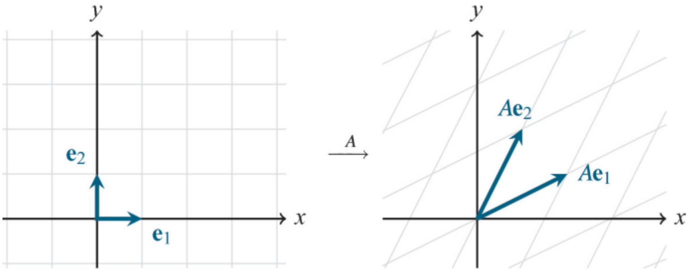
Inverse, Determinant

Hayk Aprikyan, Hayk Tarkhanyan

Recap:
When you multiply, say, a 2×2 matrix A by a vector $\mathbf{v} \in \mathbb{R}^2$, what you get is another vector $\mathbf{u} = A\mathbf{v} \in \mathbb{R}^2$. We call this $\mathbf{u}$ the **transformed version** of $\mathbf{v}$ (and we say that A is a linear transformation).



Geometric Interpretation

As we will see later, the resulting "transformed version" **u** is just the same old **v** except it is **rotated** and **scaled** to become longer or shorter (and possibly, flipped).

Geometric Interpretation

As we will see later, the resulting "transformed version" **u** is just the same old **v** except it is **rotated** and **scaled** to become longer or shorter (and possibly, flipped).

In this sense, all matrices are either just rotating vectors by some degree, or flipping them horizontally/vertically, or scale them, or do all three.

The key thing is: whatever a matrix "does" to one vector, it does the same to all other vectors too (when being multiplied with them).

Check different matrices yourself:

- `visualize-it.github.io/linear_transformations/simulation.html`
- `www.shad.io/MatVis`

Geometric Interpretation

As we will see later, the resulting "transformed version" **u** is just the same old **v** except it is **rotated** and **scaled** to become longer or shorter (and possibly, flipped).

In this sense, all matrices are either just rotating vectors by some degree, or flipping them horizontally/vertically, or scale them, or do all three.

The key thing is: whatever a matrix "does" to one vector, it does the same to all other vectors too (when being multiplied with them).

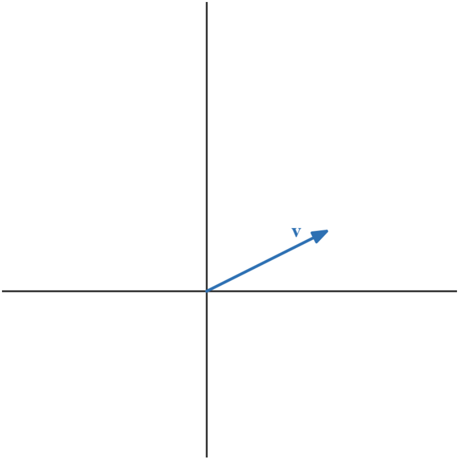
Check different matrices yourself:

- `visualize-it.github.io/linear_transformations/simulation.html`
- `www.shad.io/MatVis`

Now that we know what matrix $\times$ vector multiplication means, what about matrix $\times$ matrix multiplication? Why is it defined the way it is?

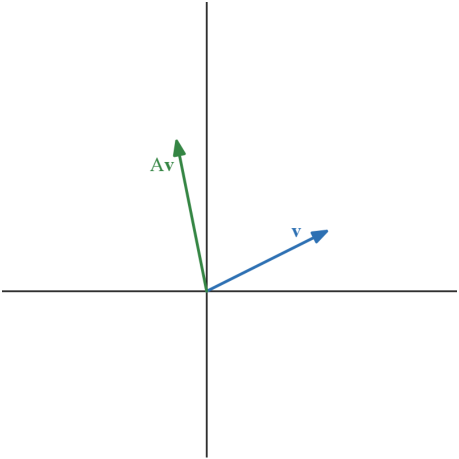
Geometric Interpretation

Suppose $\mathbf{v} \in \mathbb{R}^2$, $A \in \mathbb{R}^{2\times 2}$, $B \in \mathbb{R}^{2\times 2}$:



