MATH-GA2120 Linear Algebra II

Composition is Matrix Multiplication
Determinant of a Matrix
Determinant of a Linear Transformation

Deane Yang

Courant Institute of Mathematical Sciences
New York University

January 31, 2024

Composition is Matrix Multiplication

- Consider vector spaces U, V, W and linear maps

$$K : U \to V,\ L : V \to W$$

- Let $(e_1, \ldots, e_k)$ be a basis of U
- Let $(f_1, \ldots, f_m)$ be a basis of V
- Let $(g_1, \ldots, g_n)$ be a basis of W
- There is an m-by-k matrix M such that

$$K(e_j) = f_p M^p_j,\ 1 \le j \le k$$

- There is an n-by-m matrix N such that

$$L(f_p) = g_a N^a_p,\ 1 \le p \le m$$

- There is an n-by-k matrix P such that

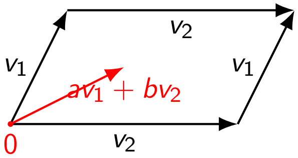
$$(L \circ K)(e_j) = g_a P^a_j,\ 1 \le j \le k$$

- On the other hand,

$$(L \circ K)(e_j) = L(K(e_j)) = L(f_p M^p_j) = L(f_p) M^p_j = g_a N^a_p M^p_j$$

- Therefore, $P^a_j = N^a_p M^p_j$.

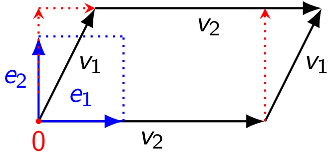
Parallelogram in Vector Space



- Let V be a 2-dimensional vector space
- Let $P(v_1, v_2)$ be the parallelogram with sides $v_1, v_2 \in V$.

$$P(v_1, v_2) = \{av_1 + bv_2 \ : \ 0 \leq a, b \leq 1\}.$$

Parallelogram With Respect To Basis



- With respect to basis (e_1, e_2)

$$v_1 = ae_1 + he_2 \text{ and } v_2 = we_1$$

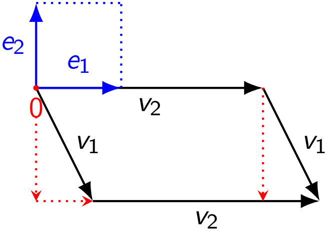
- Height is h and width is w
- Assume area of $P(e_1, e_2)$ is

$$A(e_1, e_2) = \text{area}(P(e_1, e_2)) = 1$$

- Then the area of $P(v_1, v_2)$ is

$$A(v_1, v_2) = \text{area}(P(v_1, v_2)) = hw$$

Upside Down Parallelogram With Respect To Basis



- With respect to basis (e_1, e_2)

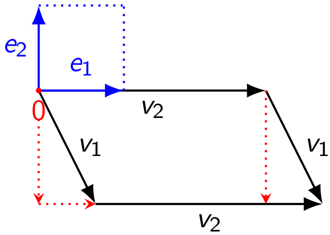
$$v_1 = ae_1 + he_2 \text{ and } v_2 = we_1,$$

  where h is negative
- Height is $|h|$ and width is w
- Then the area of $P(v_1, v_2)$ is

$$A(v_1, v_2) = |h|w,$$

  whether h is positive or negative
- Formula is awkward due to absolute value

Oriented Area of Parallelogram

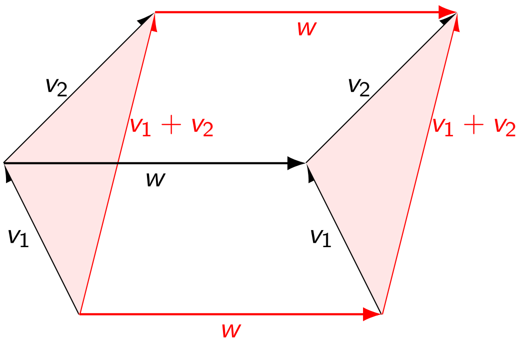


- Define oriented area to be

$$A(v_1, v_2) = hw$$

- The oriented area of $P(v_1, v_2)$ is positive if v_2 lies counterclockwise of v_1
- The oriented area of $P(v_1, v_2)$ is negative if v_2 lies counterclockwise of v_1
- Oriented area, as a function of $v_1, v_2 \in V$ has nice properties

