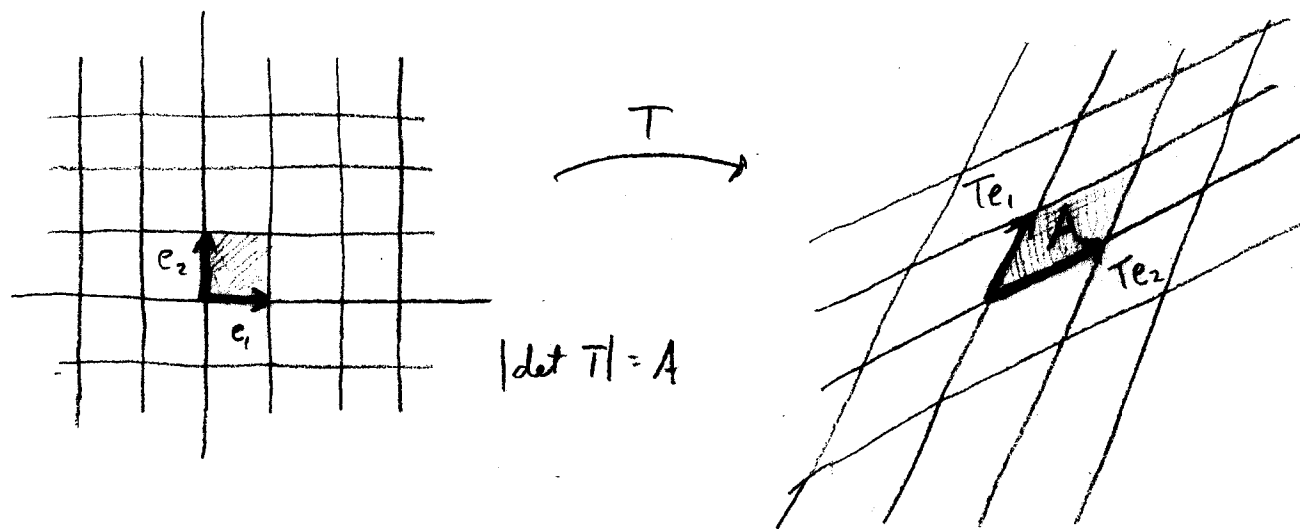


5. Determinant and trace.

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 $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $\det T = c(\det S)$.

[2]

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:

$$\pi = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{array}, \quad \pi^2 = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{array}, \quad \pi^3 = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{array} = \pi^{-1}, \quad \pi^4 = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{array}$$

Cycle notation: $\pi = (1 \ 2 \ 4 \ 3)$ meaning $1 \xrightarrow{\pi} 2 \xrightarrow{\pi} 4 \xrightarrow{\pi} 3 \xrightarrow{\pi} 1$

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

By convention, we usually omit length-1 cycles, e.g., $(12)(3) = (12)$.

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 $\pi \in S_n$,

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

Def: The signature $\text{sgn}(\pi)$ of a permutation $\pi \in S_n$ is defined as $P(\pi(x_1, \dots, x_n)) = \text{sgn}(\pi) P(x_1, \dots, x_n)$.

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

Def: A transposition is a permutation $\tau \in S_n$ such that for

$$\begin{aligned} \text{Some } j \neq k \in [n], \quad & \tau(i) = i \quad i \neq j, k \\ & \tau(j) = k \\ & \tau(k) = j. \end{aligned}$$

Prop: (i) $\text{sgn}(\pi_1 \circ \pi_2) = \text{sgn}(\pi_1) \text{sgn}(\pi_2)$

(ii) $\text{sgn}(\tau) = -1$ 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 $\text{sgn}(\pi) = (-1)^k$.

Proof: Exercise.

Multilinear forms

Def: A k-linear form is a function $f: X_1 \times \dots \times X_k \rightarrow K$ (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.

(4)

Examples:

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

(2) 2-linear forms are bilinear forms.

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

$$f(a_1 x_1 + a_2 x_2, y, z) = a_1 f(x_1, y, z) + a_2 f(x_2, y, z), \text{ and similarly}$$

$$\text{for } f(x, a_1 y + a_2 y_2, z) \text{ \& } f(x, y, a_1 z + a_2 z_2).$$

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

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

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

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

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

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

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

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

For any permutation $\pi \in S_k$, define $\pi f : X \times \dots \times X \rightarrow K$ 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.

