thus,

$$\begin{aligned} m(x) &= lcm(m_{\lambda_1 Id_1}, \dots, m_{\lambda_n Id_n}) \\ &= lcm(x-\lambda_1, \dots, x-\lambda_n) \\ &= \prod_{i=1}^{n} (x-\lambda_i). \end{aligned}$$

□

The next goal is to prove the converse to 4.3.7.

4.3.8 Lemma Suppose $P, Q : V \to V$ are linear operators such that $P+Q = Id_V$ and $PQ = 0$. Then

(a) $QP = 0$

(b) $\text{im } P = \ker Q$

(c) $\ker P = \text{im}(P)$

(d) $V = \text{im}(P) \oplus \ker(P)$

□

Proof. Exercise.

4.3.9 Lemma Let $L : V \to V$ be a linear operator on a finite-dimensional vector space. Let $f(x), g(x) \in F[x]$ be coprime polynomials such that $f(L)g(L) = 0$. Then

$$V = \ker(f(L)) \oplus \ker(g(L))$$

and $\ker(f(L))$, $\ker(g(L))$ are L-invariant.

Proof. Since the polynomials are coprime, there exists polynomials $p(x)$ and $q(x)$ such that

$$1 = p(x)f(x) + q(x)g(x) \qquad \left(\begin{array}{c}\text{Bezout's lemma} \\ \text{in } F[x]\end{array}\right)$$

Define $P := q(L)g(L)$ and $Q = p(L)f(L)$. The above implies $Id_V = P + Q$. Moreover, $PQ = 0$ since $f(L)g(L) = 0$. By the lemma, it follows that

$$V = \operatorname{im}(P) \oplus \ker(P).$$

We claim $\operatorname{im}(P) = \ker(f(L))$. Indeed,

$$\begin{aligned} \operatorname{im}(P) = \operatorname{im}(g(L)q(L)) &\subseteq \ker(f(L)) \\ &\subseteq \ker(f(L)p(L)) \\ &= \ker(Q) \\ &= \operatorname{im}(P). \end{aligned}$$

Similarly, one can show that $\ker P = \operatorname{im}(Q) = \ker(g(L))$. So we have the decomposition of V as claimed. The subspaces $\ker(f(L))$ and $\ker(g(L))$ are L-invariant. For example, if $v \in \ker f(L)$. then

$$f(L)(L(v)) = (f(L) \circ L)(v)$$

$$= (L \circ f(L))(v) = L(0) = 0.$$

So $L(v) \in \ker f(L)$. ☐

4.3.10 Theorem Let $L: V \to V$ be a linear operator on a finite-dimensional vector space. Then L is diagonalizable if and only if $m_L(x)$ has no repeated roots.

Proof. ($\Rightarrow$) Prop 4.3.7.

($\Leftarrow$) Proof by induct on the number of pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_n$. If $n=1$, then $m_L(x) = x - \lambda_1$. Thus, $0 = m_L(L) = L - \lambda_1 I$. Thus, $V = \ker(0) = \ker(L - \lambda_1 I) = E_{\lambda_1}$. Let $n \geq 2$ and assume that any operater with $n-1$ distinct eigenvalues is diagonalizable. Let L be an operator w/ pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_n$. By hypothesis,

$$m_L(x) = \prod_{i=1}^{n} (x - \lambda_i).$$
$$= (x - \lambda_1) \prod_{i=2}^{n} (x - \lambda_i).$$

Set $f(x) = x - \lambda_1$ and $g(x) = \prod_{i=2}^{n} (x - \lambda_i)$. Then $m_L(x) = f(x) g(x)$. Since $m_L(L) = 0$, Lemma 4.3.9 implies that

$$V = \ker(f(L)) \oplus \ker(g(L)).$$
$$= E_{\lambda_1} \oplus \ker(g(L)).$$

Since both subspaces are L-invariant, $L = L_1 \oplus L_2$ where $L_1 = L|_{E_{\lambda_1}}$ and $L_2 = L|_{\ker(g(L))}$. Thus,

$$cp_L(x) = cp_{L_1}(x)\, cp_{L_2}(x).$$

The only eigenvalue of L_1 is λ_1, L_2 must have eigenvalues $\lambda_2, \dots, \lambda_n$. We need to show that λ_1 is not an eigenvalue of L_2. Suppose there exists $v \in \ker(g(L))$ such that $L_2(v) = \lambda_1 v$. Then

$$0 = g(L)(v) = g(\lambda_1)v.$$

But $g(\lambda_1) \neq 0$, so $v = 0$. Finally, by inductive hypothesis, L_2 is diagonalizable so that

$$\ker(g(L)) = E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_n}.$$

This completes the proof. □

4.3.11 Example

$$A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

Easy to show that $cp_A(x) = (x-1)(x-2)^2$. So $m_A(x) \in \{(x-1)(x-2), cp_A(x)\}$. In fact, $m_A(x) = (x-1)(x-2)$ Since