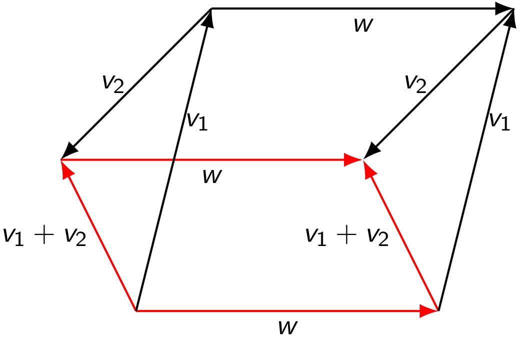
Area of Two Parallelograms with Parallel Bases



- If v_1 and v_2 both point upward relative to w, then

$$A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w)$$

Area of Two Parallelograms with Parallel Bases



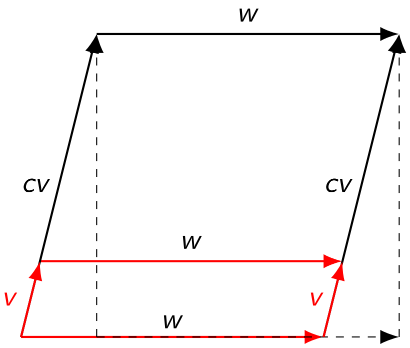
- If v_1 points upward and v_2 points downward relative to w, then $A(v_2, w) < 0$ and

$$A(v_1, w) = A(v_1 + v_2, w) - A(v_2, w)$$

and therefore

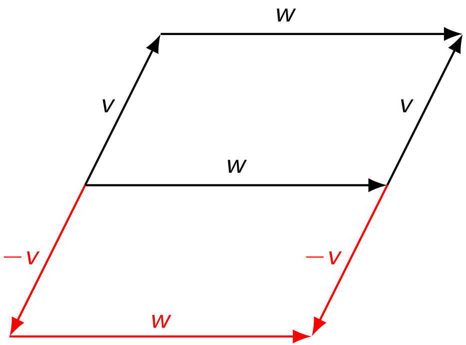
$$A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w)$$

Area of rescaled parallelogram



$$A(cv, w) = cA(v, w)$$

Area of reflected parallelogram



$$A(-v, w) = A(v, w)$$

Area Versus Oriented Area

- Definitions of area and oriented area require a basis (e_1, e_2), where we assume that

$$A(e_1, e_2) = 1$$

  - In particular, e_2 must lie counterclockwise of e_1
- The area function $|A(v, w)|$ is awkward to use
- Instead, define $A(v, w)$ to be the *oriented area* of $P(v, w)$
- Define the oriented area of $P(v, w)$ to be

$$A(v, w) = \begin{cases} \text{area of } P(v, w) & \text{if } (v, w) \text{ is positively oriented} \\ -\text{area of } P(v, w) & \text{if } (v, w) \text{ is negatively oriented} \\ 0 & \text{if } v \text{ and } w \text{ are linearly dependent} \end{cases}$$

Oriented Area of Parallelogram

- If w is held fixed, $A(v, w)$ is a linear function of v

$$A(v_1 + v_2, w) = A(v_1, w) + A(v_2, w)$$
$$A(cv, w) = cA(v, w)$$

- If v is held fixed, $A(v, w)$ is a linear function of w

$$A(v, w_1 + w_2) = A(v, w_1) + A(v, w_2)$$
$$A(v, cw) = cA(v, w)$$

- Such a function of two vectors is called **bilinear**
- For any $v \in V$, the parallelogram $A(v, v)$ has height 0 and therefore

$$A(v, v) = 0 \tag{1}$$

- Fact: Any bilinear function $A : V \times V \to \mathbb{F}$ that satisfies (1) is **antisymmmetric**
- This means that for any $v, w \in V$,

$$A(w, v) = -A(v, w)$$

2-Dimensional Antisymmetric Bilinear Function

- Let $\begin{bmatrix} e_1 e_2 \end{bmatrix}$ be a basis of V
- Let
$$A : V \times V \to \mathbb{F}$$
be an antisymmetric bilinear function such that
$$A(e_1, e_2) = 1$$
- If $v = ae_1 + be_2$ and $w = ce_1 + de_2$, then
$$\begin{aligned} A(v, w) &= A(ae_1 + be_2, ce_1 + de_2) \\ &= A(ae_1, ce_1) + A(be_2, ce_1) + A(ae_1, de_2) + A(be_2, de_2) \\ &= bcA(e_2, e_1) + adA(e_, e_2) \\ &= ad - bc \end{aligned}$$