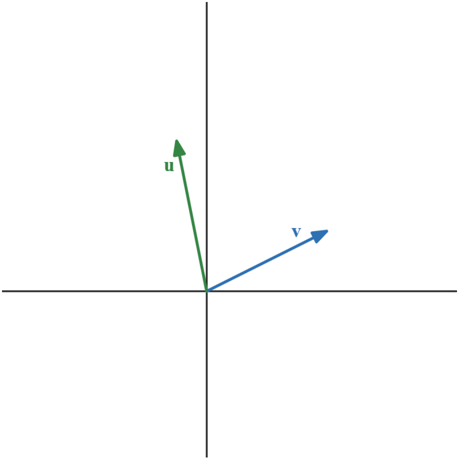
Geometric Interpretation

If we apply A on $\mathbf{v}$, we get a transformed version of $\mathbf{v}$,



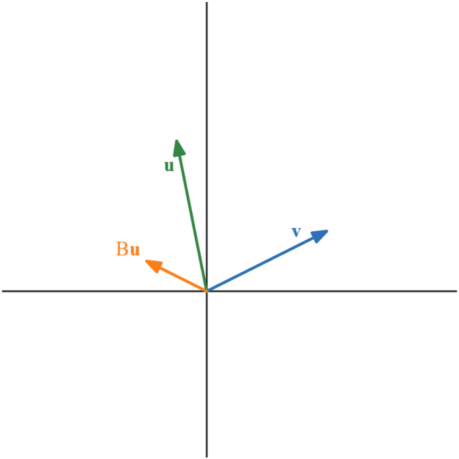
Geometric Interpretation

If we apply A on $\mathbf{v}$, we get a transformed version of $\mathbf{v}$, say $\mathbf{u}$:



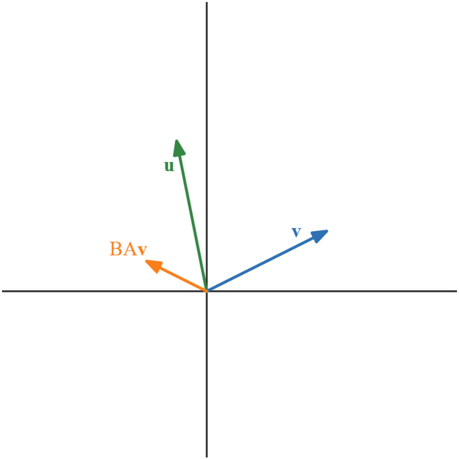
Geometric Interpretation

Now applying B on $\mathbf{u}$, we get a transformed version of $\mathbf{u}$, i.e. $B\mathbf{u}$

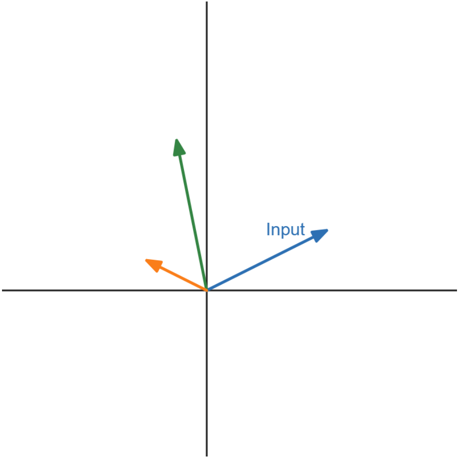


Geometric Interpretation

Now applying B on $\mathbf{u}$, we get a transformed version of $\mathbf{u}$, i.e. $B\mathbf{u} = BA\mathbf{v}$

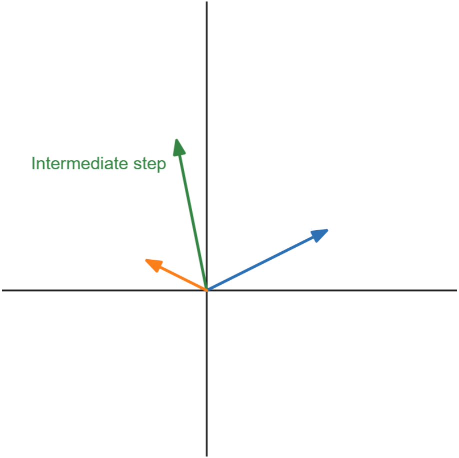


So what is the product BA? To get $(BA)(\mathbf{v})$, we do:



