Recall:

(a) The set $\mathrm{End}(V)$ of linear operators $V \to V$ is an algebra

- its a vector space (over F), so we can add and scale linear operators
- one can "multiply" operators $L, K : V \to V$

$$LK := L \circ K \quad \text{(function composition)}$$

In particular, we can take "powers" of a linear operator $L : V \to V$

$$(i \geq 1) \quad L^i := \underbrace{L \circ \cdots \circ L}_{i\text{-times}} \qquad (i = 0) \quad L^0 := \mathrm{Id}_V$$

(b) If V is finite-dimensional, one has an isomorphism of (unital) algebras

$$\begin{tikzcd} \mathrm{End}(V) \arrow[r, "\sim"] & F^{\dim V \times \dim V} \\ L \arrow[r, mapsto] & {[L]_B} \end{tikzcd}$$

where B is a fixed ordered basis. Under this isomorphism, addition, scalar mult, and composition of linear maps corresponds to addition, scalar mult, and multiplication of matrices.

(c) The vector space $F[x]$ of polynomials is also an algebra: one can multiply polynomials in the usual way. Note: multiplication of poly's is commutative.

Observe: Consider a polynomial $p(x) = \sum_{i=0}^{n} a_i x^i \in F[x]$ and a linear operator $L : V \to V$. By (a), we can define a new linear operator

$$p(L) := \sum_{i=0}^{n} a_i L^i \in \mathrm{End}(V).$$

This defines a function

$$\begin{aligned} \Phi_L : F[x] &\longrightarrow \mathrm{End}(V) \\ p(x) &\longmapsto p(L) \end{aligned}$$

This is a (unital) algebra homomorphism.

4.4.1 Proposition Let $L \in \mathrm{End}_F(V)$. The function

$$\begin{aligned} \Phi_L : F[x] &\longrightarrow \mathrm{End}(V) \\ p(x) &\longmapsto p(L) \end{aligned}$$

is an F-algebra homomorphism.

Proof. A basis for $F[x]$ is $\{x^i \mid i \geq 0\}$. There is a unique linear map $\Phi_L : F[x] \to \mathrm{End}(V)$ such that $\Phi_L(x^i) = L^i$. Thus, if $p(x) = \sum_{i=0}^{n} a_i x^i$, then

$$\Phi_L\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=0}^{n} a_i \, \Phi_L(x^i) = \sum_{i=0}^{n} a_i L^i = p(L).$$

By def, $\Phi_L(1) = \mathrm{Id}_V$. By linearity of Φ_L, it suffices to check multiplicativity on the basis. We have

$$\Phi_L(x^i x^j) = \Phi_L(x^{i+j}) = L^{i+j} = L^i L^j = \Phi_L(x^i)\Phi_L(x^j).$$

Thus, Φ_L is an algebra homomorphism. $\square$

4.1.2 Corollary Let $L \in \mathrm{End}(V)$. Then for any $p(x), q(x) \in F[x]$, the operators $p(L)$ and $q(L)$ commute.

Proof. Follows from 4.1.1 since $F[x]$ is a commutative algebra:

$$\begin{aligned} p(L)q(L) &= \Phi_L(p(x))\,\Phi_L(q(x)) = \Phi_L(p(x)q(x)) \\ &= \Phi_L(q(x)p(x)) = \Phi_L(q(x))\,\Phi_L(p(x)) \\ &= q(L)\,p(L). \end{aligned}$$

$\square$

Remark: In general, operators do not commute

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \neq \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(as long as $\mathrm{ch}\, F \neq 2$). Later we will use the fact that $p(L)$ and $q(L)$ commute.

4.2 Eigenvalues

4.2.1 Definition Let V be a vector space and $L \in \mathrm{End}(V)$. An eigenvalue of L is a scalar $\lambda \in F$ such that

$$Lv = \lambda v$$

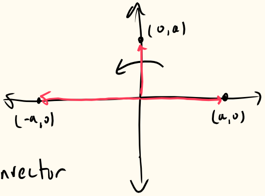
for some nonzero vector $v \in V$. Any such vector v is called an **eigenvector** of L associated to λ.

