3.5.8 Corollary There is a unique alternating n-form

$$\det : \mathbb{F}^{n\times n} = \mathbb{F}^n \times \cdots \times \mathbb{F}^n \longrightarrow \mathbb{F}$$

which satisfies $\det(I_n) = 1$. We call it the **determinant function** on $\mathbb{F}^{n\times n}$.

Proof. By 3.5.7, $\dim \Lambda^n(\mathbb{F}^n) = \binom{n}{n} = 1$. Note that the set $\tilde{I}$ from the proof contains only the identity function. So, a basis is given by $\det := \omega_{id}$. We have

$$\det(v_1, \dots, v_n) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \varphi_i(v_{\sigma(i)})$$

where $\varphi_1, \dots, \varphi_n$ is dual to the standard basis $e_1, \dots, e_n$ for $\mathbb{F}^n$. Since $I_n = (e_1 | \cdots | e_n)$,

$$\begin{aligned}
\det(I_n) &= \det(e_1, \dots, e_n) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, \varphi_1(e_{\sigma(1)}) \cdots \varphi_n(e_{\sigma(n)}) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, \delta_{1,\sigma(1)} \cdots \delta_{1,\sigma(n)} \qquad \left(\begin{array}{l} = 0 \text{ unless} \\ \sigma(i) = i \ \forall i \\ \iff \sigma = Id \end{array}\right) \\
&= \operatorname{sgn}(Id)\, \delta_{1,1} \cdots \delta_{n,n} \\
&= 1.
\end{aligned}$$

Existence is proved. Any other alternating n-form ω will satisfy $\det = \lambda\omega$ for some λ. But $\det(I_n) = 1$ forces $\lambda = 1$. Uniqueness is proved. □

3.5.9 Corollary ("Leibniz Formula" for determinant)

Let $A = (a_{ij}) \in F^{n \times n}$. Then

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Proof. If $\varphi_1, \dots, \varphi_n$ is the dual basis of $e_1, \dots, e_n$, then

$$\varphi_i \begin{pmatrix} a_{1,\sigma(i)} \\ \vdots \\ a_{n,\sigma(i)} \end{pmatrix} = a_{i,\sigma(i)} \qquad \text{for all } i.$$

So the formula follows from the formula given in the proof of 3.5.8.

$\square$

3.5.10 Example

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\sigma \in S_2} \operatorname{sgn} \sigma \prod_{i=1}^{2} a_{i,\sigma(i)}$$

$$= \operatorname{sgn}(id)\, a_{1,1} a_{22} + \operatorname{sgn}(1,2)\, a_{12} a_{21}$$

$$= a_{11} a_{22} - a_{12} a_{21}$$

the usual formula for det of a 2x2 matrix.

$\square$

3.5.11 Proposition Let $A \in F^{n\times n}$. Then $\det(A) = \det(A^T)$ and $\det : F^{n\times n} \to F$ is an alternating multilinear map on the rows of $F^{n\times n}$.

Proof. Viewing det as a function of the "rows" amounts to precomposing w/ $(-)^T : F^{n\times n} \to F^{n\times n}$. So it suffices to prove $\det(A) = \det(A^T)$. Write $A = (a_{ij})$. Then $A^T = (a_{ji})$. Hence,

$$\det(A^T) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i), i}$$

$$= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(\sigma^{-1}(i)), \sigma^{-1}(i)}$$

[Note: $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$]

$$= \sum_{\sigma^{-1}\in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{i=1}^{n} a_{i, \sigma^{-1}(i)} = \det(A). \quad \square$$

3.5.12 Proposition Let $A \in F^{n\times n}$.

(a) Suppose B is obtained from A via $C_i \leftarrow \alpha C_i$ or $R_i \leftarrow \alpha R_i$, $\alpha \neq 0$. Then $\det(B) = \alpha \det(A)$.

(b) Suppose B is obtained from A via $C_i \leftrightarrow C_j$ or $R_i \leftrightarrow R_j$, $i \neq j$. Then $\det(B) = -\det(A)$.

(c) Suppose B is obtained from A via $C_i \leftarrow C_i + \alpha C_j$ or $R_i \leftarrow R_i + \alpha R_j$, $i \neq j$. Then $\det(B) = \det(A)$.

Proof. Statements about column operations follow from Def 3.3.1, Prop 3.5.2(a), and Prop 3.5.2(b), respectively. Statements about rows follow from Prop

3.5.11.

☐

3.5.13 Proposition Let $A \in F^{n\times n}$. The following are equivalent:

(a) $\det(A) \neq 0$

(b) the columns of A are independent in F^n

(c) the rows of A are independent in F^n

(d) A is invertible

(e) $\mathrm{Rref}(A) = I_n$

(f) A is a product of elementary matrices

Proof. Equivalence of (b) – (f) can be deduced using Section 2.9. We prove (a) $\Leftrightarrow$ (b).