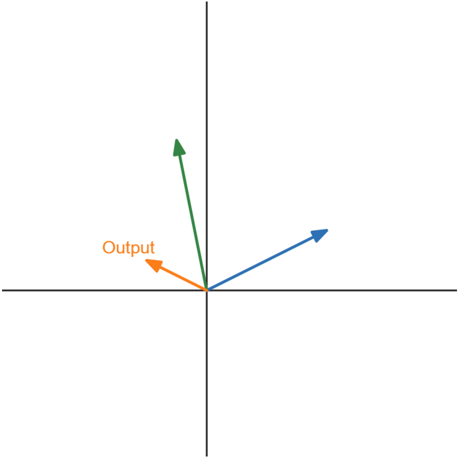
Geometric Interpretation

So what is the product BA? To get $(BA)(\mathbf{v})$, we do:



Geometric Interpretation

So what is the product BA? To get $(BA)(\mathbf{v})$, we do:



Geometric Interpretation

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!

Geometric Interpretation

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!
(Note that this means that often AB is *not the same* as BA)

Geometric Interpretation

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!
(Note that this means that often AB is *not the same* as BA)

Question

Suppose A is the matrix that rotates the vectors by 30°, B the one that rotates by 50°, and C by 260°.

Geometric Interpretation

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!
(Note that this means that often AB is *not the same* as BA)

Question

Suppose A is the matrix that rotates the vectors by 30°, B the one that rotates by 50°, and C by 260°.
What would the product matrix BA be?

Geometric Interpretation

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!
(Note that this means that often AB is *not the same* as BA)

Question

Suppose A is the matrix that rotates the vectors by 30°, B the one that rotates by 50°, and C by 260°.
What would the product matrix BA be? What about CBA?

Hence BA is the linear transformation that we get by **first applying** A, and **then applying** B!
(Note that this means that often AB is *not the same* as BA)

Question

Suppose A is the matrix that rotates the vectors by 30°, B the one that rotates by 50°, and C by 260°.
What would the product matrix BA be? What about CBA?

Which matrix leaves everything in its place (does not touch anything)?

Identity Matrix

Definition

A matrix is said to be **square** if it has the same number of rows and columns. In other words, an $n \times n$ matrix is a square matrix.

Identity Matrix

Definition

A matrix is said to be **square** if it has the same number of rows and columns. In other words, an $n \times n$ matrix is a square matrix.

Definition

A square matrix A is **symmetric** if $A^T = A$, i.e., the transpose of A is equal to A.

Identity Matrix

Definition

A matrix is said to be **square** if it has the same number of rows and columns. In other words, an $n \times n$ matrix is a square matrix.

Definition

A square matrix A is **symmetric** if $A^T = A$, i.e., the transpose of A is equal to A.

Example

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & -3 & 4 \\ 1 & 4 & 6 \end{bmatrix}$$

This matrix is both symmetric and (of course) square.

Identity Matrix

Definition

The **main diagonal** (or just the **diagonal**) of a matrix A are the terms a_{ii} for which the row and column indices are the same (a_{11}, a_{22}, $\ldots$), so from the upper left element to the lower right.

Identity Matrix

Definition

The **main diagonal** (or just the **diagonal**) of a matrix A are the terms a_{ii} for which the row and column indices are the same $(a_{11}, a_{22}, \dots)$, so from the upper left element to the lower right.

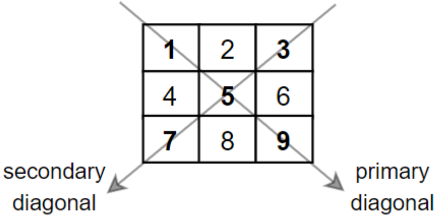
Similarly, the other diagonal from the upper right element to the lower left is called the **secondary diagonal**.

Identity Matrix

Definition

The **main diagonal** (or just the **diagonal**) of a matrix A are the terms a_{ii} for which the row and column indices are the same $(a_{11}, a_{22}, \dots)$, so from the upper left element to the lower right.

Similarly, the other diagonal from the upper right element to the lower left is called the **secondary diagonal**.