□

4.2.2 Example Let $L: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be reflection about y-axis. So $L((0,a)) = (0,a)$

Assume $a \neq 0$. $L((a,0)) = (-a,0)$

So $(0,a)$ is an eigenvector w/ eigenvalue 1 and $(a,0)$ is an eigenvector w/ eigenvalue -1



□

4.2.3 Proposition A nonzero vector $v \in V$ is an eigenvector for L associated to $\lambda \in F$ if and only if $v \in \ker(\underline{L - \lambda id})$.

Proof. Obvious.

□

4.2.4 Definition The subspace

$$E_\lambda := \ker(L - \lambda id)$$

of V is called the **eigenspace** of L associated to λ. The set E_λ consists of all the eigenvectors of L associated to λ, together with the zero vector.

□

4.2.5 Theorem Let $L \in \mathrm{End}(V)$ w/ V finite-dimensional. The following are equivalent.

(a) λ is an eigenvalue of L.

(b) $L - \lambda \mathrm{id}$ is not invertible

(c) $\det(L - \lambda \mathrm{id}) = 0$.

Proof. (b) and (c) are already equivalent by 3.5.13.

(a) $\Rightarrow$ (b) Suppose $Lv = \lambda v$ for some $v \neq 0$. Then $\mathrm{Ker}(L - \lambda \mathrm{id}) \supseteq \{v, 0\} \neq \{0\}$. So $L - \lambda \mathrm{id}$ is not injective.

(b) $\Rightarrow$ (a) Suppose $L - \lambda \mathrm{id}$ is not invertible. Then $L - \lambda \mathrm{id}$ is automatically not injective by 2.3.3. So $\mathrm{Ker}(L - \lambda \mathrm{id}) \neq \{0\}$. Thus, there exists $v \neq 0$ s.t. $Lv = \lambda v$. ☐

$K \mapsto [K]_B$

Remark: If $A = [L]_B$, then

$$[L - \lambda \mathrm{id}]_B = [L]_B - \lambda [\mathrm{id}]_B = A - \lambda I_n.$$

and

$$\det(L - \lambda \mathrm{id}) := \det([L - \lambda \mathrm{id}]_B) = \det(A - \lambda \mathrm{id})$$

Also note that $\det(L - \lambda \mathrm{id}) = 0$ iff $\det(A - \lambda \mathrm{id}) = 0$ so one can perform computations using 4.2.5 and matrices instead.

4.2.6 Theorem Let $A \in F^{n \times n}$ and let x be an indeterminate. Then $\det(xI_n - A)$ is a polynomial in $F[x]$ of degree n. We call it the characteristic polynomial of A and write

$$c_A(x) = \det(xI_n - A) \in F[x].$$

Moreover:

(a) $c_A(x)$ is monic

(b) The coefficient of x^{n-1} is $-\mathrm{tr}(A)$

(c) The constant term is $(-1)^n \det(A)$.

That is, $c_A(x)$ has the form:

$$c_A(x) = x^n - \mathrm{tr}(A)\,x^{n-1} + (\cdots) + (-1)^n \det(A).$$

Further:

(d) The roots of $c_A(x)$ are the eigenvalues of A.

Proof. Using 3.5.9,

$\{1,\ldots,n\}$

$$c_A(x) = \det(xI_n - A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \underbrace{\prod_{i=1}^{n} \left(x\,\delta_{\sigma(i),i} - a_{\sigma(i),i}\right)}_{\text{a polynomial of degree equal to the \# of fixed points } |\{i \mid \sigma(i)=i\}| \text{ of } \sigma}$$

$$= \overbrace{\prod_{i=1}^{n} (x - a_{ii})}^{\text{The only } \sigma \in S_n \text{ w/ } > n-2 \text{ fixed points is } \sigma = \mathrm{id}} + \left(\text{a bunch of polynomials of degree } \leq n-2\right)$$

$$= x^n - (a_{11} + \cdots + a_{nn})\,x^{n-1} + (\text{lower degree terms})$$

$= x^n - \operatorname{tr}(A) x^{n-1} +$ (lower degree terms).

