# MA-GY 7043: Linear Algebra II

## Areas of Parallelograms
## Volumes of Parallelopipeds
## Oriented Volume Functions
## Determinant of Linear Transformation

Deane Yang

Courant Institute of Mathematical Sciences
New York University

January 30, 2025

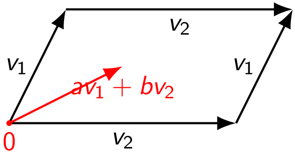
# Outline I



Determinant of Linear Map

# Parallelogram in Vector Space

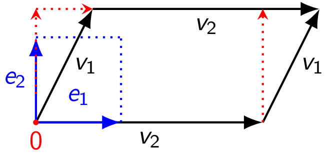





- Let $V$ be a 2-dimensional vector space
- Let $P(v_1, v_2)$ be the parallelogram with sides $v_1, v_2 \in V$.

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

# 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

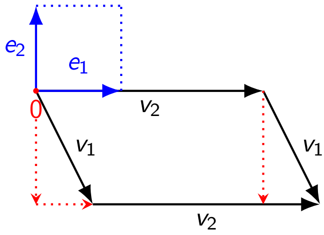
$$A(e_1, e_2) = \text{area}(P(e_1, e_2)) = 1$$

- 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 $(e_1, e_2)$

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

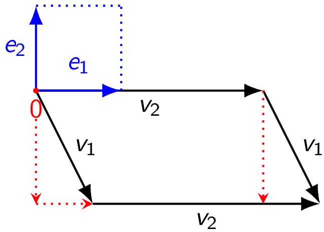
  where $h$ is negative
- Height is $|h|$ and width is $w$
- Then the area of $P(v_1, v_2)$ is

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

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

# Oriented Area of Parallelogram



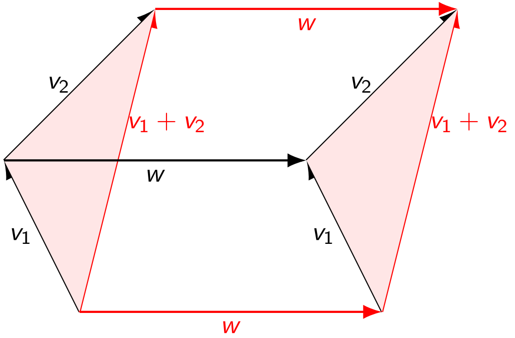



- Define oriented area to be

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

- 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

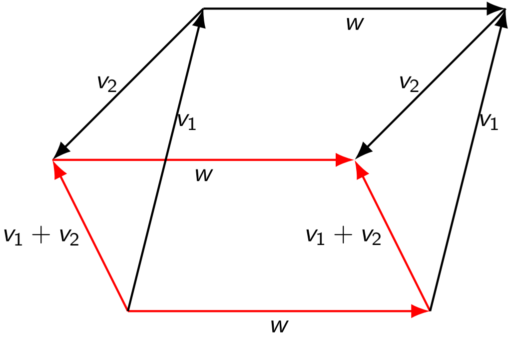
# Area of Two Parallelograms with Parallel Bases




- 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

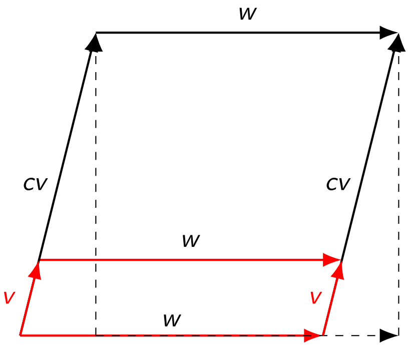
$$A(v_1, w) = A(v_1 + v_2, w) - A(v_2, w)$$

and therefore

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

# Area of rescaled parallelogram

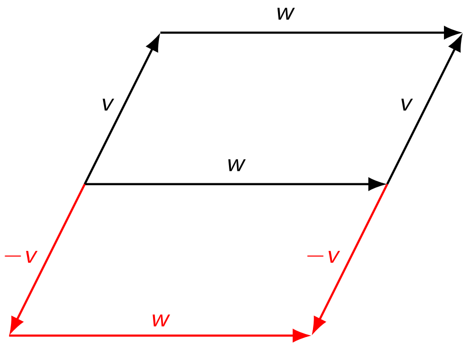





$$A(cv, w) = cA(v, w)$$

# 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

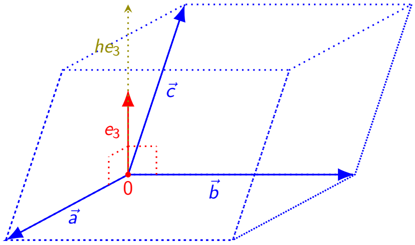
$$\begin{bmatrix} v & w \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