Identity Matrix

For example, here the main diagonal is marked with red:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Identity Matrix

For example, here the main diagonal is marked with red:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The secondary diagonal is marked with red:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Identity Matrix

Definition

The **identity matrix** of $\mathbb{R}^{n\times n}$, denoted as I_n, is the square matrix with ones on the main diagonal and zeros elsewhere.

Applying the identity matrix on vectors does not change them.

Identity Matrix

Definition

The **identity matrix** of $\mathbb{R}^{n\times n}$, denoted as I_n, is the square matrix with ones on the main diagonal and zeros elsewhere.

Applying the identity matrix on vectors does not change them.

Example

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ is the } 3 \times 3 \text{ identity matrix.}$$

Identity Matrix

Definition

The **identity matrix** of $\mathbb{R}^{n \times n}$, denoted as I_n, is the square matrix with ones on the main diagonal and zeros elsewhere.

Applying the identity matrix on vectors does not change them.

Example

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ is the } 3 \times 3 \text{ identity matrix.}$$

Therefore, we can say:

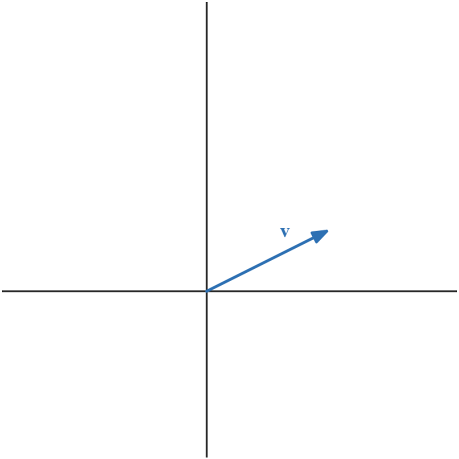
Property

For any matrix $A \in \mathbb{R}^{m \times n}$,

$$I_m A = A I_n = A$$

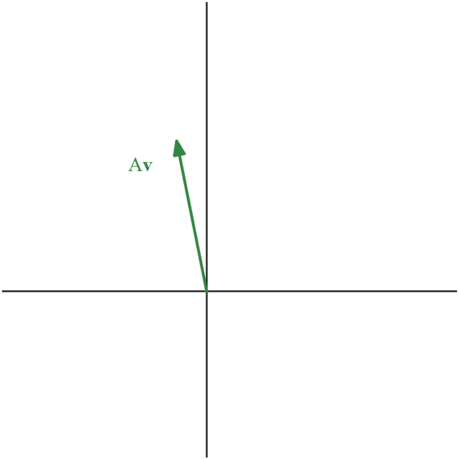
Inverse Matrix

Finally, what if we have a vector in $\mathbb{R}^n$,



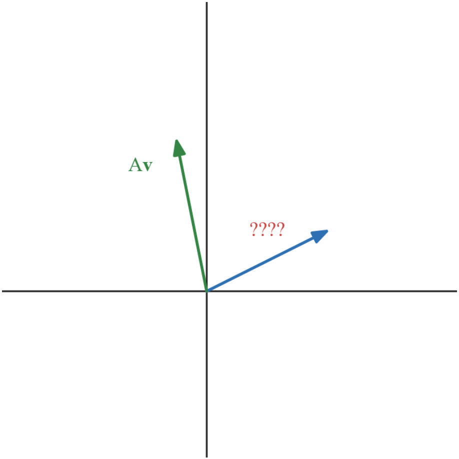
Inverse Matrix

Finally, what if we have a vector in $\mathbb{R}^n$, and we accidentally transform it?



Inverse Matrix

How to get back to the original vector?



Inverse Matrix

In other words, in terms of what we learned about matrix multiplication,

$$\textbf{what} \times A = I \quad ?$$

Inverse Matrix

In other words, in terms of what we learned about matrix multiplication,

$$\textbf{what} \times A = I \quad ?$$

We call that matrix the **inverse** of A, and we denote it by A^{-1}.

Inverse Matrix

In other words, in terms of what we learned about matrix multiplication,

$$A^{-1} \times A = I \quad !$$

We call that matrix the **inverse** of A, and we denote it by A^{-1}.