2-Dimensional Antisymmetric Bilinear Function

- This can be written as follows

$$
\begin{aligned}
A\left(\left(\begin{bmatrix} v & w \end{bmatrix}\right)\right) &= A\left(\begin{bmatrix} e_1 & e_2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \\
&= A\left(\begin{bmatrix} ae_1 + be_2 & ce_1 + de_2 \end{bmatrix}\right) \\
&= A(e_1, e_2)(ad - bc) \\
&= ad - bc
\end{aligned}
$$

- The determinant of a square 2-by-2 matrix is defined to be

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Determinant of a 2-by-2 Matrix is Equal to Oriented Area

- Let (e_1, e_2) be a basis where the oriented area of $P(e_1, e_2)$ is 1,

$$A(e_1, e_2) = 1$$

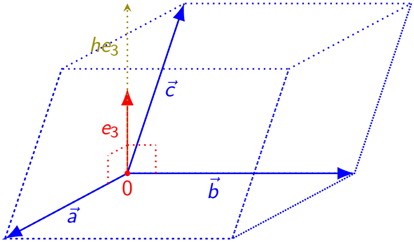
- The oriented area of the parallelogram $P(v, w)$, where

$$\begin{bmatrix} v & w \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

is

$$A(v, w) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

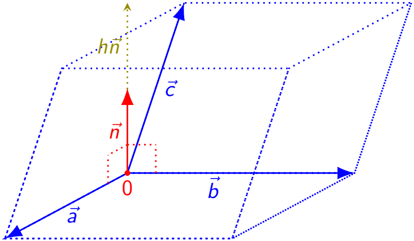
Parallelopiped spanned by 3 Vectors in 3-space



- Three linearly independent vectors $\vec{a}, \vec{b}, \vec{c}$ span a parallelopiped $P(\vec{a}, \vec{b}, \vec{c})$

$$P(\vec{a}, \vec{b}, \vec{c}) = \{s\vec{a} + t\vec{b} + u\vec{c} \ : \ 0 \leq s, t, u \leq 1\}$$

Volume of a Parallelopiped

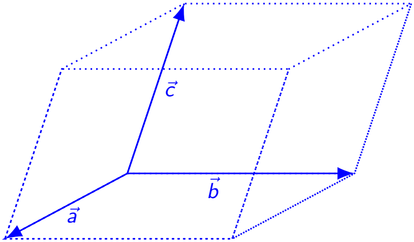


- Fix a basis (e_1, e_2, e_3) of V
  - Assume the volume of $P(e_1, e_2, e_2)$ is 1
- Assume $\vec{a}, \vec{b}$ lies in the subspace spanned by (e_1, e_2)
  - Therefore, $\vec{c} = he_3$
- If $h > 0$, then volume of parallelopiped is height times the area of the base:

$$\mathrm{vol}(P(\vec{a}, \vec{b}, \vec{c})) = h|A(\vec{a}, \vec{b})|$$

- Again, we want to avoid the absolute value

Oriented Volume of a Parallelopiped



- Define the oriented volume of $P\vec{a}, \vec{b}, \vec{c})$ to be

$$V(\vec{a}, \vec{b}, \vec{c}),$$

  where
  - $V(e_1, e_2, e_3) = 1$
  - $|V(\vec{a}, \vec{b}, \vec{c})|$ is the volume of $P(\vec{a}, \vec{b}, \vec{c})$
  - V is an antisymmetric multilinear function

Oriented Volume is the Determinant of a 3-by-3 Matrix

- Suppose $v_1, v_2, v_3 \in V$, where, using Einstein notation,

$$\begin{aligned} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} &= \begin{bmatrix} e_k A^k_1 & e_k A^k_2 & e_k A^k_3 \end{bmatrix} \\ &= \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} A^1_1 & A^1_2 & A^1_3 \\ A^2_1 & A^2_2 & A^2_3 \\ A^3_1 & A^3_2 & A^3_3 \end{bmatrix} \\ &= EA \end{aligned}$$