  is

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

# Parallelopiped spanned by 3 Vectors in 3-space

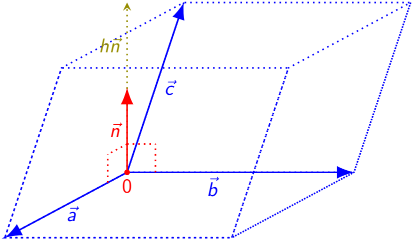





- Three linearly independent vectors $\vec{a}, \vec{b}, \vec{c}$ span a parallelopiped $P(\vec{a}, \vec{b}, \vec{c})$

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

# Volume of a Parallelopiped



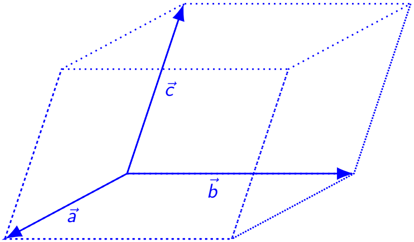

- Fix a basis $(e_1, e_2, e_3)$ of $V$
  - Assume the volume of $P(e_1, e_2, e_2)$ is 1
- Assume $\vec{a}, \vec{b}$ lies in the subspace spanned by $(e_1, e_2)$
  - Therefore, $\vec{c} = he_3$
- If $h > 0$, then volume of parallelopiped is height times the area of the base:

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

- 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,

$$\begin{aligned}\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} &= \begin{bmatrix} e_k A_1^k & e_k A_2^k & e_k A_3^k \end{bmatrix} \\ &= \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{bmatrix} \\ &= EA\end{aligned}$$

- 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 $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$
- Let $S_n$ be the set of all permutations of order $n$
- A transposition is a permutation $\tau$ 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 $\sigma$ is the composition of an even number of transpositions
  - $-1$ if $\sigma$ 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 $\sigma$ 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 Functions



- Let $V$ be an $n$-dimensional vector space
- Let
$$T : V \times \cdots \times V \to \mathbb{F}$$
be a function of $n$ vectors
- $T$ is **alternating** if for any $v_1, \ldots, v_n \in V$ and $\sigma \in S_n$,
$$T(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \epsilon(\sigma) T(v_1, \ldots, v_n)$$
- $T$ is **multilinear** if for each $1 \leq k \leq n$, $v_1, \ldots, v_n, w_k \in V$, $a, b \in \mathbb{F}$,
$$T(v_1, \ldots, av_k + bw_k, \ldots, v_n) \\ = aT(v_1, \ldots, v_k, \ldots, v_n) + bT(v_1, \ldots, w_k, \ldots, v_n)$$
- Let $\Lambda^n V^*$ denote the set of all alternating multilinear functions on $V$
- Each $T \in \Lambda^n V^*$ is an **oriented volume function** of $V$

# Alternating Multilinear Function on Permutation of Basis



- Let $T \in \Lambda^n V^*$
- 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 $\phi \in \mathrm{End}(n) \backslash S_n$, then it is not injective and therefore

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

- For any $v, v_3, \ldots, v_n$,

$$T(v, v, v_3, \ldots, v_n) = -T(v, v, v_3, \ldots, v_n) = 0$$

- If $v_1, \ldots, v_n$ are linearly dependent and

$$v_1 = a^2 v_2 + \cdots + a^n v_n,$$

then

$$\begin{aligned} T(v_1, \ldots, v_n) &= T(a^2 v_2 + \cdots + a^n v_n, v_2, \ldots, v_n) \\ &= a^2 T(v_2, v_2, \ldots, v_n) + \cdots + a^n T(v_n, v_2, \ldots, v_n) \\ &= 0 \end{aligned}$$

