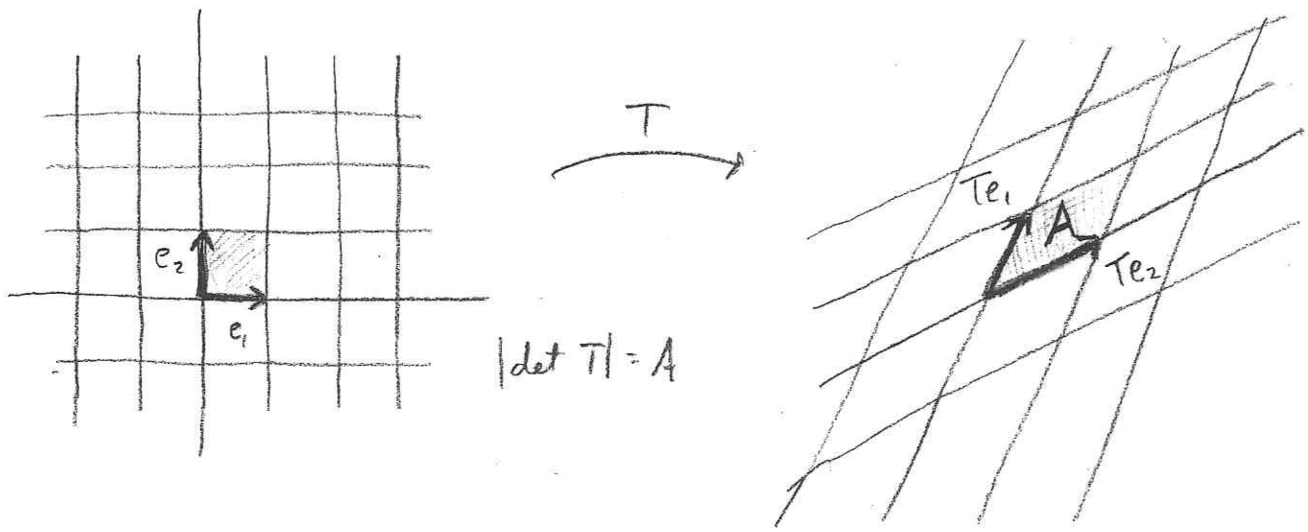
5. Determinant and trace.

The concept of the determinant of a linear map is simple – we will give an unofficial geometric definition as motivation, and then formalize it.

Def: (Unofficial). Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear map, and $[0,1]^n$ the unit n-cube. The determinant of T is the "signed volume" of $T([0,1]^n)$.



Our unofficial definition has some interesting properties:

- $\det T = 0$ iff T is not invertible
- $\det(TS) = (\det T)(\det S)$
- If T and S differ by swapping two columns, then $\det T = -\det S$.
- $\det T$ is "linear" in each column – i.e, if T is obtained from S by multiplying a column by c, then $\det T = c(\det S)$.

Permutations:

Def: Let $[n] := \{1, \dots, n\}$. A permutation is a bijection $\pi : [n] \longrightarrow [n]$. The set of all permutations of n elements is denoted S_n, and is a group.

We can describe a permutation by a table, or more concisely, by cycle notation:

Example: $\pi : \{1,2,3,4\} \longrightarrow \{1,2,3,4\}$.

i	1	2	3	4
$\pi(i)$	2	4	1	3

Table notation:

$$\pi = \frac{1234}{2413}, \quad \pi^2 = \frac{1234}{4321}, \quad \pi^3 = \frac{1234}{3142} = \pi^{-1}, \quad \pi^4 = \frac{1234}{1234}$$

Cycle notation: $\pi = (1\ 2\ 4\ 3)$ meaning $1 \to 2 \to 4 \to 3$

$$\pi^2 = (1\ 4)(2\ 3), \quad \pi^3 = (1\ 3\ 4\ 2), \quad \pi^4 = (1)(2)(3)(4).$$

By convention, we usually omit length-1 cycles, e.g., $(1 2)(3) = (1 2)$.