- The determinant of A is defined by the equation

$$V(v_1, v_2, v_3) = E \det A$$

- In particular, since $V(e_1, e_2, e_2) = 1$,

$$\det I = 1$$

Permutations

- A permutation is a bijective map $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$
- Let S_n be the set of all permutations of order n
- A transposition is a permutation τ that switches two elements and leaves the others unchanged.
  - Example: $\tau : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$, where

$$\tau(1) = 1,\ \tau(2) = 4,\ \tau(3) = 3,\ \tau(4) = 2$$

- Every permutation is a composition of transpositions
  - Example: The permutation $\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1$ can be written as

$$\sigma = \tau_1 \circ \tau_2, \text{ where}$$

$$\tau_1(1) = 2,\ \tau_2(2) = 1,\ \tau_2(3) = 3$$
$$\tau_2(1) = 1,\ \tau_1(2) = 3,\ \tau_1(3) = 2$$

Parity or Sign of a Permutation

- Given any permutation $\sigma \in S_n$, its parity or sign, which we will write as $\epsilon(\sigma)$, is defined to be
  - 1 if σ is the composition of an even number of transpositions
  - -1 if σ is the composition of an odd number of transpositions
- Easy consequences
  - If $\sigma \in S_n$ is a transposition, then $\epsilon(\sigma) = -1$
  - For any $\sigma, \tau \in S_n$, $\epsilon(\sigma \circ \tau) = \epsilon(\sigma)\epsilon(\tau)$
  - If σ is the identity map, then $\epsilon(\sigma) = 1$
  - For any $\sigma \in S_n$, $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$ because

$$\sigma = \tau_1 \circ \cdots \circ \tau_N \implies \sigma^{-1} = \tau_N \circ \cdots \circ \tau_1$$

Existence of Sign Function

- We have stated the properties that the sign function

$$\epsilon : S_n \to \{-1, 1\}$$

- Claim: There exists a unique function satisfying these properties
- This is the consequence of the following:
  - A permutation is never both the composition of an even number of transpositions and the composition of an odd number of transpositions
- There are straightforward elementary proofs of this
- There are also many sophisticated proofs

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

- Let $\mathrm{End}(n)$ denote the space of all maps

$$\phi : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

- Observe that $S_n \subset \mathrm{End}(n)$
- We can extend the function $\epsilon : S_n \to \{-1, 1\}$ to a function

$$\epsilon : \mathrm{End}(n) \to \{-1, 0, 1\},$$

where, if $\phi \in S_n$, then $\epsilon(\phi)$ is as defined before and

$$\epsilon(\phi) = 0 \text{ if } \phi \notin S_n$$

Alternating Multilinear Function on Permutation of Basis

- Suppose $T : V \times \cdots V \to \mathbb{F}$ is an alternating multilinear function with n inputs on an n-dimensional vector space
- Let $(e_1, \ldots, e_n)$ be a basis of V
- If $\phi \in S_n$ is a transposition, then

$$T(e_{\phi(1)}, \ldots, e_{\phi(n)}) = -T(e_1, \ldots, e_n)$$

- If $\phi_1, \phi_2 \in S_n$, then

$$\begin{aligned} &T(e_{\phi_2 \circ \phi_1(1)}, \ldots, e_{\phi_2 \circ \phi_1(n)}) \\ &\qquad = T(e_{\phi_2(1)}, \ldots, e_{\phi_2(n)}) T(e_{\phi_1(1)}, \ldots, e_{\phi_1(n)}) \end{aligned}$$

- Therefore, for any $\phi \in S_n$,

$$T(e_{\phi(1)}, \ldots, e_{\phi(n)}) = \epsilon(\phi) T(e_1, \ldots, e_n) = \epsilon(\phi) a$$

