VISPAD INSTITUTE OF TECHNOLOGY

Nea onnim no sua a ohu,

Matrices

$$\begin{bmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \end{bmatrix} \begin{bmatrix} b11 & b12 \\ b21 & b22 \\ b31 & b32 \end{bmatrix} = \begin{bmatrix} c11 & c12 \\ c21 & c22 \end{bmatrix}$$

```
c11 = a11 * b11 + a12 * b21 + a13 * b31
c12 = a11 * b12 + a12 * b22 + a13 * b32
c21 = a21 * b11 + a22 * b21 + a23 * b31
c22 = a21 * b12 + a22 * b22 + a23 * b32
```

Fall 2025 💻

(October – December Virtual Internship)

At the end of the session the students or candidates should be able to understand and work with:

- Matrices and Matrix Arithmetic
- Transposes and Symmetric Matrices
- Linear Systems and Vectors
- Image, Kernel, Rank, Nullity
- Superposition Principles for Linear Systems
- Matrix Inverses
- Linear transformation
    - Determinants

Matrix Inverses

- Many problems in machine learning involve solving systems of linear equations, often represented in the compact form $Ax = b$ where A is a matrix of coefficients, x is a vector of unknowns, and b is a vector of target values or outcomes.

- In linear algebra, solving equation: $5x = 10$ will involve multiplying both sides by the reciprocal of 5, which is $^1/_5$ or 5^{-1}.
- This gives $(5^{-1})5x = (5^{-1})10$
- This simplifies to: $1x = 2$
- The 5^{-1} is the multiplicative inverse of 5 because $5 \times 5^{-1} = 1$.

- The ***matrix inverse*** "undoes" the effect of ***matrix***.
- For a given square matrix A, its inverse, denoted as A^{-1}, is a matrix such when multiplied by A (in either order), the result is the identity matrix I.

$$AA^{-1} = A^{-1}A = I$$

- The identity matrix I acts like the number 1 in matrix multiplication

$$AI = IA = A$$

- The matrix inverse is defined only for square matrices.
- However, not all square matrices have an inverse.
- Matrices that do have an inverse are called invertible or non-singular.

- Matrices that do not have an inverse are called non-invertible or singular.

Properties of the Inverse

- **Inverse of the inverse:** The inverse of A^{-1} is A itself.

  $(A^{-1})^{-1} = A$

- **Inverse of a Product:** The inverse of a product of two invertible matrices is the product of their inverses in reverse order. $(AB)^{-1} = B^{-1}A^{-1}$

❑**Inverse of a Transpose:** The inverse of the transpose of a matrix is the transpose of its inverse.

$$(A^T)^{-1} = (A^{-1})^T$$

Matrix Mappings and Linear Mappings

- **Functions** are a fundamental concept in mathematics.
- A function f is a rule that assigns to every element x of an initial set called the **domain** of the function a unique value y in another set called the **codomain** of f.

- If f is a function with domain U and codomain V, then we say that f maps U to V and denote this by $f: U \to V$

Matrix Mappings

- Matrix-vector multiplication behaves like a function whose domain is $\mathbb{R}^n$ and whose codomain is $\mathbb{R}^m$.
- For any $m \times n$ matrix A and vector $\vec{x} \in \mathbb{R}^n$, the product $A\vec{x}$ is a vector in $\mathbb{R}^m$.

For any $A \in M_{m\times n}(\mathbb{R})$, we define a function $f_A: \mathbb{R}^n \to \mathbb{R}^m$ called matrix mapping corresponding to A by

$$f_A(\vec{x}) = A\vec{x}, \text{ for all } \vec{x} \in \mathbb{R}^n$$

Although a matrix mapping sends vectors to vectors, it is much more common to view functions as mappings points to points.

What is a linear transformation?

- In machine learning, our interest mainly lies in transforming the data.
- A neural network is just a function composed of smaller parts (known layers), transforming the data to a new feature space in every step.

- One of the key components of models in machine learning are linear transformations.
- *Linear transformation* and *linear mapping* mean exactly the same thing.

- Let U and V be two vector spaces (over the scalar field), and let $f: U \to V$ be a function between them.
- We say that f is linear if $f(a\boldsymbol{x} + b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$ holds for all vectors $\boldsymbol{x}, \boldsymbol{y} \in U$ and all scalars $\boldsymbol{a}, \boldsymbol{b}$

- A linear transformation is a mapping between two vector spaces that preserves the algebraic structure: **addition** and **scalar multiplication.**
- Functions between vector spaces are often called **transformations.**

1.Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 0 & 1 \end{bmatrix}$. Find $f_A(-1,4)$

2.Let $A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 0 & 2 & 6 \\ 3 & 2 & 1 & 7 \end{bmatrix}$. Find $f_A(-3,1,0,1)$

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and find the values of

$f_A(1,0), f_A(0,1)$, and $f_A(x_1, x_2)$

$$f_A(1, 0) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

$$f_A(0, 1) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

$$f_A(x_1, x_2) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

- Clearly, we can see a relationship between the image of the standard basis vectors in $\mathbb{R}^2$ and the image of any vector $\vec{x}$

Let $\overrightarrow{e_1}, \dots, \vec{e}_n$ be the standard basis vector of $\mathbb{R}^n$. If $A \in M_{m\times n}$ and $f_A\colon \mathbb{R}^n \to \mathbb{R}^m$ is the corresponding matrix mapping, then for any vector $\vec{x} \in \mathbb{R}^n$ we have

$$f_A(\vec{x}) = x_1 f_A(\vec{e}) + x_2 f_A(\vec{e}_2) + \cdots + x_n f_A(\vec{e}_n)$$

The images of the standard basis vectors are just the columns of A, we see that the image of any vector $\vec{x} \in \mathbb{R}^n$.

- Linearity is essentially combining two properties in one: $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) = f(\boldsymbol{y})$ and $f(a\boldsymbol{x}) = af(\boldsymbol{x})$ for all vectors $\boldsymbol{x}, \boldsymbol{y}$ and all scalars a.

$$f(a\boldsymbol{x}+b\boldsymbol{y}) = f(a\boldsymbol{x}) + f(b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$$

Properties of Linear Transformation

1. $f(0) = 0$ holds for every linear transformation.