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

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

Thus, A is diagonalizable. □

4.4 Jordan Canonical Form

In this section, whenever we consider a linear operator $L: V \to V$ on a vector space V, we will assume that the field F contains the roots of $cp_L(x)$. In that case, the characteristic polynomial will always split into a product of linear factors

$$cp_L(x) = \prod_{i=1}^{n} (x - \lambda_i)^{d_i}.$$

Still, the operator may not be diagonalizable since the geometric multiplicities could be smaller than the algebraic multiplicities.

A field F is algebraically closed if every nonconstant polynomial $p(x) \in F[x]$ has a root in F. By induction and the Euclidean algorithm, all the roots of $p(x)$ are in F and $p(x)$ splits into a product of linear factors. So, for instance, the results of this section will apply to every operator $L: V \to V$ when V is defined over an algebraically closed field, like $\mathbb{C}$. However, the hypothesis may still be satisfied when F is not algebraically closed. For example, consider $F = \mathbb{Z}_2$

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in F^{2\times 2} \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in F^{2\times 2}$$

Then $cp_A(x) = (x+1)^2$ and $cp_B(x) = x^2 + x + 1$.

So the results would apply to A, but not to B, since $cp_B(0) = 1$ and $cp_B(1) = 1$, i.e., $cp_B(x)$ has no roots in F.

Idea: With the above hypothesis, the only obstruction to diagonalization is that the eigenspaces could be too small to give a direct sum decomposition of V. We will replace E_λ by a larger subspace $V_{(\lambda)}$ of "generalized eigenvectors". We will have

- $E_\lambda \subseteq V_{(\lambda)}$, $\dim V_{(\lambda)}$ = algebraic mult of λ
- $V = \bigoplus_{i=1}^{n} V_{(\lambda_i)}$
- $V_{(\lambda)}$ is L-invariant.

Thus, we will obtain a direct sume decomposition of L. The matrices representind $L|_{V_{(\lambda_i)}}$ will be fairly simple.

4.4.1 Definition Let $L: V \to V$ be a linear operator and λ an eigenvalue of L. A generalized eigenvector for λ is a nonzero vector $v \in V$ such that

$$(L - \lambda Id_V)^i v = 0$$

for some $i \geq 1$. The rank of a generalized eigenvector v is the smallest positive integer i such that $(L - \lambda I)^i v = 0$. The set of all generized eigenvectors for λ, together with zero vector, is denoted by $V_{(\lambda)}$. □

Observe: the rank 1 generalized eigenvectors are just the ordinary eigenvectors.

4.4.2 Lemma Let $L : V \to V$ be a linear operator and λ an eigenvalue. Then $V_{(\lambda)}$ is an L-invariant subspace of V.

Proof. By definition, $V_{(\lambda)} = \bigcup_{i \geq 1} \ker (L - \lambda \mathrm{Id}_V)^i$.

In general, the union of subspaces is not a subspace. But in this case, the union is over a nested chain of subspaces

$$\{0\} \subseteq E_\lambda \subseteq \ker (L - \lambda \mathrm{Id}_V) \subseteq \ker (L - \lambda \mathrm{Id}_V)^2 \subseteq \cdots$$

One can easily show in this situation that the union is a subspace. To show that $V_{(\lambda)}$ is L-invariant, let $v \in V_{(\lambda)}$. Then $(L - \lambda \mathrm{Id})^i v = 0$ for some $i \geq 1$. Then

$$(L - \lambda \mathrm{Id})^i (L(v)) \underset{\uparrow \text{Corollary 4.1.2}}{=} L((L - \lambda \mathrm{Id})^i v) = L(0) = 0.$$