# Alternating Multilinear Function With Respect to Basis



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

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

# Space of Alternating Multilinear Functions



- If $S, T \in \Lambda^n V^*$ and $a, b \in \mathbb{F}$, then $aS + bT \in \Lambda^n V^*$
- Therefore, $\Lambda^n V^*$ is a vector space
- If $T \in \Lambda^n V^*$ is nonzero, then there exists a basis $(e_1, \dots, e_n)$ such that
$$T(e_1, \dots, e_n) \neq 0$$
- 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
$$S(e_1, \dots, e_n) = cT(e_1, \dots, e_n),$$
which implies that for any $v_1, \dots, v_n \in V$,
$$S(v_1, \dots, v_n) = cT(v_1, \dots, v_n)$$
- $\Lambda^n V^*$ is a 1-dimensional vector space
- Any $T \in \Lambda^n V^* \backslash \{0\}$ is a basis of $\Lambda^n V^*$

# Pullback of Oriented Volume Function by Linear Map



- Let $V$ and $W$ be $n$-dimensional vector spaces
- Let $L : V \to W$ be a linear map
- Let $T \in \Lambda^n W^*$ be an oriented volume function
- Use $L$ and $T$ to define $S \in \Lambda^n V^*$ as follows:

$$S(v_1, \ldots, v_n) = T(L(v_1), \ldots, L(v_n))$$

- $S$ is called the **pullback** of $T$ by the linear map $L$ and denoted $L^* T$
- $L^*(T) \neq 0$ if and only if $L$ is invertible
- The pullback is a linear map

$$L^* : \Lambda^n W^* \to \Lambda^n V^*$$

# Composition of Pullbacks



- Let
$$L_1 : V_0 \to V_1 \text{ and } L_2 : V_1 \to V_2$$
be linear maps
- Then for each $T \in \Lambda^n V_2^*$ and $v_1, \ldots, v_n \in V_0$,
$$\begin{aligned}(L_2 \circ L_1)^*(T)(v_1, \ldots, v_n) &= T(L_2 \circ L_1(v_1), \ldots, L_2 \circ L_1(v_n)) \\ &= T(L_2(L_1(v_1)), \ldots, L_2(L_1(v_n))) \\ &= L_2^*(T)(L_1(v_1), \ldots, L_1(v_n)) \\ &= L_1^*(L_2^*(T))(v_1, \ldots, v_n) \\ &= (L_1^* \circ L_2^*)(T)(v_1, \ldots, v_n)\end{aligned}$$
- Therefore, $(L_2 \circ L_1)^* = L_1^* \circ L_2^*$

# Determinant of a Linear Map



- Let $S \in \Lambda^n V^* \backslash \{0\}$ and $T \in \Lambda^n W^*$
- There exists a constant $c$ such that

$$L^*(T) = cS$$

- $c \neq 0$ if and only if $L$ is invertible
- $(L^{-1})^*(S) = c^{-1}S$
- $c$ depends on $L$, $S$, and $T$

# Determinant of Composition of Linear Maps



- Let $V_0, V_1, V_2$ be $n$-dimensional vector spaces
- Let
$$T_0 \in \Lambda^n V_0^*,\ \ T_1 \in \Lambda^n V_1^* \backslash \{0\},\ \ T_2 \in \Lambda^n V_2^*$$
- Let $L_1 : V_0 \to V_1$ and $L_2 : V_1 \to V_2$
- There exists a constants $c_1, c_2, c_3 \in \mathbb{F}$ such that
$$L_1^*(T_1) = c_1 T_0,\ \ L_2^*(T_2) = c_2 T_1,\ \ (L_2 \circ L_1)^*(T_2) = c_3 T_0$$
- Therefore,
$$\begin{aligned} c_3 T_0 &= (L_2 \circ L_1)^*(T_2) \\ &= (L_1^* \circ L_2^*)(T_2) \\ &= L_1^*(L_2^*(T_2)) \\ &= L_1^*(c_2 T_1) \\ &= c_2 L_1^*(T_1) \\ &= c_2 c_1 T_0 \end{aligned}$$
- It follows that $c_3 = c_1 c_2$