- If $\phi \in \mathrm{End}(n) \backslash S_n$, then it is not injective and therefore

$$T(e_{\phi(1)}, \ldots, e_{\phi(n)}) = 0 = \epsilon(\phi) T(e_1, \ldots, e_n)$$

Alternating Multilinear Function With Respect to Basis

- If $(e_1, e_2, \ldots, e_n)$ is a basis of V and $v_k = e_j a^j_k,\ 1 \le k \le n$,

$$
\begin{aligned}
T(v_1, \ldots, v_n) &= T(e_{j_1} a^{j_1}_1, \ldots, e_{j_n} a^{j_n}_n) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} T(e_{j_1}, \cdots, e_{j_n}) a^{j_1}_1 \cdots a^{j_n}_n \\
&= \sum_{\phi \in \mathrm{End}(n)} T(e_{\phi(1)}, \cdots, e_{\phi(n)}) a^{\phi(1)}_1 \cdots a^{\phi(n)}_n \\
&= \sum_{\phi \in \mathrm{End}(n)} \epsilon(\phi) T(e_1, \cdots, e_n) a^{\phi(1)}_1 \cdots a^{\phi(n)}_n \\
&= T(e_1, \ldots, e_n) \sum_{\phi \in \mathrm{End}(n)} \epsilon(\phi) a^1_{\phi(1)} \cdots a^n_{\phi(n)} \\
&= T(e_1, \ldots, e_n) \sum_{\phi \in S_n} \epsilon(\phi) a^1_{\phi(1)} \cdots a^n_{\phi(n)}
\end{aligned}
$$

Space of Alternating Multilinear Functions

- Let V be an n-dimensional vector space
- Let $\Lambda^n V^*$ denote the space of antisymmetric alternating functions with n inputs
- If $(e_1, \ldots, e_n)$ is a basis, then $T \in \Lambda^n V^*$, then T is uniquely determined by the value of

$$T(e_1, \ldots, e_n)$$

- If $T \in \Lambda^n V^*$ is nonzero, then for any $S \in \Lambda^n V^*$, there exists a constant $c \in \mathbb{F}$ such that

$$S = cT$$

- Specifically,

$$S = \left(\frac{S(e_1, \ldots, e_n)}{T(e_1, \ldots, e_n)} \right) T$$

- $\Lambda^n V^*$ is a 1-dimensional vector space

Determinant of an n-by-n Matrix

- Let $(e_1, \ldots, e_n)$ be the standard basis of $\mathbb{F}^n$
- There is a unique antisymmetric multilinear function

$$D : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$$

  with n inputs such that

$$D(e_1, \ldots, e_n) = 1,$$

- The formula for D is

$$D(v_1, \ldots, v_n) = \sum_{\phi \in \mathrm{End}(n)} b^1_{\phi(1)} \cdots b^n_{\phi(n)} = \sum_{\sigma \in S_n} b^1_{\sigma(1)} \cdots b^n_{\sigma(n)}$$

  where each $v_k = (b^1_k, \ldots, b^n_k) = e_j b^j_k$
- Let M be an n-by-n matrix with columns $C_1, \ldots, C_n$
- The determinant of M is defined to be

$$\det M = D(C_1, \ldots, C_n)$$

Determinant of an n-by-n Matrix

- The determinant function $det : M_{n\times n} \to \mathbb{F}$ is the unique function such that
  - $\det M$ is an alternating multilinear function of the columns of M
  - $\det I = 1$
- If
  $$M = \begin{bmatrix} M_1^1 & \cdots & M_n^1 \\ \vdots & & \vdots \\ M_1^n & \cdots & M_n^n \end{bmatrix},$$
  then
  $$\det M = \sum_{\phi\in \mathrm{End}(n)} \epsilon(\phi) M^1_{\phi(1)} \cdots M^n_{\phi(n)} = \sum_{\sigma\in S_n} \epsilon(\sigma) M^1_{\sigma(1)} \cdots M^n_{\sigma(n)}$$
- This formula can be useful in proofs
- To calculate the determinant of a specific matrix M, it is usually easier to use the properties of an alternating multilinear function

Multiplicative Property of the Determinant

A fundamental property of the determinant is that if A and B are n-by-n matrices, then

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

Proof of Multiplicative Property

- Given a matrix A, define the function $D_A : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ to be
$$D_A(v_1, \ldots, v_n) = \det AM,$$
where $(v_1, \ldots, v_n)$ are the columns of M
- If $C_1, \ldots, C_N$ are the columns of B, then the k-th column of AB is AC_k
- Therefore,
$$D_A(C_1, \ldots, C_n) = \det AB = D(AC_1, \ldots, AC_n)$$
- From this it is easy to see that D_A is an antisymmetric multilinear function of the columns of B
- Therefore,
$$\begin{aligned} D_A(C_1, \ldots, C_n) &= D_A(e_1, \ldots, e_n))D(C_1, \ldots, C_n) \\ &= (\det(AI))(D(B)) = (\det A)(\det B) \end{aligned}$$