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

Example:

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

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

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

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

Example: $f(x_1, x_2) = l_1(x_1)l_2(x_2) - l_1(x_2)l_2(x_1).$

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

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

Proof: Exercise.

(c)

Theorem 5.2: Every alternating multilinear form is skew-symmetric.

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

$$\begin{aligned} \text{Thus, } 0 &= g(x_i + x_j, x_i + x_j) = g(x_i, x_i) + g(x_j, x_j) + g(x_i, x_j) + g(x_j, x_i) \\ &= g(x_i, x_j) + g(x_j, x_i) \end{aligned}$$

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

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

Remark: 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.

Theorem 5.3 If $x_1, \dots, x_n$ are linearly dependent and f is an alternating k -linear form, then $f(x_1, \dots, x_n) = 0$

Proof: 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 $f(x_1, \dots, x_n)$ 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: $k=n$.

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

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

Pick $y_1, \dots, y_n \in X$, write $y_i = a_{i1}x_1 + \dots + a_{in}x_n$.

$$\text{Now, } f(y_1, \dots, y_n) = f(\sum a_{1j}x_j, \dots, \sum a_{nj}x_j)$$

$$= \sum_{\pi \in S_n} C_\pi f(x_{\pi(1)}, \dots, x_{\pi(n)})$$

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

$$\text{If } f(x_1, \dots, x_n) = 0 \Rightarrow f(y_1, \dots, y_n) = 0 \quad \forall y_1, \dots, y_n \in X \Rightarrow f = 0 \quad \zeta \quad \square$$

Remark: The proof of Theorem 5.4 yields an important corollary:

Theorem 5.5: Any two alternating n -linear forms are linearly dependent.

(8)

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

Pick any $y_1, \dots, y_n \in X$. As before:

$$f(y_1, \dots, y_n) = \sum_{\pi \in S_n} c_\pi \cdot \text{sgn}(\pi) f(x_{1_\pi}, \dots, x_{n_\pi}) = a \cdot f(x_1, \dots, x_n) \in K.$$

$$\begin{aligned} g(y_1, \dots, y_n) &= \sum_{\pi \in S_n} d_\pi \cdot \text{sgn}(\pi) f(x_{1_\pi}, \dots, x_{n_\pi}) = b \cdot f(x_1, \dots, x_n) \\ &= \lambda \cdot a \cdot f(x_1, \dots, x_n) \end{aligned}$$

Thus, $g(y_1, \dots, y_n) = \lambda f(y_1, \dots, y_n)$ for any $y_1, \dots, y_n \in X$

$$\Rightarrow g = \lambda f.$$

□

* Thus, the subspace of alternating n -linear forms is at most one-dim.

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 $k \leq n$.

(9)

Base case: ($k=1$): Any $l \neq 0$ 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 $y_{k+1} \in X$ not in subspace $Y := \text{Span}\{y_1, \dots, y_k\}$.

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

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

[Example: If $k=3$, then

$$g(x_1, x_2, x_3, x_4) = -f(x_1, x_2, x_3) l(x_4) + f(x_4, x_2, x_3) l(x_1) + f(x_1, x_4, x_3) l(x_2) + f(x_1, x_2, x_4) l(x_3)$$

Clearly, this is k -linear.

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

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

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

Case 1: ($j = k+1$)

$$\begin{aligned} g(x_1, \dots, x_{k+1}) &= (i - k + 1) f(x_1, \dots, x_k) l(x_{k+1}) - f(x_1, \dots, x_k) l(x_{k+1}) \\ &= 0 \quad (\text{because } x_i = x_{k+1}). \quad \checkmark \end{aligned}$$

(10)

Case 2: $(j \leq k)$

$$\begin{aligned}
 g(x_1, \dots, x_{k+1}) &= [(i \ j) f(x_1, \dots, x_k) - f(x_1, \dots, x_k)] x_{k+1} \\
 &= 0 \quad (\text{because } f \text{ is alternating})
 \end{aligned}$$

□

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

Define a new alternating n -linear form $\bar{T}f: X^n \rightarrow K$ 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: $\bar{T}: f \mapsto \lambda f$ for some $\lambda \in K$.