Inverse Matrix

Question

Assume the matrix $A \in \mathbb{R}^{n \times n}$ does the following when applied on a vector:

1. scales the vector up 2 times in the horizontal direction,
2. then rotates it by 30° clockwise,
3. then squishes it down 3 times in the vertical direction,
4. and then flips it horizontally (around the x-axis) $\sim$

Inverse Matrix

Question

Assume the matrix $A \in \mathbb{R}^{n \times n}$ does the following when applied on a vector:

1. scales the vector up 2 times in the horizontal direction,
2. then rotates it by 30° clockwise,
3. then squishes it down 3 times in the vertical direction,
4. and then flips it horizontally (around the x-axis) $\sim$

Given $\mathbf{v} = A\mathbf{u}$, could we recover the original $\mathbf{u}$?

Inverse Matrix

Question

Assume the matrix $A \in \mathbb{R}^{n \times n}$ does the following when applied on a vector:

1. scales the vector up 2 times in the horizontal direction,
2. then rotates it by 30° clockwise,
3. then squishes it down 3 times in the vertical direction,
4. and then flips it horizontally (around the x-axis) $\sim$

Given $\mathbf{v} = A\mathbf{u}$, could we recover the original $\mathbf{u}$?

The answer is yes, i.e. the matrix A has an inverse. As we will see soon, only some square matrices actually have an inverse.

Trace

Definition

The **trace** of a square matrix A, denoted as $\mathrm{tr}(A)$, is the sum of the elements on its main diagonal.

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

Trace

Definition

The **trace** of a square matrix A, denoted as $\mathrm{tr}(A)$, is the sum of the elements on its main diagonal.

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

Example

If

$$A = \begin{bmatrix} 2 & 5 & 1 \\ 0 & -3 & 4 \\ 7 & 2 & 6 \end{bmatrix}$$

then

$$\mathrm{tr}(A) = 2 + (-3) + 6 = 5$$

Trace

Definition

The **trace** of a square matrix A, denoted as $\text{tr}(A)$, is the sum of the elements on its main diagonal.

$$\text{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

Example

If

$$A = \begin{bmatrix} 2 & 5 & 1 \\ 0 & -3 & 4 \\ 7 & 2 & 6 \end{bmatrix}$$

then

$$\text{tr}(A) = 2 + (-3) + 6 = 5$$

Note that only square matrices have a trace.

Trace

Trace Properties

For any matrices A and B, and any scalar c, the trace of a matrix satisfies the following properties:

- $\mathrm{tr}(cA) = c \cdot \mathrm{tr}(A)$
- $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$
- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
- $\mathrm{tr}(A^T) = \mathrm{tr}(A)$

Determinant of a 2×2 Matrix

Determinant Formula

For a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is given by

$$\det(A) = ad - bc$$

Determinant of a 2×2 Matrix

Determinant Formula

For a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is given by

$$\det(A) = ad - bc$$

Example

For the matrix

$$A = \begin{bmatrix} 2 & 5 \\ -3 & 4 \end{bmatrix}$$

the determinant is $\det(A) = (2)(4) - (5)(-3) = 8 + 15 = 23.$

Determinant of a 3×3 Matrix

Determinant Formula

For a 3×3 matrix

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

the determinant is given by

$$\det(C) = aei + bfg + cdh - ceg - bdi - afh$$

Determinant of a 3 × 3 Matrix

Determinant Formula

For a 3×3 matrix

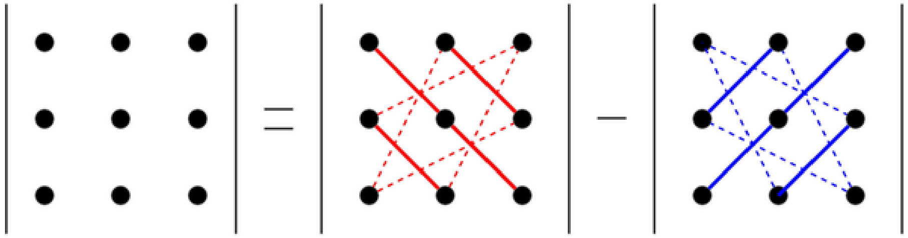
$$C = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

the determinant is given by

$$\det(C) = aei + bfg + cdh - ceg - bdi - afh$$

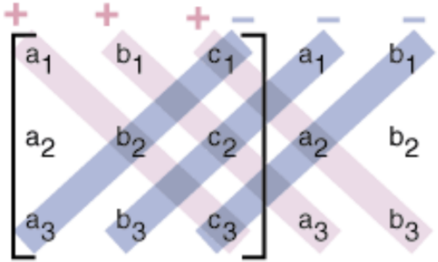
Forget that formula–remember the algorithm!