(a) $\Rightarrow$ (b) Follows by contrapositive from Prop 3.5.3

(b) $\Rightarrow$ (a) The columns $c_1, \dots, c_n$ of A form a basis for F^n. Assume $[e_i]_{(c_1,\dots,c_n)} = (\beta_{1j}, \dots, \beta_{nj})$. Then

$$1 = \det(e_1, \dots, e_n) = \det\left(\sum_{i=1}^{n} \beta_{i1} c_1, \dots, \sum_{i=1}^{n} \beta_{in}, c_n\right)$$

(expand using multilinearity of det)

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \beta_{i,\sigma(i)} \det(c_1, \dots, c_n)$$

Divide away the constant to see $\det(c_1, \dots, c_n) \neq 0$.

☐

3.5.14 Proposition Let $A, B \in F^{nn}$. Then $\det(AB) = \det(A)\det(B)$. In particular, if A is invertible, then $\det(A^{-1}) = \det(A)^{-1}$.

Proof. If A is invertible, then

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

So it suffices to prove that det is mult. plicative. Easy exercise using elementary matrices.

$\square$

Remark: By 3.5.14, similar matrices have the same determinant:

$$\det(B) = \det(P^{-1}AP) = \det(P)^{-1}\det(A)\det(P) = \det(A).$$

3.5.15 Definition The determinant of a linear operator $L : V \longrightarrow V$ on a finite-dimensional vector space V is defined to be

$$\det(L) := \det([L]_B)$$

where B is any basis for V. This is well-defined by 3.5.14.

$\square$

3.5.16 Proposition If $A \in F^{k\times k}$ and $B^{(n-k)\times(n-k)}$, $n>k$, then

$$\det(A \oplus B) = \det(A)\det(B).$$

Here $A \oplus B$ is the $n \times n$ "block diagonal" matrix

$$A \oplus B := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

Proof. Write $A = (a_{ij})$, $B = (b_{ij})$, and $C = (c_{ij})$. Define

$$T_n := \left\{ \sigma \in S_n \mid \sigma(\{1, \dots, k\}) \subseteq \{1, \dots, k\}, \sigma(\{k+1, \dots, n\}) \subseteq \{k+1, \dots, n\} \right\}$$

Then $c_{i,\sigma(i)} = 0$ whenever $\sigma \notin T_n$. thus,

$$\begin{aligned}
\det(A \oplus B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} c_{i,\sigma(i)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{k} c_{i\sigma(i)} \prod_{i=k+1}^{n} c_{i\sigma(i)} \\
&= \sum_{\rho \in S_k} \sum_{\tau \in S_{n-k}} \operatorname{sgn}(\rho\tau) \prod_{i=1}^{k} c_{i,\rho\tau(i)} \prod_{i=k+1}^{n} c_{i\rho\tau(i)} \\
&= \sum_{\rho \in S_k} \operatorname{sgn}(\rho) \prod_{i=1}^{k} a_{i\rho(i)} \sum_{\tau \in S_{n-k}} \operatorname{sgn}(\tau) \prod_{i=1}^{n-k} b_{i,\tau(i)} \\
&= \det(A)\det(B).
\end{aligned}$$

□

More generally, one can prove

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A)\det(B).$$

However, in general,

$$\det \begin{pmatrix} A & C \\ D & B \end{pmatrix} \neq \det(A)\det(B).$$

Chapter IV: Canonical Forms

4.1 Introduction

Goal: Given a linear operator $L: V \to V$ on a finite-dimensional vector space V, we want to find an ordered basis for V such that $[L]_B$ is as simple as possible.

Motivating Example:

What are the simplest matrices we can think of?

- identity matrix $I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$
- scalar multiples of identity
  $$\lambda I_n = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}$$
- diagonal matrices $\operatorname{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

Suppose $B = (b_1, \dots, b_n)$ is a basis for V such that

$$[L]_B = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Then we already know a lot about L:

(a) Geometrically, L is expanding or contracting in the direction of each basis vector:

$$L(b_i) = \lambda_i b_i \qquad , \; i = 1, \dots, n.$$

(b) Certain invariants of L can be read directly off of the matrix:

- $\det(L) = \det(\{L\}_{B}) = \prod_{i=1}^{n} \lambda_i$
- $\operatorname{tr}(L) = \operatorname{tr}(\{L\}_{B}) = \sum_{i=1}^{n} \lambda_i$
- $\operatorname{im}(L) = \operatorname{span}\{ b_i \mid \lambda_i \neq 0 \}$
- $\ker(L) = \operatorname{span}\{ b_i \mid \lambda_i = 0 \}$

Questions:

- Can every operator be represented by a diagonal matrix for some basis?
- If not, for which operators does such a basis exist?
- How can we compute such a basis, if it exists?
- If such a basis does not exist, can we find a different kind of "simple" matrix to represent L?