This proves that $c_A(x)$ is a polynomial of degree n and claims (a) and (b). Next,

$$\text{constant term of } c_A(x) = c_A(0)$$
$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (-a_{\sigma(i),i})$$
$$= (-1)^n \det(A).$$

Follows from 4.2.5 and the remark below it since $\det(xI_n - A) = 0$ iff $\det(A - xI_n) = 0$.

$\square$

Remark: If $L : V \longrightarrow V$ is a linear operator, then the characteristic polynomial is defined to be

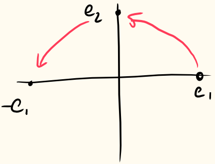
$$c_L(x) := \det(x\,\mathrm{id} - L)$$
$$= \det(xI_n - [L]_B) = c_{[L]_B}(x)$$

where B is any ordered basis for V. We have already seen that this does not depend on the choice of basis. Thm 4.2.6 can be translated into a statement about L.

Remark: Since a polynomial over an field of degree n has at most n roots, Thm 4.2.6 implies that an operator $L : V \to V$ has at most n 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 about the origin by $\pi/2$ radians. Set $E = (e_1, e_2)$

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



So

$$\begin{aligned} c_L(x) &= \det\left(xI_2 - [L]_E\right) \\ &= \det\begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix} \\ &= x^2 + 1. \end{aligned}$$

This polynomial has no roots in $\mathbb{R}$! So it has no eigenvalues. Geometrically, it makes sense - rotation does not expand or contract in any direction. ☐

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

$$\begin{aligned} c_A(x) &= \det\left(xI_3 - A\right) \\ &= \det\begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix} \\ &= (x-3)\begin{vmatrix} x-2 & 1 \\ -2 & x \end{vmatrix} - (-1)\begin{vmatrix} -2 & 1 \\ -2 & x \end{vmatrix} + 1\begin{vmatrix} -2 & x-1 \\ -2 & -2 \end{vmatrix} \\ &= (x-3)(x^2 - x + 2) + (-2x + 2) + (4 + 2(x-1)) \\ &= x^3 - 5x^2 + 8x - 4 \\ &= (x-1)(x-2)^2 \end{aligned}$$

So A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$ (w/ mult 2).

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

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

(row reduce)
$$\sim \begin{pmatrix} -2 & -1 & 1 \\ -2 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} -2 & -1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \underbrace{\begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\text{rref}}$$

So $rk(I-A) = 2$ and $\dim E_1 = 1$.

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

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

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

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

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

So $rk(2I-A) = 2$ and $\dim E_2 = 1$

Specifically, $E_2 = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}\right\}$

$\square$

4.2.9 Example

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

$$c_A(x) = \det(xI - A) = \det\begin{pmatrix} x+1 & 0 & 1 \\ 0 & x+1 & 1 \\ 1 & 1 & x+1 \end{pmatrix}$$

$$= (x+1)\left((x+1)^2+1\right) + (x+1)$$
$$= (x+1)^3$$

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 $\operatorname{rank}(I-A) = 2$ and $\dim E_1 = 1$.

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

$\square$

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. If $k=1$, then $v_1 \neq 0$ is an independent set. Suppose $k \geq 1$ and assume $\sum_{i=1}^{k+1} \alpha_i v_i = 0$. Apply $L - \lambda_{k+1} \mathrm{id}$ to the equation:

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

By inductive hypothesis, $\alpha_i(\lambda_i - \lambda_{k+1}) = 0$ for all $i = 1, \dots, k$. But $\lambda_i \neq \lambda_{k+1}$ by hypothesis, so $\alpha_i = 0$ for all $i = 1, \dots, k$. This implies $0 = \alpha_{k+1} v_{k+1}$. But $v_{k+1} \neq 0$ so $\alpha_{k+1} = 0$. ☐

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

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

Thus,

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

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

Proof. Suffices to show that

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

Everything else follows from facts about direct sums. If $v_i \in E_{\lambda_i}$ and $v_i \in \sum_{i \neq j} E_{\lambda_j}$. Then

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

for some $v_j \in E_{\lambda_j}$, $j \neq i$. This forces $v_i = 0$ by 4.2.10.