Def: Let $x_1, \dots, x_n$ be n variables. Their discriminant is defined as $P(x_1, \dots, x_n) = \prod_{i<j} (x_i - x_j)$.

For a permutation $\pi \in S_n$,

$$P(\pi(x_1, \dots, x_n)) := \prod_{i<j} (x_{\pi(i)} - x_{\pi(j)}) = \pm P(x_1, \dots, x_n).$$

<u>Def:</u> The <u>signature</u> $sgn(\pi)$ of a permutation $\pi \in S_n$ is defined as $P(\pi(x_1,\dots,x_n)) = sgn(\pi)\, P(x_1,\dots,x_n)$.

Clearly, $sgn(\pi) = \pm 1$.

<u>Def:</u> A <u>transposition</u> is a permutation $\tau \in S_n$ such that for some $j \neq k \in [n]$,

$$\tau(i) = i \qquad i \neq j, k$$
$$\tau(j) = k$$
$$\tau(k) = j.$$

<u>Prop:</u> (i) $sgn(\pi_1 \circ \pi_2) = sgn(\pi_1)\, sgn(\pi_2)$

(ii) $sgn(\tau) = -1$ for any transposition

(iii) Every permutation $\pi \in S_n$ can be written as a composition of transpositions: $\pi = \tau_k \circ \dots \circ \tau_1$

(iv) This decomposition is not unique, but the parity of k is.

(v) If $\pi = \tau_k \circ \dots \circ \tau_1$, then $sgn(\pi) = (-1)^k$.

<u>Proof:</u> Exercise.

Multilinear forms

<u>Def:</u> A <u>k-linear form</u> is a function $f: X_1 \times \dots \times X_k \to K$ (we'll usually assume $X_1 = \dots = X_k = X$) that is linear in each coordinate, i.e., upon fixing $k-1$ arguments, it is linear in the remaining argument.

Examples:

(1) 1-linear forms are linear functions in X'

(2) 2-linear forms are bilinear forms.

(3) A 3-linear form $X \times X \times X \to K$ has identities such as

$f(a_1 x_1 + a_2 x_2, y, z) = a_1 f(x_1, y, z) + a_2 f(x_2, y, z)$, and similarly

for $f(x, a_1 y_1 + a_2 y_2, z)$ & $f(x, y, a_1 z_1 + a_2 z_2)$.

Theorem 5.1: The set of k-linear forms is a vector space of dimension n^k. (Assuming $\dim X = n$)

Proof: (sketch) Verify that a basis consists of functions

$\{ f_{j_1 \cdots j_k}(x_{i_1}, \dots, x_{i_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k} : 1 \le j_\ell \le n \}$. $\square$

Example: If $f : X \times X \to K$, $\dim X = 2$ with basis $\{x_1, x_2\}$,

then
$$\begin{aligned} f(u,v) &= f(a_1 x_1 + a_2 x_2,\ b_1 x_1 + b_2 x_2) \\ &= a_1 b_1 f(x_1, x_1) + a_1 b_2 f(x_1, x_2) + a_2 b_1 f(x_2, x_1) + a_2 b_2 f(x_2, x_2) \\ &:= c_{11} f(x_1, x_1) + c_{12} f(x_1, x_2) + c_{21} f(x_2, x_1) + c_{22} f(x_2, x_2). \end{aligned}$$

Basis: $f_{11}(x_i, x_j) = \begin{cases} 1 & i=1,\ j=1 \\ 0 & \text{else} \end{cases}$ $\quad f_{12}(x_i, x_j) = \begin{cases} 1 & i=1,\ j=2 \\ 0 & \text{else} \end{cases}$

$f_{21}(x_i, x_j) = \begin{cases} 1 & i=2,\ j=1 \\ 0 & \text{else} \end{cases}$ $\quad f_{22}(x_i, x_j) = \begin{cases} 1 & i=2,\ j=2 \\ 0 & \text{else} \end{cases}$

For any permutation $\pi \in S_k$, define $\pi f : X \times \cdots \times X \to K$ by

$$(\pi f)(x_1, \dots, x_k) = f(x_{\pi(1)}, \dots, x_{\pi(k)})$$

Note that for any k-linear form f, πf is k-linear.