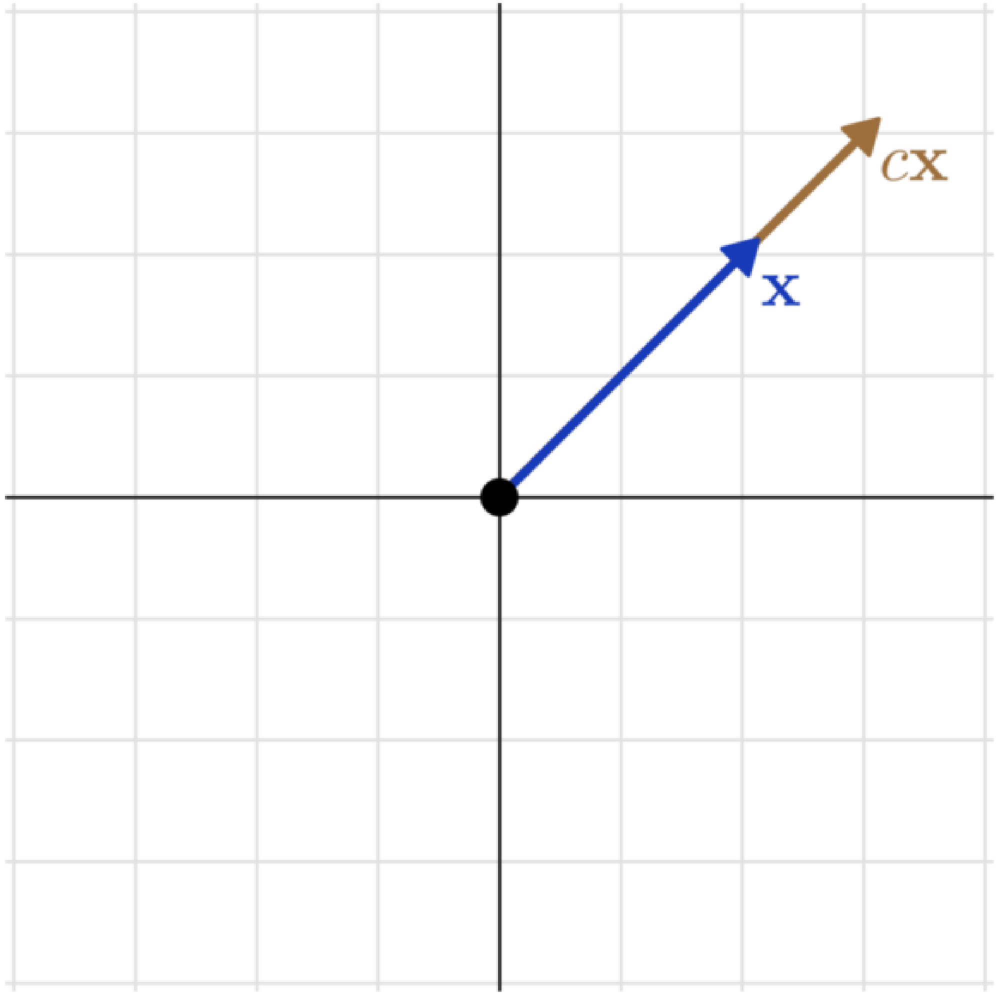
$$\begin{aligned} f(x) &= f(x+0) \\ &= f(x) + f(0) \end{aligned}$$

1.The composition of linear transformation is still linear.

$$
\begin{aligned}
f(g(ax+by) &= f(ag(x)+bg(y)) \\
&= af\big(g(x)\big) + bf\big(g(y)\big)
\end{aligned}
$$

Shows for any linear f and g and scalars a and b

For any scalar c, the scaling transformation $f(\boldsymbol{x}) = c\boldsymbol{x}$ is linear.



$$
\begin{aligned}
c(ax + by) &= c(ax) + c(by) \\
&= a(cx) + b(cy)
\end{aligned}
$$

Linear transformation and matrices

- Let $f: U \to V$ be a linear transformation between two vector spaces U and V.
- Suppose that $\{u_1, \dots, u_m\}$ is a basis in U, while $\{v_1, \dots, v_n\}$ is a basis in V.

- Every $\boldsymbol{x} \in U$ can be written in the form

$$x = \sum_{i=1}^{m} x_i \boldsymbol{u}_i \text{ the linearity of } f \text{ implies}$$

$$f\left(\sum_{j=1}^{m} x_j \boldsymbol{u}_j\right) = \sum_{j=1}^{m} x_j f\left(\boldsymbol{u}_j\right),$$

meaning that $f(x)$ is a linear combination of $f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_m)$.

- Every linear transformation is completely determined by the images of basis vectors.

- Suppose that for every u_j, we have

$$f(u_j) = \sum_{i=1}^{n} a_{i,j} v_i$$

for some scalars $a_{i,j}$

- These $n \times m$ numbers completely describe f

$$f \leftrightarrow A_f = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{bmatrix},$$

- Linear transformations are represented by matrices.
- For every $x = \sum_{j=1}^{m} x_j \boldsymbol{u}_j$, we have

$$
\begin{aligned}
f(\mathbf{x}) &= \sum_{j=1}^{m} x_j f(\mathbf{u}_j) \\
&= \sum_{j=1}^{m} x_j \sum_{i=1}^{n} a_{i,j} \mathbf{v}_i \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{i,j} x_j \right) \mathbf{v}_i .
\end{aligned}
$$

The image of $\boldsymbol{x}$ can be expressed as $A_f\boldsymbol{x}$:

$$f(\mathbf{x}) = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m}, \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^m a_{1,j}x_j \\ \sum_{j=1}^m a_{2,j}x_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}x_j \end{bmatrix}.$$

- Functions can be **added** and **composed**.
- Because of the connection between linear transformations and matrices, matrix operations are inherited from the corresponding function operations.

- Matrix addition can be defined so that the matrix of the sum of two linear transformation is the sum of the corresponding matrices.
- If $f, g\colon U \to V$ are two linear tranformations with matrices, $f \leftrightarrow A$ and $g \leftrightarrow B$

$$(f+g)(\mathbf{u}_j) = f(\mathbf{u}_j) + g(\mathbf{u}_j) = \sum_{i=1}^{n} (a_{i,j} + b_{i,j})\mathbf{v}_i.$$

- The corresponding matrices can be added together elementwise: $A + B = (a_{i,j} + b_{i,j})_{i,j=1}^{n,m}$.

- Multiplication between matrices is defined by the composition of the corresponding transformation.
- Let $f, g: U \to U$ be two linear transformations, mapping U onto itself.