Determinant of a 3×3 Matrix



Determinant of a 3×3 Matrix

Alternatively,



$$\det A = (a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3) - (a_3 b_2 c_1 + b_3 c_2 a_1 + c_3 a_2 b_1)$$

Determinant of a 3×3 Matrix

Example

For the matrix

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\det(C) = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 3 \cdot 5 \cdot 7 - 2 \cdot 4 \cdot 9 - 1 \cdot 6 \cdot 8 = 0$$

Determinant of a 3×3 Matrix

Example

For the matrix

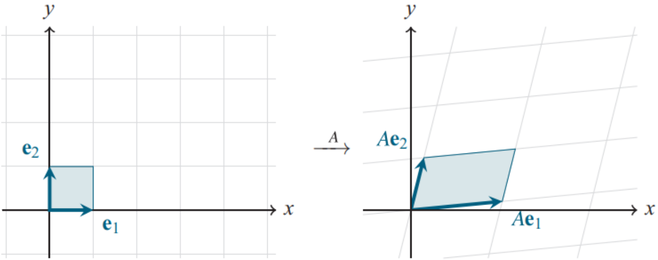
$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\det(C) = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 3 \cdot 5 \cdot 7 - 2 \cdot 4 \cdot 9 - 1 \cdot 6 \cdot 8 = 0$$

But what does the determinant show, and how do we need it?

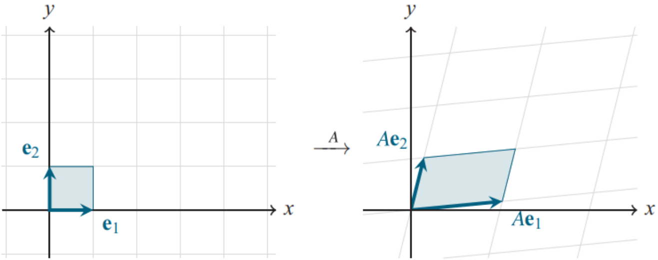
Determinant

If we take, for example, the so-called "unit square" formed by the vectors $\mathbf{e}_1 = [1 \ \ 0]$ and $\mathbf{e}_2 = [0 \ \ 1]$, we can see that their transformed versions, $A\mathbf{e}_1$ and $A\mathbf{e}_2$, form a parallelogram:



Determinant

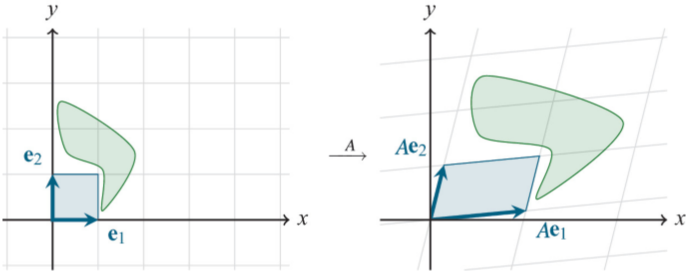
If we take, for example, the so-called "unit square" formed by the vectors $\mathbf{e}_1 = [1 \;\; 0]$ and $\mathbf{e}_2 = [0 \;\; 1]$, we can see that their transformed versions, $A\mathbf{e}_1$ and $A\mathbf{e}_2$, form a parallelogram:



Then $\det(A)$ is the area of that parallelogram.