<u>Def:</u> A k-linear form is <u>symmetric</u> if $\pi f = f$ for all $\pi \in S_k$.

<u>Examples:</u>

(i) $f(x_1, x_2) = l_1(x_1) l_2(x_2) + l_1(x_2) l_2(x_1)$ for fixed $l_1, l_2 \in X'$

(ii) $f(x_1, \dots, x_k) = \sum_{k \in S_k} \pi f(x_1, \dots, x_k)$

(iii) Standard "dot product": $f(u, v) = u \cdot v$.

<u>Def:</u> A k-linear form is <u>skew-symmetric</u> if $\tau f = -f$ for every transposition $\tau \in S_k$.

<u>Example:</u> $f(x_1, x_2) = l_1(x_1) l_2(x_2) - l_1(x_2) l_2(x_1)$.

<u>Def:</u> A k-linear form is <u>alternating</u> if $f(x_1, \dots, x_k) = 0$ if $x_i = x_j$ for some $i \neq j$.

<u>Prop:</u> The set of alternating (resp. symmetric, or skew-symmetric) k-linear forms is a subspace

<u>Proof:</u> Exercise.

<u>Theorem 5.2</u>: Every alternating multilinear form is skew-symmetric.

<u>Proof</u>: Pick $i \neq j$, define $g(x_i, x_j) = f(x_1, \ldots, x_n)$ (i.e., the other entries are fixed). Note that g is bilinear, alternating.

Thus, $0 = g(x_i + x_j, x_i + x_j) = g(x_i, x_i) + g(x_j, x_j) + g(x_i, x_j) + g(x_j, x_i)$
$= g(x_i, x_j) + g(x_j, x_i)$

$\Rightarrow g(x_i, x_j) = -g(x_j, x_i)$

$\Rightarrow \tau f = -f$ for $\tau = (i\ j)$

<u>Remark</u>: The converse "almost holds":

Suppose $f = -f$. Then $(1+1)f = 0 \Rightarrow f = 0$ or $1+1=0$ (e.g., if $K = \mathbb{Z}_2$). Thus, the converse holds for all fields of "characteristic" not equal to 2.

<u>Theorem 5.3</u> If $x_1, \ldots, x_k$ are linearly dependent and f is an alternating k-linear form, then $f(x_1, \ldots, x_n) = 0$

<u>Proof</u>: If $x_i = 0$, result is trivial.

Otherwise, we can write $x_j = \sum_{i \neq j} a_i x_i$ for some x_j, and replace x_j in $f(x_1, \ldots, x_k)$ by this sum.

Use multilinearity to break apart $f(x_1, \dots, x_k)$ into a sum, each one having some x_i in the j^{th} position, $i \neq j$.

All of these terms are zero since f is alternating. $\square$

The converse holds in one special case: $k = n$.

<u>Theorem 5.4</u>: If f is a non-zero alternating n-linear form, and $x_1, \dots, x_n$ are linearly independent, then $f(x_1, \dots, x_n) \neq 0$.

<u>Proof</u>: By assumption, $x_1, \dots, x_n$ is a basis for X.