Transpose of a Matrix

- Given a matrix $M \in \mathrm{gl}(n, m, \mathbb{F})$, its transpose is the matrix $M^T \in \mathrm{gl}(m, n, \mathbb{F})$ that switches the rows and columns
- In other words,
$$(M^T)^j_k = M^k_j$$
- Or
$$\begin{bmatrix} M^1_1 & \cdots & M^1_m \\ \vdots & & \vdots \\ M^n_1 & \cdots & M^n_m \end{bmatrix}^T = \begin{bmatrix} M^1_1 & \cdots & M^n_1 \\ \vdots & & \vdots \\ M^1_m & \cdots & M^n_m \end{bmatrix}$$
- If $M \in \mathcal{M}_{n\times m}$, then $M^T \in \mathrm{gl}(m, n, \mathbb{F})$
- For any $A \in \mathcal{M}_{k\times m}$ and $B \in \mathrm{gl}(m, n, \mathbb{F})$, then $AB \in \mathcal{M}_{k\times n}$ and
$$(AB)^T = B^T A^T \in \mathcal{M}_{n\times k}$$

Determinant of Matrix Equals Determinant of Its Tranpose

- Given any square matrix M,
$$\det M^T = \det M$$
- Proof 1: Use the formula for the determinant
$$\begin{aligned}\det M &= \sum_{\sigma \in S_n} \epsilon(\sigma) M_1^{\sigma(1)} \cdots M_n^{\sigma(n)} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) M^1_{\sigma^{-1}(1)} \cdots M^n_{\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) M^1_{\sigma(1)} \cdots M^n_{\sigma(n)} \\ &= \det M^T\end{aligned}$$
- Proof 2: Use the following facts:
  - Any matrix M can be written as $M = PLU$, where
    - P is a permutation matrix and $\det P = \det P^T$
    - L is a lower triangular matrix
    - U is an upper triangular matrix
  - Transpose of a triangular matrix is a triangular matrix with same determinant

$\det M \neq 0 \iff M$ is Invertible

- If the columns of M are linearly dependent, then one column can be written as a linear combination of the others, e.g.,

$$C_n = a^1 C_1 + \cdots + a^{n-1} C_{n-1}$$

- It follows that

$$\begin{aligned}
\det M &= D(C_1, \ldots, C_n) \\
&= D(C_1, \ldots, C_{n-1}, a^1 C_1 + \cdots + a^{n-1} C_{n-1} \\
&= a^k D(C_1, \ldots, C_{n-1}, C_k) \\
&= 0
\end{aligned}$$

- Recall that M is invertible
  - if and only if $\dim \ker M = 0$
  - if and only if the columns of M are linearly independent
- Conclusion:
  - M is invertible if and only if $\det M \neq 0$
  - Equivalently, M is singular if and only if $\det M = 0$

Determinant of a Linear Transformation

- Consider a linear transformation $L : V \to V$
- Given a basis $E = (e_1, \ldots, e_n)$ of V, let M be the matrix such that $L(E) = EM$
- Given another basis F, let N be the matrix such that $L(F) = FN$
- Observe that if $F = ET$, where T is an invertible matrix,

$$L(F) = L(ET) = L(E)T = EMT = ET(T^{-1}MT) = F(T^{-1}MT)$$

- It follows that $N = T^{-1}MT$
- Define the determinant of L to be the following: If E is a basis and $L(E) = EM$, then

$$\det(L) = \det(M)$$

- By the calculation above, this definition is independent of the basis used

Abstract Definition of $\det(L)$

- Let $L : V \to V$ be a linear map
- Let $D \in \Lambda^n V^*$
- Consider the function

$$\begin{aligned} D_L : V \times \cdots \times V &\to \mathbb{F} \\ (v_1, \ldots, v_n) &\mapsto D(L(v_1), \ldots, L(v_n)) \end{aligned}$$

- D_L is an antisymmetric multilinear function
- Therefore, there exists $c \in \mathbb{F}$ such that $D_l = cD$
- Define $\det(L) = c$