Def: This scalar is the determinant of T , denoted $\det T$.

The determinant satisfies the following:

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

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

$$\begin{array}{ccc}
 X^n & \xrightarrow{T \times \dots \times T} & X^n \\
 f \downarrow & & \downarrow \\
 K & \xrightarrow{\lambda} & K
 \end{array}$$

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

Example: 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 $\det 0 = 0$, $\det I = 1$.

Theorem 5.7: For any two linear maps $A, B: X \rightarrow X$,

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

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

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

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

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

$$\Rightarrow \det(AB) = (\det A)(\det B). \quad \square$$

Corollary: $A: X \rightarrow X$ is invertible iff $\det A \neq 0$.

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

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

[12]

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

$$0 \neq (\det A) f(x_1, \dots, x_n) = f(Ax_1, \dots, Ax_n)$$

↑ Thm 5.4

↑ def'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.

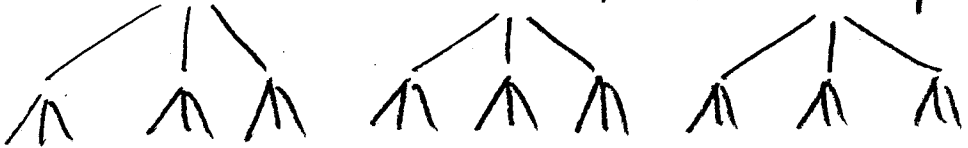
Examples:

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

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

3x3:
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & 0 \\ - & - & - \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ - & - & - \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ - & - & - \end{vmatrix}$$



Non-zero terms: (one for each permutation matrix)

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

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

n x n Formula: $\det A = \sum_{\pi \in S_n} (\text{sgn } \pi) a_{1, \pi(1)} a_{2, \pi(2)} \cdots a_{n, \pi(n)}$

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

$$= \det A^T.$$

Thus, the determinant can also be thought of as an alternating

n-linear function of its row vectors. It is the unique

such "normalized" function, in that $f(e_1, \dots, e_n) = 1$ ($\det I = 1$).

Lemma 5.8: 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 $\det A = \det B$.

(14)

Proof: Exercise. (HW)

Lemma

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

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

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

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_j .

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 $1 \leq j \leq n$. Then

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

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

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

$$\begin{aligned} \text{By multilinearity, } \det A &= f(a_1, \dots, a_n) \\ &= f(a_{11}e_1 + \dots + a_{n1}e_n, a_2, \dots, a_n) \\ &= a_{11}f(e_1, a_2, \dots, a_n) + \dots + a_{n1}f(e_n, a_2, \dots, a_n). \end{aligned}$$

Now apply Corollary 5.10. □

Application: Systems of equations.

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

$$\text{and } A = (a_1, \dots, a_n) = (Ae_1, \dots, Ae_n).$$

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

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

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

$$\begin{aligned} \text{By multilinearity, } \det A_k &= \sum_{j=1}^n x_j \det(a_1, \dots, a_{k-1}, \underbrace{a_j}_{(a_j)} a_{k+1}, \dots, a_n) \\ &= x_k \det(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n) \\ &= x_k \det A. \end{aligned}$$

$$\text{Since } A \text{ is invertible, } \det A \neq 0 \Rightarrow x_k = \frac{\det A_k}{\det A}$$

(16)

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

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

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

Theorem 5.12: If A is invertible, then its inverse A^{-1} has the

$$\text{form } (A^{-1})_{ki} = (-1)^{i+k} \frac{\det A_{ik}}{\det A}.$$

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

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

Note: This formula is computationally impractical for computing inverses.

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

Theorem 5.13:

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

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

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

Proof: (a) is obvious.

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

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

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

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

$$\begin{aligned} \det A &= \det (SBS^{-1}) = (\det S)(\det B)(\det S^{-1}) \\ &= (\det B)(\det S)(\det S)^{-1} = \det B. \quad \checkmark \end{aligned}$$

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

Since similar matrices represent the same linear maps 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.