Pick $y_1, \dots, y_n \in X$, write $y_i = a_{i_1} x_1 + \dots + a_{in} x_n$.

$$
\begin{aligned}
\text{Now, } f(y_1, \dots, y_n) &= f\left(\sum a_{1j} x_{1j}, \dots, \sum a_{nj} x_{nj} \right) \\
&= \sum_{\pi \in S_n} C_\pi \cdot f(x_{\pi(1)}, \dots, x_{\pi(n)}) \\
&= \sum_{\pi \in S_n} C_\pi \cdot \operatorname{sgn}(\pi)\, f(x_1, \dots, x_n).
\end{aligned}
$$

$$
\begin{aligned}
\text{If } f(x_1, \dots, x_n) = 0 &\implies f(y_1, \dots, y_n) = 0 \quad \forall y_1, \dots, y_n \in X \\
&\implies f \equiv 0. \quad \lightning
\end{aligned}
$$

$\square$

<u>Remark</u>: The proof of Theorem 5.4 yields an important corollary:

<u>Theorem 5.5</u>: Any two alternating n-linear forms are linearly dependent.

Proof: Let f, g be alternating n-linear forms, and let $x_1, \dots, x_n$ be a basis of X.

Let $\lambda \in K$ be such that $g(x_1, \dots, x_n) = \lambda \cdot f(x_1, \dots, x_n)$.

Claim: $g(y_1, \dots, y_n) = \lambda \cdot f(y_1, \dots, y_n)$ for all $y_1, \dots, y_n \in X$.

Indeed, $f(y_1, \dots, y_n) = \sum_{\pi \in S_n} C_\pi \cdot \mathrm{sgn}(\pi) \cdot f(x_1, \dots, x_n)$

and $g(y_1, \dots, y_n) = \sum_{\pi \in S_n} C_\pi \cdot \mathrm{sgn}(\pi) \cdot g(x_1, \dots, x_n)$

$$= \sum_{\pi \in S_n} C_\pi \cdot \mathrm{sgn}(\pi) \cdot \lambda \cdot f(x_1, \dots, x_n)$$

$$= \lambda f(x_1, \dots, x_n) \quad \checkmark \qquad \square$$

Cor: The subspace of alternating n-linear forms is at most one-dimensional.

Still need to show there is a non-zero alternating n-linear form.

Remark: There is no non-zero alternating $(n+1)$-linear form. (Why?)

Theorem 5.6: The vector space of alternating n-linear forms is one-dimensional.

Proof: It suffices to construct a non-zero alternating n-linear form. We'll fix n, then use induction on $k \le n$.

Motivation: Recall that $f(x_1, x_2) = l_1(x_1) l_2(x_2) - l_1(x_2) l_2(x_1)$ is alternating.

<u>Base case:</u> $(k=1)$: Any $\ell \neq 0$ in X' is an alternating 1-linear form. ✓

Now, suppose f is a non-zero alternating k-linear form, $k < n$.

We'll construct a non-zero alternating $(k+1)$-linear form g.

Pick $y_1, \dots, y_k$ such that $f(y_1, \dots, y_k) \neq 0$.

Find $y_{k+1} \in X$ <u>not</u> in subspace $Y := \mathrm{Span}\{y_1, \dots, y_k\}$.

Pick $\ell \in Y^\perp$, so that $\ell(y_{k+1}) \neq 0$. (Note that $\ell(y_1), \dots, \ell(y_k) = 0$)

Define $g(x_1, \dots, x_{k+1}) = -f(x_1, \dots, x_k)\,\ell(x_{k+1}) + \sum_{i=1}^{k} (i\ k{+}1)\, f(x_1, \dots, x_k)\,\ell(x_{k+1})$.

[<u>Example</u>: If $k=3$, then

$$g(x_1, x_2, x_3, x_4) = -f(x_1, x_2, x_3)\,\ell(x_4)$$
$$+ f(x_4, x_2, x_3)\,\ell(x_1) + f(x_1, x_4, x_3)\,\ell(x_2) + f(x_1, x_2, x_4)\,\ell(x_3)\Big]$$

Clearly this is k-linear.