- $f \circ g$, can be expressed as $f\left(g(u_j)\right)$ in terms of all of the basis vectors $u_1, \dots, u_n$

$$
\begin{aligned}
(fg)(\mathbf{u}_j) = f(g(\mathbf{u}_j)) &= f\Big(\sum_{k=1}^{n} b_{k,j}\mathbf{u}_k\Big) \\
&= \sum_{k=1}^{n} b_{k,j} f(\mathbf{u}_k) \\
&= \sum_{k=1}^{n} b_{k,j} \sum_{i=1}^{n} a_{i,k}\mathbf{u}_i \\
&= \sum_{i=1}^{n} \Big(\sum_{k=1}^{n} a_{i,k} b_{k,j}\Big)\mathbf{u}_i.
\end{aligned}
$$

$$
AB = \left(\sum_{k=1}^{n} a_{i,k} b_{k,j}\right)_{i,j=1}^{n}.
$$

Inverting linear transformations

$$2x_1 + x_2 = 5$$
$$x_1 - 3x_2 = -8$$

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ -8 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

- ❑The above system can be written in the form $Ax = \boldsymbol{b}$
- ❑This is called linear equations.
- ❑How do we write the solution of such equation?

- If there would be a matrix A^{-1} such that $A^{-1}A$ is the identity matrix I
- Multiplying the equation $A\boldsymbol{x} = \boldsymbol{b}$ from the left by A^{-1} would yield the solution in the form $\boldsymbol{x} = A^{-1}\boldsymbol{b}$
- The matrix A^{-1} is called the ***inverse matrix*** of A.
- It might not always exist, but when it does, it is extremely important for several reasons.

Let $f: U \to V$ be a linear transformation between the vector spaces U and V.

We say that f is invertible if there is a linear transformation f^{-1} such that $f^{-1} \circ f$ and $f \circ f^{-1}$ are the identity function

$$f^{-1}\big(f(\boldsymbol{u})\big) = \boldsymbol{u},$$

$$f\big(f^{-1}(\boldsymbol{v})\big) = \boldsymbol{v}$$

holds for all $u \in U, v \in V$. f^{-1} is called the inverse of f

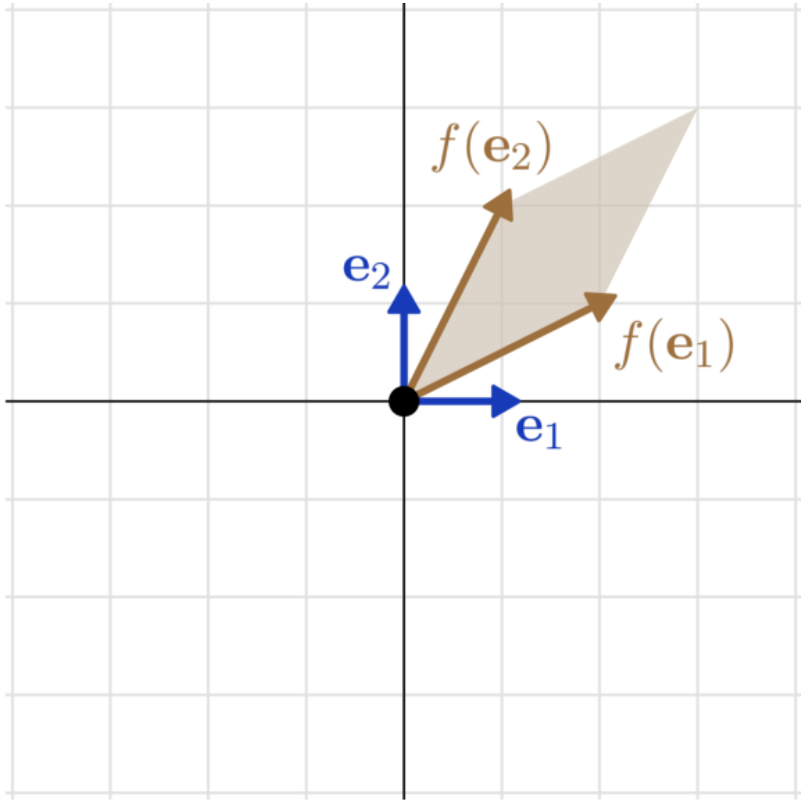
- Not all linear transformations are invertible.
- If f maps all vectors to the zero vector, you cannot define inverse.

Let $f: U \to V$ be a linear transformation and let $u_1, \dots, u_n$ be a basis in U. Then f is invertible if and only if $f(u_1), \dots, f(u_n)$ is a basis in V.

Change of basis

- ❑Any linear transformation can be described with the images of the basis vector.
- ❑This gives us the matrix representation that we use all the time.
- ❑Different bases yield different matrices for the same transformation.

- Let's take a look at $f: \mathbb{R}^2 \to \mathbb{R}^2$, which maps $e_1 = (1, 0)$ to the vector $(2, 1)$ and $e_2 = (0, 1)$ to $(1, 2)$
- Its matrix in the standard orthonormal basis $E = \{e_1, e_2\}$ is given by $A_{f,E} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$