Thus, $L(v) \in V_{(\lambda)}$. ☐

4.4.3 Lemma Let $L : V \to V$ be a linear operator and let λ be an eigenvalue of L. Then λ is the unique eigenvalue of the operator

$$L|_{V(\lambda)} : V(\lambda) \to V(\lambda).$$

Proof. Clearly, λ is an eigenvalue since $E_\lambda \subseteq V_{(\lambda)}$. Suppose $0 \neq v \in V_{(\lambda)}$ satisfies $Lv = \mu v$ for some $\mu \in F$. Then $(L - \lambda \mathrm{Id}_V)^i v = 0$ for some $i \geq 1$. Set $p(x) = (x-\lambda)^i$, so that $p(L)v = 0$. Then

$$0 = p(L)(v) = p(\mu)v = (\mu - \lambda)^i v.$$

Since $v \neq 0$, $(\mu-\lambda)^i = 0$. This forces $\mu = \lambda$. $\square$

4.4.4 Lemma Let $L : V \longrightarrow V$ be a linear operator and λ an eigenvalue. Then $V_{(\lambda)} = \ker(L - \lambda \mathrm{Id}_V)^d$ where d is the algebraic multiplicity of λ.

(i.e., the chain of subspaces terminates after d-steps

$$\{0\} \subseteq E_\lambda \subseteq \ker(L - \lambda \mathrm{Id}_V)^2 \subseteq \cdots \subseteq \ker(L - \lambda \mathrm{Id})^d = V_{(\lambda)}\text{)}$$

Proof. The characteristic polynomial of $L|_{V_{(\lambda)}}$ is given by

$$cp_{L|_{V(\lambda)}}(x) = (x - \lambda)^{\dim(V(\lambda))}$$

according to Lemma 4.4.3. By Cayley-Hamilton,

$$(cp_{L|_{V(\lambda)}}(x) - \lambda)^{\dim(V(\lambda))} = 0.$$

Let W be a complement of $V_{(\lambda)}$ so that $V = V_{(\lambda)} \oplus W$. Then with respect to appropriate bases,

$$[L] = \begin{pmatrix} [L|_{V(\lambda)}] & B \\ 0 & D \end{pmatrix}$$

for some matrices B and D. Thus,

$$cp_L(x) = cp_{L|_{V(\lambda)}}(x) \cdot cp_D(x).$$
$$= (x-\lambda)^{\dim(V(\lambda))} \; cp_D(x).$$

Thus, $\dim(V(\lambda)) \leq d$ where d is the algebraic mult of λ as an eigenvalue of L. Thus,

$$(L-\lambda Id_V)^d\Big|_{V(\lambda)} = (L|_{V(\lambda)} - \lambda Id_{V(\lambda)})^d = 0$$

This show that

$$V(\lambda) \subseteq \ker(L-\lambda Id_V)^d \subseteq V(\lambda).$$

$\square$

4.4.5 Theorem Let $L: V \to V$ be a linear operator on a finite-dimensional F-vector space V and assume F contains the roots of $cp_L(x)$. Then

$$V = V_{(\lambda_1)} \oplus \cdots \oplus V_{(\lambda_n)}$$

where $\lambda_1, \ldots, \lambda_n$ are the pairwise distinct eigenvalues of L. Further, $\dim V_{(\lambda_i)} =$ the algebraic multiplicity of λ_i.

Proof. Induction on the number of pairwise distinct eigenvalues. If $n=1$, then $cp_L(x) = (x-\lambda_1)^d$. By Cayley-Hamilton,

$$(L-\lambda_1 Id_V)^d = 0.$$

Thus, $V = \ker(L-\lambda_1 Id_V)^d = V(\lambda)$.

Lemma 4.4.4.

Let $n \geq 2$ and assume V decomposes as the direct sum of generalized eigenspaces whenever L is an operator w/ $n-1$ distinct eigenvalues. Assume L has pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_n$. Then

$$cp_L(x) = \prod_{i=1}^{n} (x-\lambda_i)^{d_i}.$$

Set $f(x) = (x-\lambda_1)^{d_1}$ and $g(x) = \prod_{i=2}^{n} (x-\lambda_i)^{d_i}$. Then f and g are coprime and $f(L)g(L) = 0$ by Cayley-Hamilton. By Lemma 4.3.9,

$$\begin{aligned} V &= \ker f(L) \oplus \ker g(L) \\ &= \ker (L - \lambda_1 Id_V)^{d_1} \oplus \ker g(L) \\ &\overset{\text{Lemma 4.4.4}}{=} V_{(\lambda_1)} \oplus \ker g(L). \end{aligned}$$

Set $L_1 := L|_{V(\lambda_1)}$ and $L_2 = L|_{\ker g(L)}$. Then $L = L_1 \oplus L_2$ and $cp_L(x) = cp_{L_1}(x)\, cp_{L_2}(x)$. Notice:

① λ_1 is not an eigenvalue of L_2. If $v \in \ker g(L)$, such that $Lv = \lambda_1 v$, then $0 = g(L)(v) = g(\lambda_1)v$. But $g(\lambda_1) \neq 0$ so $v = 0$.