<u>Non-zero:</u> $g(y_1, \dots, y_{k+1}) = -\overbrace{f(y_1, \dots, y_k)}^{\neq 0}\,\overbrace{\ell(y_{k+1})}^{\neq 0} \neq 0$ ✓

<u>Alternating:</u> Consider vectors $x_1, \dots, x_{k+1}$ such that $x_i = x_j$ for some $i < j$.

Need to prove $g(x_1, \dots, x_{k+1}) = 0$.

<u>Case 1:</u> $(j = k+1)$

$$g(x_1, \dots, x_{k+1}) = (i\ k{+}1)\, f(x_1, \dots, x_k)\,\ell(x_{k+1}) - f(x_1, \dots, x_k)\,\ell(x_{k+1})$$
$$= 0 \quad (\text{because } x_i = x_{k+1}). \quad ✓$$

<u>Case 2</u>: $(j \leq k)$

$$g(x_1,\dots,x_{k+1}) = \left[(i\ j)\, f(x_1,\dots,x_k) - f(x_1,\dots,x_k)\right] \ell(x_{k+1})$$

$$= 0 \quad \text{(because } f \text{ is alternating)} \qquad \square$$

Let f be an alternating n-linear form on X, and $T: X \to X$ be a linear map.

Define a new alternating n-linear form $\bar{T}f : X^n \longrightarrow K$ as

$$(\bar{T}f)(x_1,\dots,x_n) := f(Tx_1,\dots,Tx_n).$$

* This defines a linear map $\bar{T}$ on the (one-dimensional) space of all such forms: $\bar{T}: f \longmapsto \lambda f$ for some $\lambda \in K$.

<u>Def</u>: This scalar is the <u>determinant</u> of T, denoted $\det T$.

The determinant satisfies the following.

<u>Universal property</u>: Given a linear map $T: X \to X$, there exists a unique scalar $\lambda \in K$ such that for every alternating n-linear form f,

$$f(Tx_1,\dots,Tx_n) = \lambda f(x_1,\dots,x_n)$$

$$\begin{tikzcd} X^n \arrow[r, "T\times\cdots\times T"] \arrow[d, "f"'] & X^n \arrow[d] \\ K \arrow[r, "\lambda"] & K \end{tikzcd}$$

<u>Remark</u>: This definition is not only independent of matrices, but also of choice of basis!

<u>Example</u>: If $Tx = cx$, then

$$(\overline{T}f)(x_1, \dots, x_n) = f(cx_1, \dots, cx_n) = c^n f(x_1, \dots, x_n).$$

Thus, $\det T = c^n$. It follows that $\det 0 = 0$, $\det I = 1$.

<u>Theorem 5.7</u>: For any two linear maps $A, B : X \to X$,

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

<u>Proof</u>: Write $C = AB$. For an alternating n-linear form f:

$$(\overline{C}f)(x_1, \dots, x_n) = f(ABx_1, \dots, ABx_n)$$
$$= (\overline{A}f)(Bx_1, \dots, Bx_n) = \overline{B}\,\overline{A}f(x_1, \dots, x_n).$$

Thus, $\overline{C} = \overline{B}\,\overline{A}$. We have $\overline{C}f = (\det C)f$ and

$$\overline{C}f = \overline{B}\,\overline{A}f = (\det B)\overline{A}f = (\det B)(\det A)f.$$

$$\Rightarrow \det(AB) = (\det A)(\det B).$$ □

<u>Corollary</u>: $A : X \to X$ is invertible iff $\det A \neq 0$.

<u>Proof</u>: If A is invertible, then $AA^{-1} = I$

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

Conversely, if $\det A \neq 0$, and $x_1, \dots, x_n$ is a basis of X, and f a non-zero alternating n-linear form on X, then

$$0 \underset{\uparrow\ \text{Thm 5.4}}{\neq} (\det A) f(x_1, \dots, x_n) \underset{\uparrow\ \text{def'n}}{=} f(Ax_1, \dots, Ax_n)$$

Since this is non-zero, $Ax_1, \dots, Ax_n$ is linearly independent by Thm 5.3. Thus A is invertible. $\square$

Determinants and matrices

The determinant of an $n \times n$ matrix can be thought of as an alternating n-linear function of its column vectors.