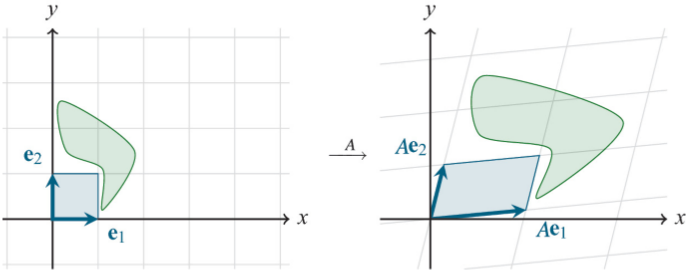
Determinant

More generally, after we apply the transformation A (play that animation in your head), the area of *any shape* gets scaled by the factor of $\det(A)$:



Determinant

More generally, after we apply the transformation A (play that animation in your head), the area of *any shape* gets scaled by the factor of $\det(A)$:



So the determinant shows how much the matrix scales up everything in average. Note that it is defined **only** for square matrices.

Determinant

Determinant Properties

Let $A, B \in \mathbb{R}^{n \times n}$ be square matrices of the same size, and let $c \in \mathbb{R}$ be any scalar. Then:

- $\det(cA) = c^n \cdot \det(A)$ *(where n is the size of the matrix)*
- $\det(AB) = \det(A) \cdot \det(B)$ *(multiplicativity)*
- $\det(I) = 1$
- If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(A^T) = \det(A)$ *(invariance under transpose)*
- If all numbers on some row or some column of A are zero, then $\det(A) = 0$
- If $\det(A) < 0$, then A flips the space around.

Determinant

Determinant Properties

Let $A, B \in \mathbb{R}^{n \times n}$ be square matrices of the same size, and let $c \in \mathbb{R}$ be any scalar. Then:

- $\det(cA) = c^n \cdot \det(A)$ *(where n is the size of the matrix)*
- $\det(AB) = \det(A) \cdot \det(B)$ *(multiplicativity)*
- $\det(I) = 1$
- If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(A^T) = \det(A)$ *(invariance under transpose)*
- If all numbers on some row or some column of A are zero, then $\det(A) = 0$
- If $\det(A) < 0$, then A flips the space around.

It would be an exercise of huge importance to attempt proving these properties (except the last three) by playing the matrices in your head.

Determinant

Finally,

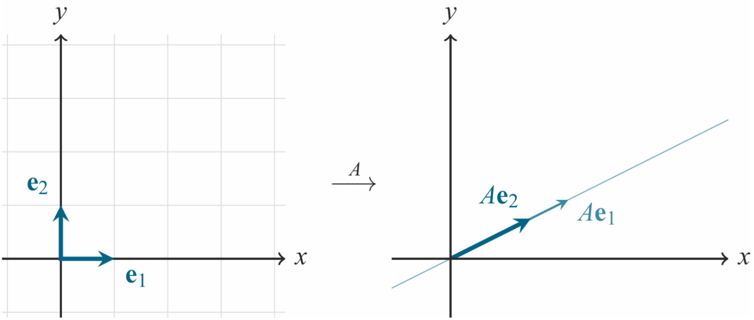
Question

What does it mean if $\det A = 0$?

Finally,

Question

What does it mean if $\det A = 0$?

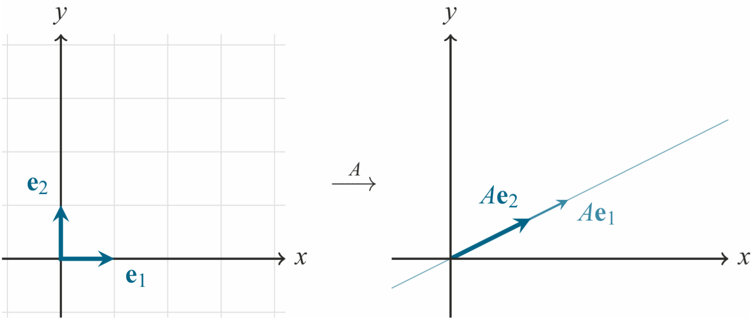


Determinant

Finally,

Question

What does it mean if $\det A = 0$?



Theorem

A square matrix A has an inverse if and only if its determinant is not zero.

Inverse Matrix

Formula for 2x2

For a 2×2 invertible matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse A^{-1} can be calculated using the formula:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Inverse Matrix

Formula for 2x2

For a 2×2 invertible matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse A^{-1} can be calculated using the formula:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Example

Given $A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ with $\det A = (2 \times 4) - (3 \times 1) = 5$, we can calculate the inverse as follows:

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix}$$