- Say $P = \{p_1 = (1,1), p_2 = (-1, 1)\}$

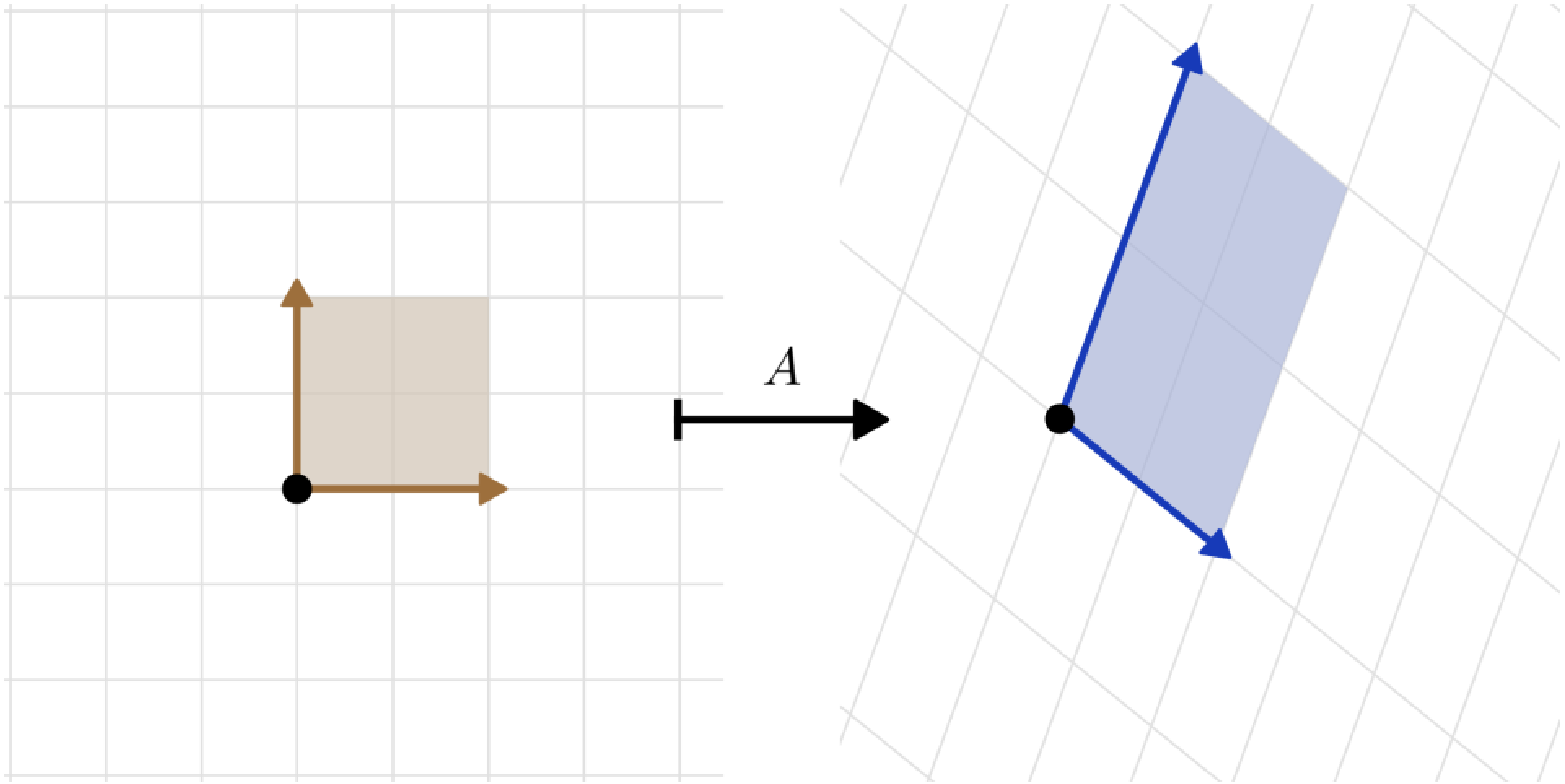
$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

In other words, $f(p_1) = 3p_1 + 0p_2$ and $f(p_2) = 0p_1 + p_2$

$$A_{f,P} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Linear transformations in the Euclidean plane

- ❑ Linear transformation can be described by the image of a basis set.
- ❑ From a geometric viewpoint, they are functions mapping parallelepipeds to parallelepipeds.
- ❑ Because of the linearity, you can imagine this as distorting the grid determined by the bases.



A

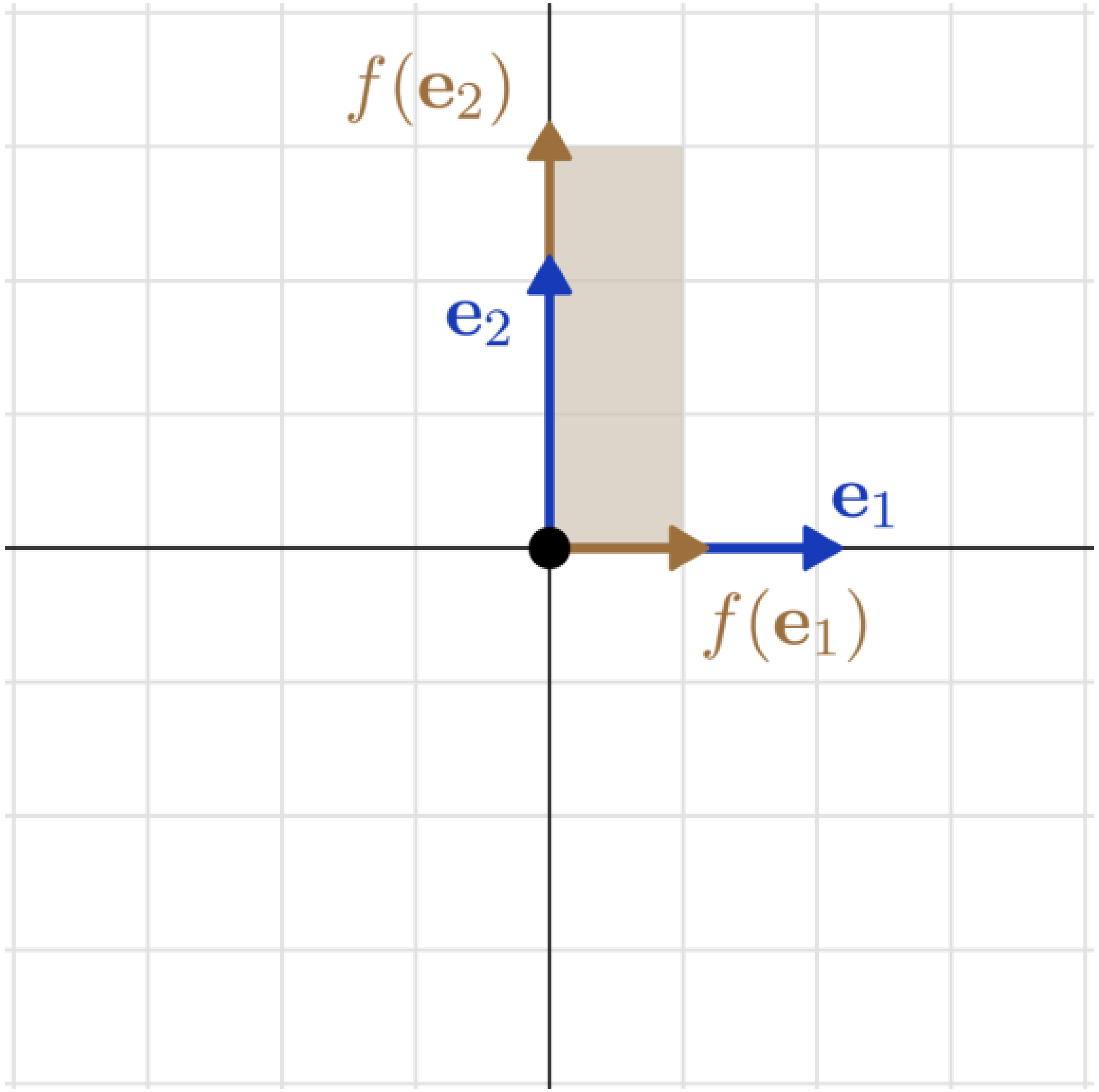
Geometric maps includes:

- ❑ Stretching
- ❑ Shearing
- ❑ Rotation
- ❑ Reflection
- ❑ Projection

Stretching

$$A = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

- Linear transformation such as this can be visualized by plotting the image of the unit square determined by the standard basis $\boldsymbol{e}_1 = (1, 0), \boldsymbol{e}_2 = (0, 1)$



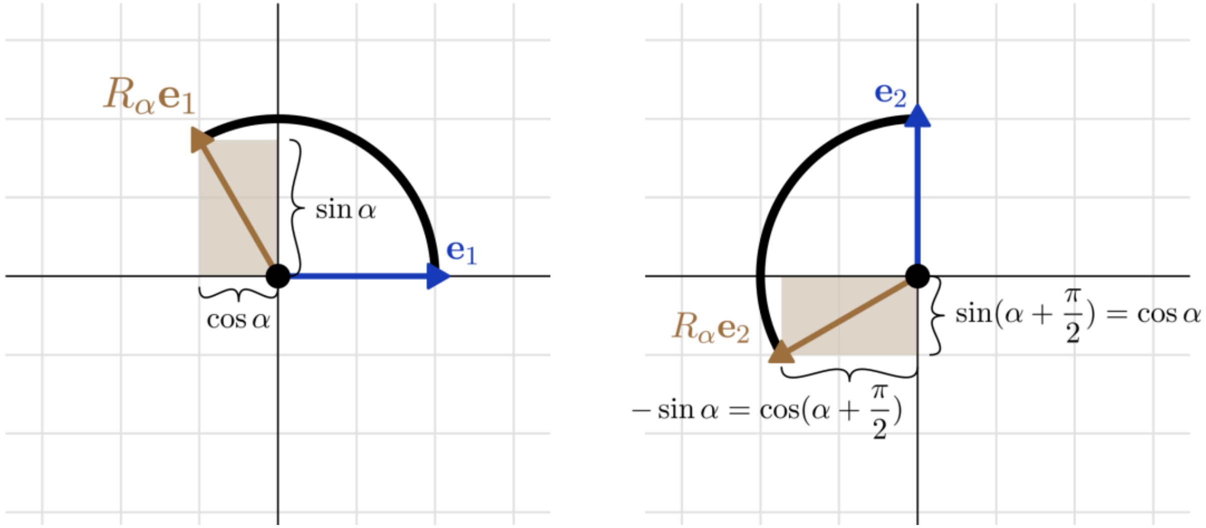
$f(\mathbf{e}_2)$
$\mathbf{e}_2$
$\mathbf{e}_1$
$f(\mathbf{e}_1)$

Rotations

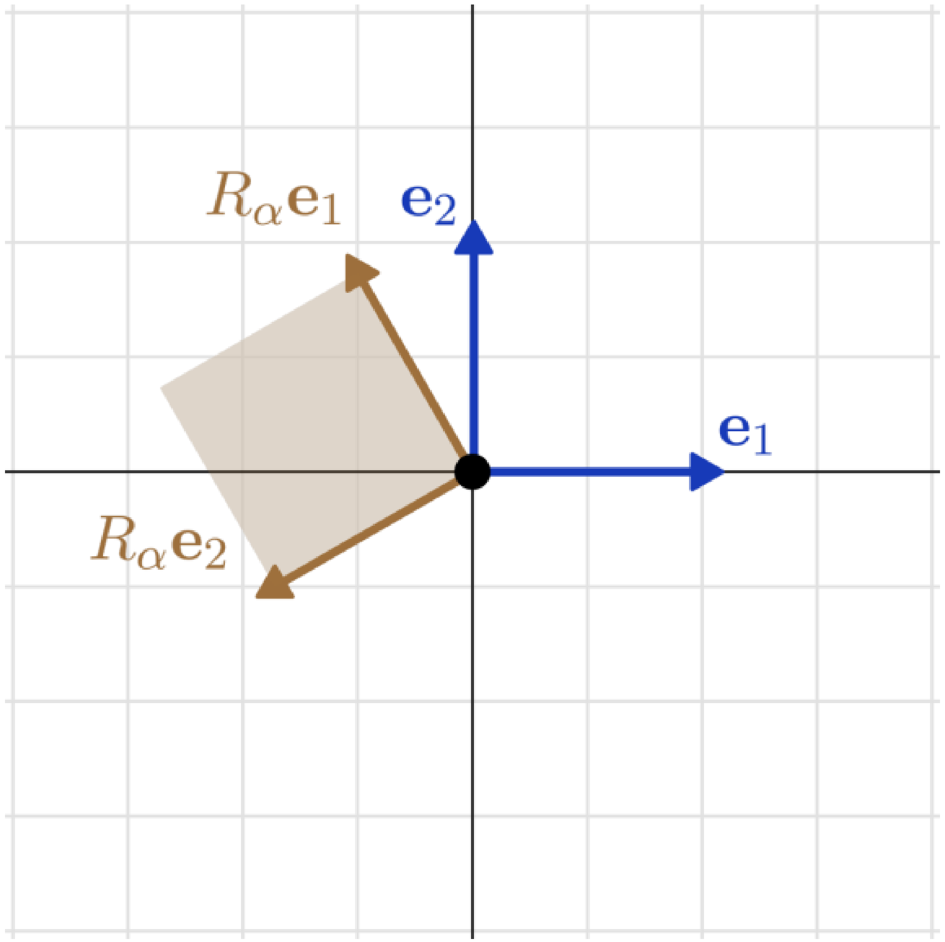
$$R_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

- The rotation of $(1,0)$ is given by $(\cos\alpha, \sin\alpha)$, while

  the rotation of $(0, 1)$ is $\left(\cos\left(\alpha + \frac{\pi}{2}\right), \sin\left(\alpha + \frac{\pi}{2}\right)\right)$



$R_\alpha \mathbf{e}_1$
$\sin\alpha$
$\mathbf{e}_1$
$\cos\alpha$
$\mathbf{e}_2$
$R_\alpha \mathbf{e}_2$
$\sin(\alpha + \frac{\pi}{2}) = \cos\alpha$
$-\sin\alpha = \cos(\alpha + \frac{\pi}{2})$