② λ_1 is the unique eigenvector of L_1 by Lemma 4.4.3.

Together ① + ② implies that $cp_{L_1}(x) = (x-\lambda_1)^{d_1}$ and $cp_{L_2}(x) = (x-\lambda_2)^{d_2} \cdots (x-\lambda_n)^{d_n}$. Thus, $d_1 = \deg(cp_{L_1}(x)) = \dim V_{(\lambda_1)}$ and by inductive hypothesis $\ker g(L) = V_{(\lambda_2)} \oplus \cdots \oplus V_{(\lambda_n)}$ and $\dim V_{(\lambda_i)} = d_i$, $i = 2, \dots, n$. This completes the proof. □

Remark: By definition, $V_{(\lambda)} = \bigcup_{i \geq 1} \ker(L - \lambda Id_V)^i$.
Then we proved that actually
$$V_{(\lambda)} = \ker(L - \lambda Id)^d$$
where d = algebraic multiplicity of λ. We can do even better:

4.4.6 Proposition Let λ be an eigenvalue of L. The smallest positive integer m such that
$$V_{(\lambda)} = \ker(L - \lambda Id_V)^m$$
is the multiplicity of λ as a root of the minimal polynomial $m_L(x)$.

Proof. Consider $L|_{V_{(\lambda)}} : V_{(\lambda)} \longrightarrow V_{(\lambda)}$. Then

$$\begin{aligned} V_{(\lambda)} = \ker(L - \lambda Id_V)^i &\Leftrightarrow V_{(\lambda)} \subseteq \ker(L - \lambda Id_V)^i \\ &\Leftrightarrow (L - \lambda Id_V)^i (V_{(\lambda)}) = \{0\} \\ &\Leftrightarrow (L - \lambda Id_V)^i \big|_{V(\lambda)} = 0 \\ &\Leftrightarrow (L|_{V_{(\lambda)}} - \lambda Id_{V_{(\lambda)}})^i = 0 \\ &\Leftrightarrow \text{the minimal polynomial of } L|_{V(\lambda)} \text{ divides } (x-\lambda)^i. \end{aligned}$$

Since λ is the unique eigenvalue of $L|_{V_{(\lambda)}}$ (Lem 4.4.3) we have $m_{L|_{V_{(\lambda)}}}(x) = (x-\lambda)^k$ for some k. Thus, the smallest positive integer w/ the desired property is k, and k is the multiplicity of λ as a root of the minimal polynomial of $L|_{V_{(\lambda)}}$. By Theorem 4.4.5 the multiplicity of λ as a root of $m_{L|_{V_{(\lambda)}}}(x)$ coincides w/ the multiplicity of λ as a root of $m_L(x)$, hence the claim is proved. □

Recap: If $\lambda_1, \dots, \lambda_n$ are the distinct eigenvalues of L, then

① $V_{(\lambda_i)} \subseteq V$ is L-invariant

② $V = V_{(\lambda_1)} \oplus \cdots \oplus V_{(\lambda_n)}$

③ $\dim V_{(\lambda_i)}$ = algebraic multiplicity of λ_i.

④ $\{0\} \subseteq \ker(L-\lambda_i Id) \subseteq \ker(L-\lambda_i Id)^2 \subseteq \cdots \subseteq \ker(L-\lambda_i Id)^m = V_{(\lambda_i)}$

where m is the multiplicity of λ_i as a root of the minimal polynomial. □

Remark:

- The multiplicity of λ in $m_L(x)$ is 1 iff $V_{(\lambda)} = E_\lambda$ by ④.
- The geometric multiplicity of λ equals the algebraic multiplicity iff $V_{(\lambda)} = E_\lambda$ by ②.

Thus, our characterizations of diagonalizable operators are both equivalent to the statement that $V_{(\lambda)} = E_\lambda$ for all λ. In particular, the material in this section is just a refinement of what we saw in the diagonalizable case.