# Determinant of a Linear Transformation



- Let $T \in \Lambda^n V^*$
- Let $L : V \to V$ be a linear transformation
- Therefore, the pullback is a map
$$L^* : \Lambda^n V^* \to \Lambda^n V^*$$
- Since $\dim(\Lambda^n V^*) = 1$, for any $T \in \Lambda^n V^* \backslash \{0\}$, there exists $c \in \mathbb{F}$ such that
$$L^*(T) = cT$$
- Since $L^*$ is linear, if $S = aT$, then
$$L^*(S) = L^*(aT) = aL^*(T) = acT = cS$$
- The constant $c$ does not depend on $T \in \Lambda^n V^*$
- $c$ is called the **determinant** of $L$, denoted $\det(L)$, and defines a function
$$\det : \mathcal{L}(V) \to \mathbb{F}$$
- If $I : V \to V$ is the identity map, then $\det(I) = 1$

# Determinant of Composition of Linear Transformations



- Let $L_1 : V \to V$ and $L_2 : V \to V$ be linear transformations
- Given $T \in \Lambda^n V^*$,

$$\begin{aligned}(\det(L_2 \circ L_2))T &= (L_2 \circ L_1)^*(T) \\ &= L_1^*(L_2^*(T)) \\ &= (\det(L_1))L_2^*(T) \\ &= (\det(L_1))(\det(L_2))T\end{aligned}$$

- It follows that

$$\det(L_2 \circ L_1) = \det(L_1)\det(L_2)$$

- In particular, if $L : V \to V$ is invertible and $L^{-1}$ is the inverse map, then

$$1 = \det(I) = \det(L^{-1} \circ L) = \det(L^{-1})\det(L)$$

- Therefore, $\det(L) \neq 0$ and

$$\det(L^{-1}) = \frac{1}{\det(L)}$$

# Determinant of an $n$-by-$n$ Matrix



- Let $E = (e_1, \dots, e_n)$ be the standard basis of $\mathbb{F}^n$
- A matrix $M \in \mathrm{gl}(n, \mathbb{F})$ defines a linear map
$$L_M : \mathbb{F}^n \to \mathbb{F}^n,$$
where, for each $k \in \{1, \dots, n\}$,
$$L_M(e_k) = C_k$$
is the $k$-th column of $M$
- The determinant of the matrix $M$ is defined to be the determinant of the linear transformation $L_M$
- Recall that if $M_1, M_2 \in \mathrm{gl}(n, \mathbb{F})$ and $v = Ea \in \mathbb{F}^n$, then
$$\begin{aligned} L_{M_2} \circ L_{M_1}(Ea) &= L_{M_2}(L_{M_1}(Ea)) \\ &= L_{M_2}(EM_1 a) \\ &= EM_2 M_1 a \end{aligned}$$
- Therefore,
$$\det(M_2)\det(M_1) = \det(M_2 M_1)$$

# Formula for Determinant of Matrix



- There exists a unique $D \in \Lambda^n(\mathbb{F}^n)^*$ such that
$$D(e_1, \dots, e_n) = 1$$
- By the definition of $\det(M)$,
$$\begin{aligned} \det(M) &= \det(M) D(e_1, \dots, e_n) \\ &= \det(L_M) D(e_1, \dots, e_n) \\ &= D(L_M(e_1), \dots, L_M(e_n)) \\ &= D(C_1, \dots, C_n), \end{aligned}$$
where $C_1, \dots, C_n$ are the columns of $M \in \mathrm{gl}(n, \mathbb{F})$
- Since $C_k = e_j M^j_k$, it follows that
$$\begin{aligned} \det(M) &= D(C_1, \dots, C_n) \\ &= D(e_{j_1} M^{j_1}_1, \dots, e_{j_n} M^{j_n}_n) \\ &= D(e_1, \dots, e_n) \sum_{\phi \in S_n} \epsilon(\sigma) M^1_{\phi(1)} \cdots M^n_{\phi(n)} \\ &= D(e_1, \dots, e_n) \quad \sum \quad \epsilon(\sigma) M^1_{\phi(1)} \cdots M^n_{\phi(n)} \end{aligned}$$

# 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 Transpose



- Lemma: Given any square matrix $M$,

$$\det M^T = \det M$$

- Proof 1: Use the formula for the determinant

$$\begin{aligned} \det M &= \sum_{\sigma \in S_n} \epsilon(\sigma) M_1^{\sigma(1)} \cdots M_n^{\sigma(n)} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) M_{\sigma^{-1}(1)}^1 \cdots M_{\sigma^{-1}(n)}^n \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) M_{\sigma(1)}^1 \cdots M_{\sigma(n)}^n \\ &= \det M^T \end{aligned}$$

- Proof 2: Use the following facts to be proved later:
  - 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