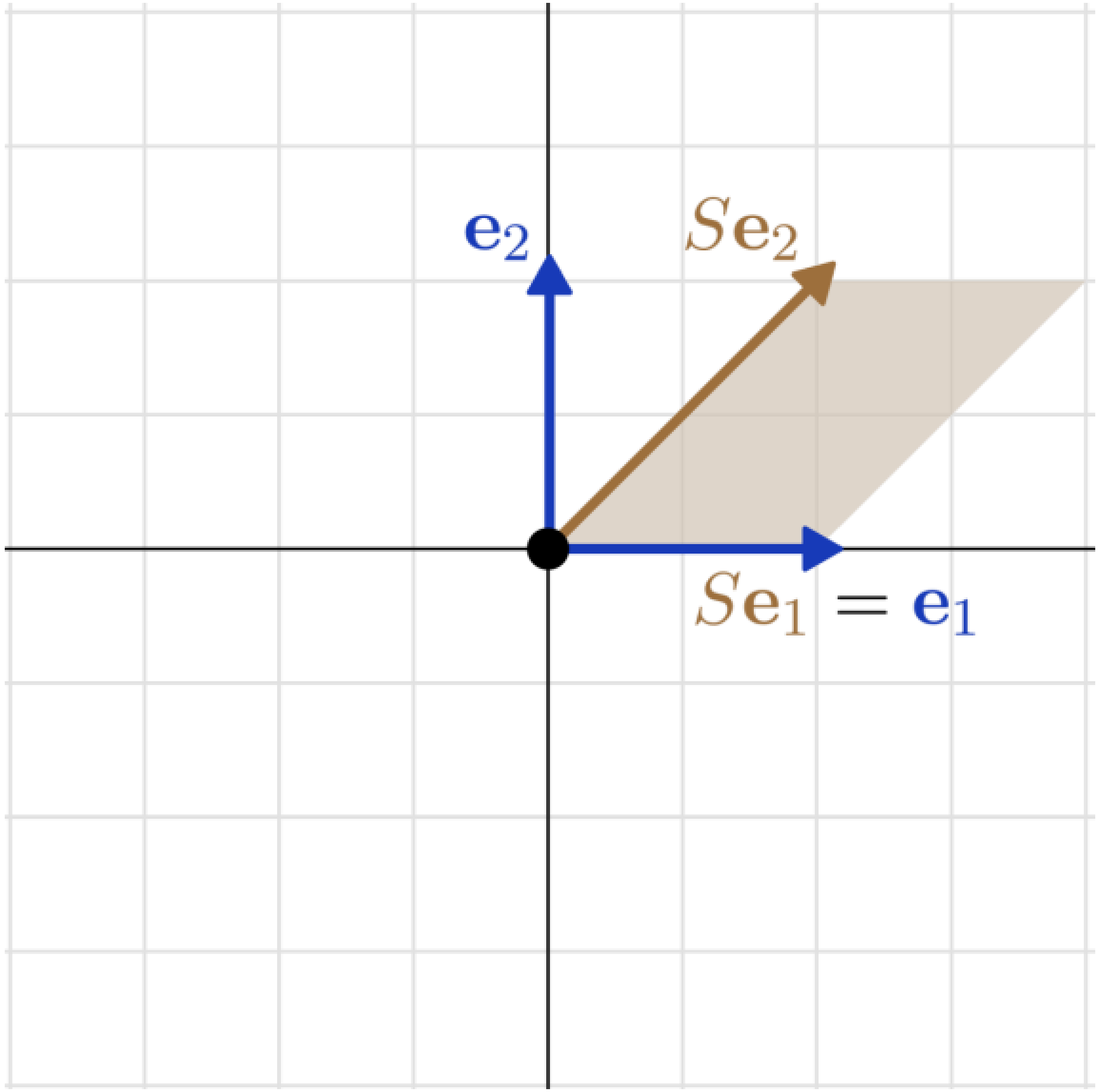


$R_\alpha \mathbf{e}_1$
$\mathbf{e}_2$
$\mathbf{e}_1$
$R_\alpha \mathbf{e}_2$

Shearing

$$\mathrm{S}_x = \begin{bmatrix} 1 & k_x \\ 0 & 1 \end{bmatrix}, \quad \mathrm{S}_y = \begin{bmatrix} 1 & 0 \\ k_y & 1 \end{bmatrix}, \quad \mathrm{S} = \begin{bmatrix} 1 & k_x \\ k_y & 1 \end{bmatrix},$$

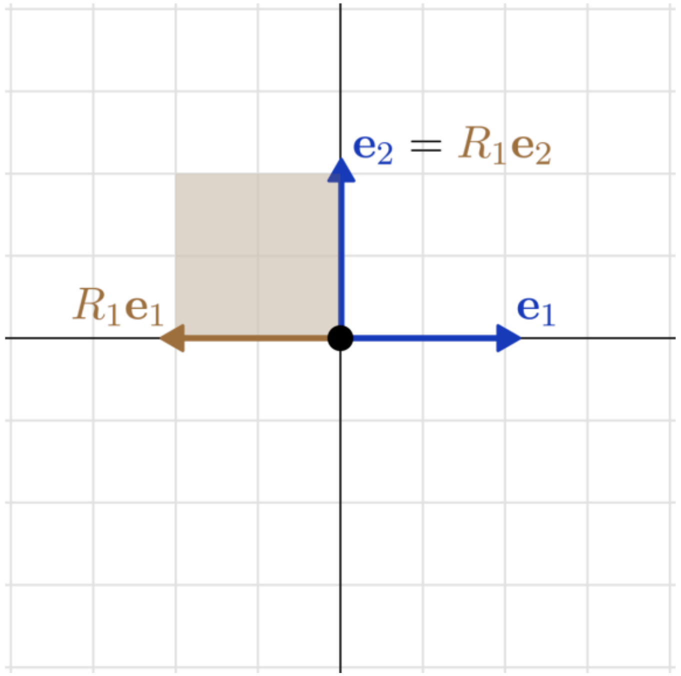
where S_x, S_y and S represent shearing transformation in the $x, y,$ and in both directions.


$\mathbf{e}_2$
$S\mathbf{e}_2$
$S\mathbf{e}_1 = \mathbf{e}_1$

Reflection

$$R_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

acts as reflections with respect to the x and the y axes.



- In general, reflections can be easily defined in higher dimensional spaces.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

- This is reflection in $\mathbb{R}^3$ that flips $\boldsymbol{e}_3$ to the opposite direction.