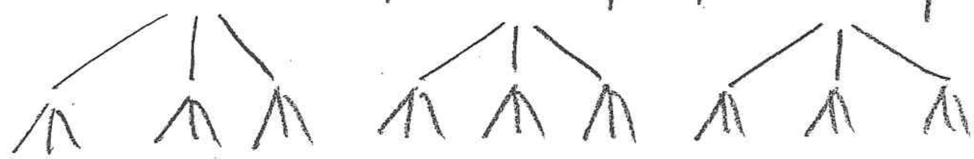
Examples:

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Denote $\det A$ by $|A|$.

Note that $\begin{vmatrix} a+a' & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b \\ c & d \end{vmatrix}$ by bilinearity

Now,
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$
$$= \left(\begin{vmatrix} a_{11} & 0 \\ a_{21} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} \right) + \left(\begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} \\ 0 & a_{22} \end{vmatrix} \right)$$
$$= a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{12}a_{22} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = \boxed{a_{11}a_{22} - a_{12}a_{22}}$$

<u>3×3:</u> $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & 0 \\ & — & \\ & — & \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ & — & \\ & — & \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ & — & \\ & — & \end{vmatrix}$



Non-zero terms: (one for each <u>permutation</u> matrix)

$$\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{33} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

<u>n×n Formula:</u> $$\det A = \sum_{\pi \in S_n} (\operatorname{sgn} \pi)\, a_{1,\pi(1)}\, a_{2,\pi(2)} \cdots a_{n,\pi(n)}$$

(by symmetry) $$= \sum_{\pi \in S_n} (\operatorname{sgn} \pi)\, a_{\pi(1),1}\, a_{\pi(2),2} \cdots a_{\pi(n),n}$$

$$= \det A^T.$$

Thus, the determinant can also be thought of as an alternating n-linear function of its <u>row vectors</u>. It is the unique such "normalized" function, in that $f(e_1, \dots, e_n) = 1$ ($\det I = 1$).

<u>Lemma 5.8:</u> Let $A = [c_1, \dots, c_n]$ be a matrix (c_i is a column vector), and B be obtained by adding $k c_i$ to the j^{th} column of A, $i \neq j$. Then $\det A = \det B$.

Proof: Exercise. (HW)

lemma

Lemma 5.9: Let A be an $n \times n$ matrix whose first column is e_1:

$$A = \begin{pmatrix} 1 & x \; x \; \cdots \; x \\ 0 & \\ \vdots & A_{11} \\ 0 & \end{pmatrix}. \qquad \text{Then } \det A = \det A_{11}.$$

Proof: By lemma 5.8, $\det A = \det \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & A_{11} \\ 0 & \end{pmatrix}$.

Define $f(A_{11}) = \det \begin{pmatrix} 1 & 0 \\ 0 & A_{11} \end{pmatrix}$, a function of an $(n-1)\times(n-1)$ matrix A_{11}.

Clearly, f is an alternating n-linear form with $f(I) = 1$, so it must be the determinant function.

Corollary 5.10: Let A be a matrix whose j^{th} column is e_i.

Then $\det A = (-1)^{i+j} \det A_{ij}$, where A_{ij} is the $(n-1)\times(n-1)$ matrix obtained by striking out the i^{th} row & j^{th} column of A (called the (i,j)-minor of A.)

Proof: Exercise.

Theorem 5.11: Let A be an $n \times n$ matrix and $1 \le j \le n$. Then

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} (\det A_{ij}) \qquad (\text{The "Laplace expansion"})$$

Proof: To simplify notation, take $j=1$.