To answer these questions, we will use the algebra of linear operators.

Recall:

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

- it is a vector space – one can add and scale linear operators
- one can "multiply" operator $L, K : V \to V$:

$$LK := L\circ K$$

- The operations are all compatible.

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

$$L^i := \underbrace{L\circ L\circ \cdots \circ L}_{i\text{-times}}, \quad L^0 := \mathrm{Id}_V.$$

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

$$\begin{aligned} \mathrm{End}_F(V) &\longrightarrow F^{\dim V \times \dim V} \\ L &\longmapsto [L]_B \end{aligned}$$

where B is a fixed ordered basis. Under this isomorphism, addition, scalar mult, and composition of operators 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 polynomials is commutative.

Observe: Consider a polynomial $p(x) = \sum_{i=1}^{n} a_i x^i \in F[x]$ and a linear operator $L: V \to V$. By (a), we can define a new operator

$$p(L) := \sum_{i=1}^{n} a_i L^i \in \mathrm{End}(V).$$

this defines a function

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

This is actually an algebra homomorphism!

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

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

is an F-algebra homomorphism.

Proof. A basis for $F[x]$ is $\{x^i \mid i \geq 0\}$. There is a unique linear map Φ_L sending $x^i \longmapsto L^i$ for $i \geq 0$. Thus, if $p(x) = \sum_{i=1}^{n} a_i x^i$, then

$$\Phi_L(p(x)) = \sum_{i=1}^{n} a_i \, \Phi_L(x^i) = \sum_{i=1}^{n} a_i L^i = p(L).$$

So Φ_L is linear as claimed. It suffices by linearity to show multiplicativity on a basis. We have

$$\begin{aligned} \Phi_L(x^i \cdot x^j) = \Phi_L(x^{i+j}) &= L^{i+j} \\ &= L^i L^j \\ &= \Phi_L(x^i)\,\Phi_L(x^j). \end{aligned}$$

So Φ_L is an algebra homomorphism. □

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 immediately from 4.1.1 since $F[x]$ is a commutative algebra.

□

Remark: In general, operators do not commute:

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

We will need to fact that $p(L)$ and $q(L)$ commute later.

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

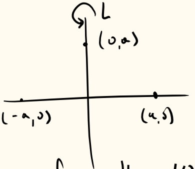
$$Lv = \lambda v$$

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 $L((0,a)) = (0,a)$

$L((a,0)) = -(a,0)$



So $(0,a)$ is an eigenvector w/ eigenval 1 and $(a,0)$ is an eigenvector w/ eigenvalue -1 for all $a \neq 0$.

$\square$

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

Proof. Obvious.

$\square$

4.2.4 Definition The subspace

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

of eigenvectors corresponding to λ, together w/ the zero vector, is called the eigenspace of L associated λ.

$\square$

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 $\ker(L - \lambda id) \supseteq \{v\} \neq \{0\}$ so $L - \lambda id$ is not injective.

(b) $\Rightarrow$ (a) Suppose $L - \lambda id$ is not invertible. By 2.3.3 L is not injective and not surjective. So $\ker(L - \lambda id) \neq \{0\}$. There exists $v \neq 0$ s.t. $Lv = \lambda v$.

□

Remark: If $A = [L]_B$, then $[L - \lambda id]_B = [L]_B - \lambda [id]_B$
$= A - \lambda I_n$

and

$$\det(L - \lambda id) = \det(A - \lambda I_n)$$

Also note, $\det(L - \lambda id) = 0$ iff $\det(A - \lambda I_n) = 0$. So one can perform the necessary computations in 4.2.5 using matrices.

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

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

Moreover:

(a) $cp_A(x)$ is monic

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

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

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

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

Further:

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

Proof. Using 3.5.9,

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

$$= \overbrace{\prod_{i=1}^{n} (x - a_{ii})}^{\substack{\text{The only permutation} \\ \text{fixing } > n-2 \text{ pts is} \\ id}} + (\text{a bunch of polynomials of degree } \le n-2)$$

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

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

This proves (a) and (b) and that $cp_A(x)$ has degree n. To prove (c),

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

Claim (d) follows from 4.2.5 and the remark below it. □

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

$$\begin{aligned} cp_L(x) &:= \det(x\,\mathrm{id} - L) \\ &= \det([x\,\mathrm{id} - L]_B) \end{aligned}$$

where B is any basis. We have already proved that this does not depend on the choice of basis. Thm 4.2.6 can be translated directly into a statement about L. For instance, the roots of $cp_L(x)$ are precisely the eigenvalues of L.