□

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.

□

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

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

where $\lambda_1, \dots, \lambda_n$ are pairwise distinct and λ_i appears on the diagonal w/ multiplicity d_i. That is

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

$$= \quad xI - \begin{pmatrix} \lambda_1 I_{d_1} & & & 0 \\ & \lambda_2 I_{d_2} & & \\ & & \ddots & \\ 0 & & & \lambda_n I_{d_n} \end{pmatrix}$$

Then $C_A(x) = \prod_{i=1}^{n} (x - \lambda_i)^{d_i}$. Thus, if L is diagonalizable, then $C_A(x)$ splits into a product of linear polynomials.

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

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

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

$$= \begin{pmatrix} (\lambda_i - \lambda_1) I_{d_1} & & 0 \\ & \ddots & \\ 0 & & (\lambda_i - \lambda_n) I_{d_n} \end{pmatrix}$$

This matrix is diagonal w/ precisely d_i zeros on the diagonal, so $\dim E_{\lambda_i} = d_i$. $\square$

4.2.13 Definition Let $L : V \longrightarrow V$ 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 $c_L(x)$. $\square$

Remark:

The above remark shows that if L is diagonalizable, then for each eigenvalue λ, the algebraic and geometric multiplicities are equal. $\square$

4.2.14 Proposition The geometric multiplicity of λ is at most 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. Call this basis B. Then

$$[L]_B = \begin{pmatrix} \lambda I_k & B \\ 0 & D \end{pmatrix}$$

for some matrices B and C. Then

$$c_L(x) = \det\left(x I_n - [L]_B \right)$$

$$= \det \begin{pmatrix} (x-\lambda)I_k & -B \\ 0 & xI_{n-k}-D \end{pmatrix}$$

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

$$= (x-\lambda)^k c_D(x)$$

thus, k is less than or equal to the algebraic multiplicity. □

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

Proof. (a) ⇒ (d) See the remarks above.

(d) ⇒ (c) If $c_L(x) = \prod_{i=1}^{n}(x-\lambda_i)^{d_i}$, then

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

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

This implies $V = \sum_{i=1}^{n} E_{\lambda_i}$.

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

(b) ⇒ (a) Obvious.

□

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

Method for diagonalizing a matrix Let $A \in F^{n\times n}$.

(1) Compute $c_A(x)$ 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 $c_A(x) = \prod_{i=1}^{n} (x-\lambda_i)^{d_i}$, then determine the dimension of each eigenspace $E_{\lambda_1}, \dots, E_{\lambda_n}$. For each i, compute $\ker(\lambda_i I - A)$ using row-reduction and rank-nullity. If $\dim E_{\lambda_i} < d_i$ for some i, A is not diagonalizable.

(3) If you make it to step three, A is diagonalizable. In that case, A is similar to a diagonal matrix D. You can find an invertible matrix P s.t. $D = P^{-1}AP$ as follows: compute a basis B_i

for each E_{λ_i}. Then $B = (B_1, \dots, B_n)$ is a basis for V consisting of eigenvectors. Set $P = T^{E}_{B}$ where E is the standard basis for F^n. Then

$$\begin{aligned} P^{-1} A P &= T^{B}_{E} \, [A\cdot]_{E} \, T^{E}_{B} \qquad \ulcorner A\cdot : F^n \longrightarrow F^n \\ & \qquad\qquad\qquad\qquad\qquad\quad x \longmapsto Ax \\ &= [A\cdot]^{B}_{B} \\ &= \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_n \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_n \end{pmatrix} \end{aligned}$$

$\square$