Write $a_1 = a_{11}e_1 + \cdots + a_{n1}e_n$, $A = [a_1, \cdots, a_n]$

By multilinearity, $\det A = f(a_1, \ldots, a_n)$

$$= f(a_{11}e_1 + \cdots + a_{n1}e_n, a_2, \ldots, a_n)$$

$$= a_{11} f(e_1, a_2, \ldots, a_n) + \cdots + a_{n1} f(e_n, a_2, \ldots, a_n),$$

Now apply Corollary 5.10. $\square$

Application: Systems of equations.

Consider a system $Ax = u$, A invertible, $x = \sum_{j=1}^{n} x_j e_j$,

and $A = (a_1, \ldots, a_n) = (Ae_1, \ldots, Ae_n)$.

$Ax = u$ can be written $\sum_{j=1}^{n} x_j a_j = u$.

Let $A_k = (a_1, \ldots, a_{k-1}, u, a_{k+1}, \ldots, a_n)$ for each k.

$$= \left(a_1, \ldots, a_{k-1}, \sum_{j=1}^{n} x_j a_j, a_{k+1}, \ldots, a_n\right)$$

By multilinearity, $\det A_k = \sum_{j=1}^{n} x_j \det(a_1, \ldots, a_{k-1}, a_j, a_{k+1}, \ldots, a_n)$

$$= x_k \det(a_1, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots, a_n)$$

$$= x_k \det A.$$

Since A is invertible, $\det A \neq 0 \implies x_k = \dfrac{\det A_k}{\det A}$

Now apply the Laplace expansion (Thm 5.11) to A_k:

$$\det A_k = \sum_{i=1}^{n} (-1)^{i+k} \det A_{ik}\, u_i$$

$$\Rightarrow x_k = \sum_{i=1}^{n} (-1)^{i+k} \frac{\det A_{ik}}{\det A} u_i$$ This is "<u>Cramer's rule</u>".

<u>Theorem 5.12</u>: If A is invertible, then its inverse A^{-1} has the form $(A^{-1})_{ki} = (-1)^{i+k} \dfrac{\det A_{ik}}{\det A}$.

<u>Proof</u>: Consider a system $Ax = u \Rightarrow x = A^{-1}u$.

By Cramer's rule, $\sum_{i=1}^{n} (-1)^{i+k} \dfrac{\det A_{ik}}{\det A} u_i = x_k = (A^{-1}u)_k = \sum_{i=1}^{n} (A^{-1})_{ki}\, u_i$. □

<u>Note</u>: This formula is computationally impractical for computing inverses.

<u>Def</u>: The <u>trace</u> of an $n \times n$ matrix is $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$.

<u>Theorem 5.13</u>:

(a) Trace is linear: $\operatorname{tr}(kA) = k(\operatorname{tr} A)$.

$$\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B.$$

(b) Trace is "commutative": $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

<u>Proof</u>: (a) is obvious.

(b) $(AB)_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}$ and $(BA)_{ii} = \sum_{k=1}^{n} b_{ik} a_{ki}$.

Thus, $\operatorname{tr}(AB) = \sum_{i,k} a_{ik} b_{ki} = \sum_{i,k} b_{ik} a_{ki} = \operatorname{tr}(BA)$. $\square$

Theorem 5.14: Similar matrices have the same determinant & trace.

Proof: Suppose $A = SBS^{-1}$.

$$\det A = \det(SBS^{-1}) = (\det S)(\det B)(\det S^{-1})$$
$$= (\det B)(\det S)(\det S)^{-1} = \det B. \checkmark$$

$$\operatorname{tr} A = \operatorname{tr}(SBS^{-1}) = \operatorname{tr}(S^{-1}SB) = \operatorname{tr} B. \checkmark \qquad \square$$

Since similar matrices represent the same linear map but with a different choice of basis, it is well-founded to speak of the determinant and trace as functions of linear maps, not just matrices.