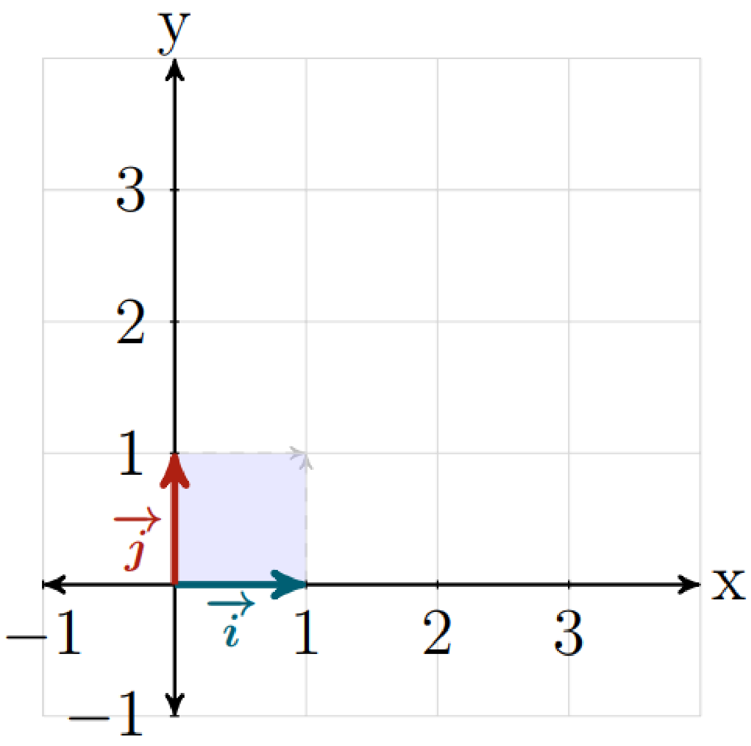
- Reflections can flip orientations multiple times. The transformation given by

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

flips $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, changing the orientation twice.

Determinants

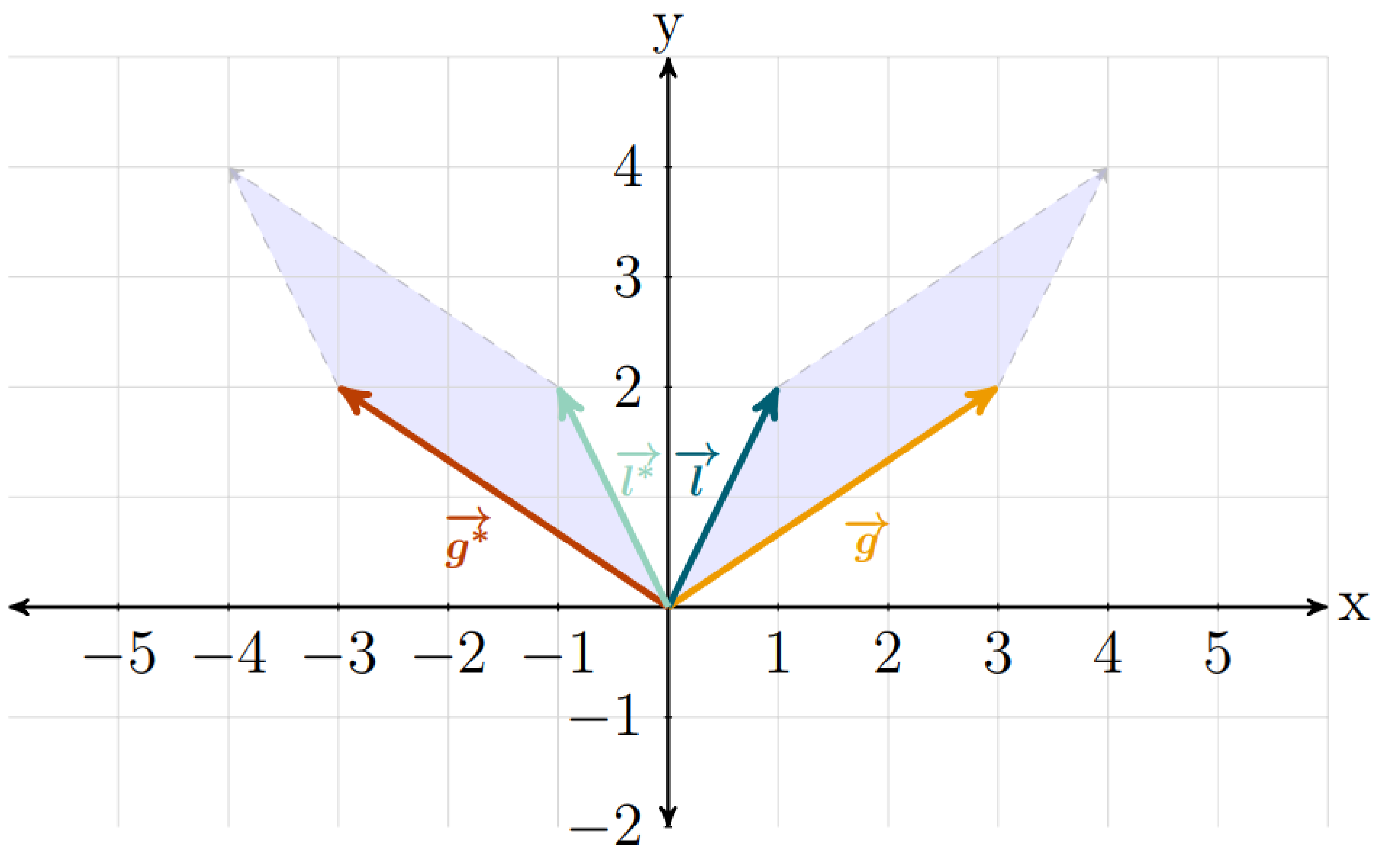
- ❑Let's use the standard basis $\vec{\imath} = (1,0)^T$ and $\vec{\jmath} = (0, 1)^T$ and form a square.



- A linear transformation can be represented by the matrix: $L = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.
- Applying the mapping to L to $\vec{i}$ and $\vec{j}$ will result into two new vectors, $\overrightarrow{i^*}$ and $\overrightarrow{j^*}$, with values $(2, 0)^T$ and $(0, 2)^T$ respectively.

- The scalar that represents the change in area size due to the transformation is called the **determinants.**
- The **determinants** is a function that receives a square matrix and outputs a scalar that does not need to be strictly positive.
- It can take negative values or even zero.

- Consider the transformation $z = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$
- Let's consider two new vectors: $\vec{g} = \binom{3}{2}, \vec{l} = \binom{1}{2}$



- The image on the previous slide represent a determinant with a negative value.
- The negative sign indicates that the orientation of the vectors has changed.

- Consider a generic matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

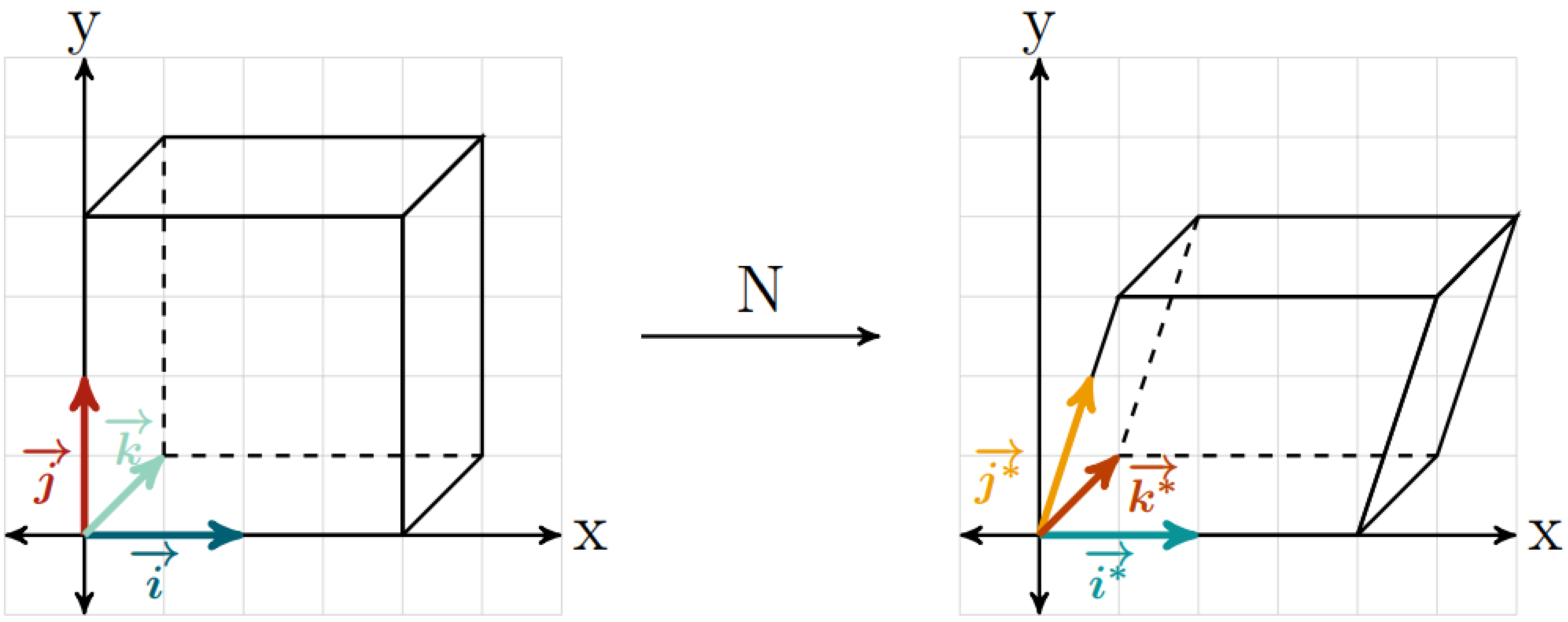
- The notation for the determinant is $\det(M)$ or $|M|$, and the formula for its computation is:

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

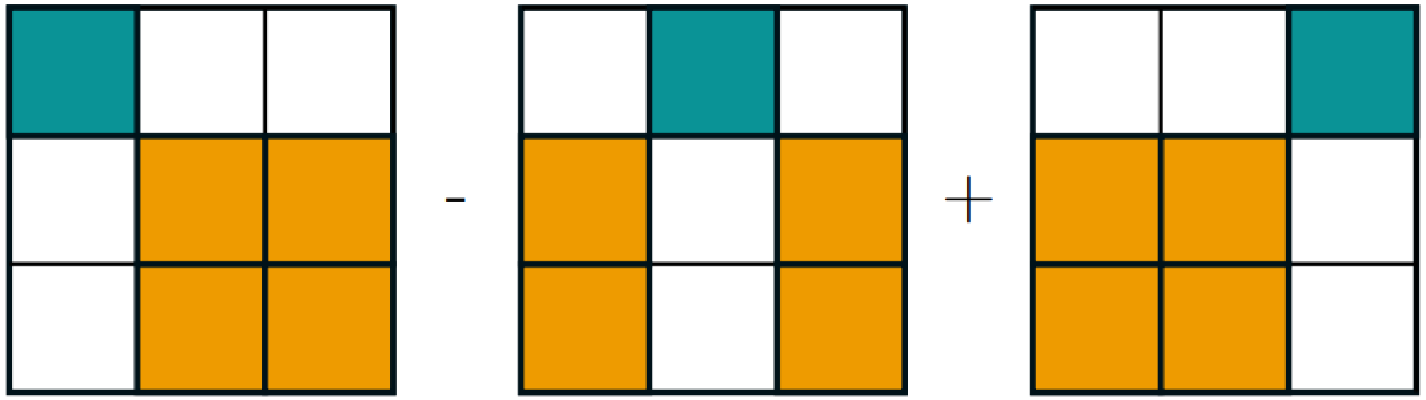
- Consider a generic 3×3 matrix, N such that:

$$N = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

- The determinant will reflect the change in volume of a transformed parallelepiped.



y
x
$\overrightarrow{j}$
$\overrightarrow{k}$
$\overrightarrow{i}$
N
y
x
$\overrightarrow{j^*}$
$\overrightarrow{k^*}$
$\overrightarrow{i^*}$



$$\det(N) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Fundamental properties of the determinants

1.$\in \mathbb{R}^{n\times n}$ be two matrices. Then $\det AB = detAdetB$

2.Let $A \in \mathbb{R}^{n\times n}$ be an arbitrary invertible matrix. Then $detA^{-1} = (detA)^{-1}$

3.Let $A \in \mathbb{R}^{n\times n}$ be an arbitrary matrix. Then $detA = detA^T$

4. Let $A \in \mathbb{R}^{n \times n}$ be a matrix that has two identical rows or columns. Then $\det A = 0$
5. Let $A \in \mathbb{R}^{n \times n}$ be a matrix. Then it columns are linearly dependent if and only if $\det A = 0$. Similarly, the rows of A are linearly dependent if and only if $\det A = 0$

6. Let $A \in \mathbb{R}^{n \times n}$ be a matrix with a constant zero column (or row). Then $\det A = 0$
7. The linear transformation $A \in \mathbb{R}^{n \times n}$ is invertible if and only if $\det A \neq 0$