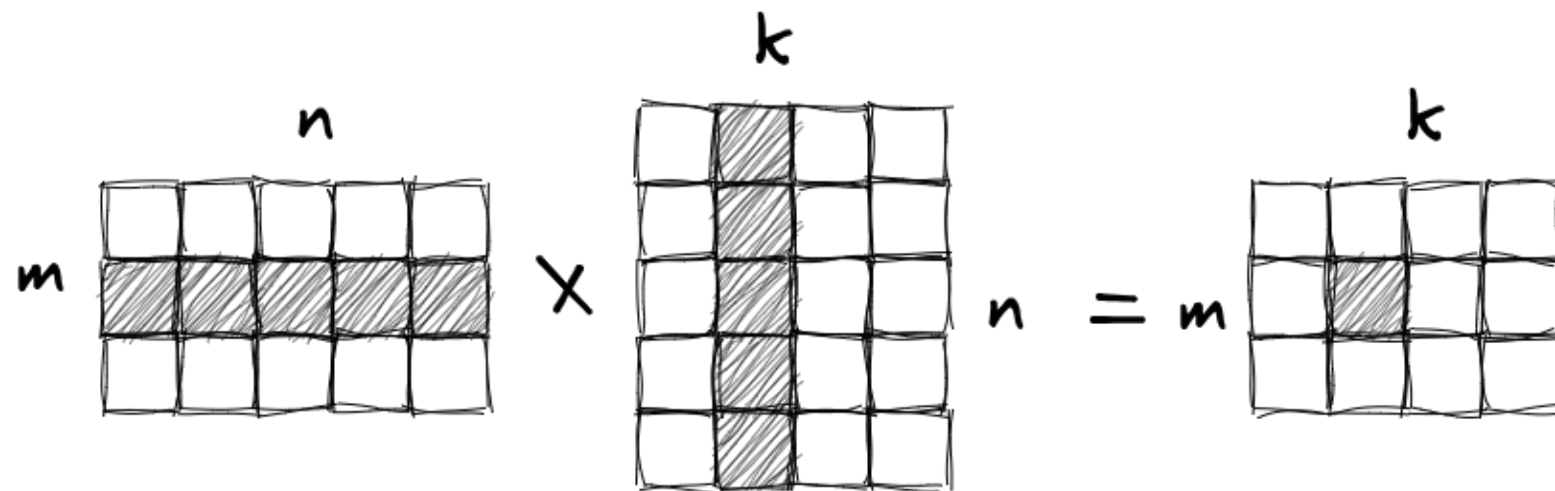


# Linear algebra

## Vectors and Matrices



Week 6

Middlesex University Dubai; CST4050: Winter21  
Instructor: Dr. Ivan Reznikov

# Plan

- Vectors
- Matrices
- Determinant

# Why need linear algebra?

All data science, especially machine learning, depends on linear algebra and statistics. Knowing math is crucial to understanding how neural networks and other ML algorithms work.

$$\begin{aligned} E(Y) &= \int_0^{\infty} \frac{y^{\alpha} e^{-y/\beta} dy}{\beta^{\alpha} \Gamma(\alpha)} = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} \int_0^{\infty} y^{\alpha} e^{-y/\beta} dy \\ &= \frac{1}{\beta^{\alpha} \Gamma(\alpha)} [\beta^{\alpha+1} \Gamma(\alpha+1)] = \frac{\beta \alpha \Gamma(\alpha)}{\Gamma(\alpha)} = \alpha \beta. \end{aligned}$$

# 2D array

Most common form of data organization for machine learning is a 2D array.

Rows represent observations (records, items, data points).

Columns represent attributes (features, variables).

Natural to think of each sample as a vector of attributes, and whole array as a matrix

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	
0	75.0	0	582	0	20	1	265000.00	
1	55.0	0	7861	0	38	0	263358.03	
2	65.0	0	146	0	20	0	162000.00	vector K
3	50.0	1	111	0	20	0	210000.00	
4	65.0	1	160	1	20	0	327000.00	vector N

# Vectors

Vector is a n-tuple of values (usually real numbers).

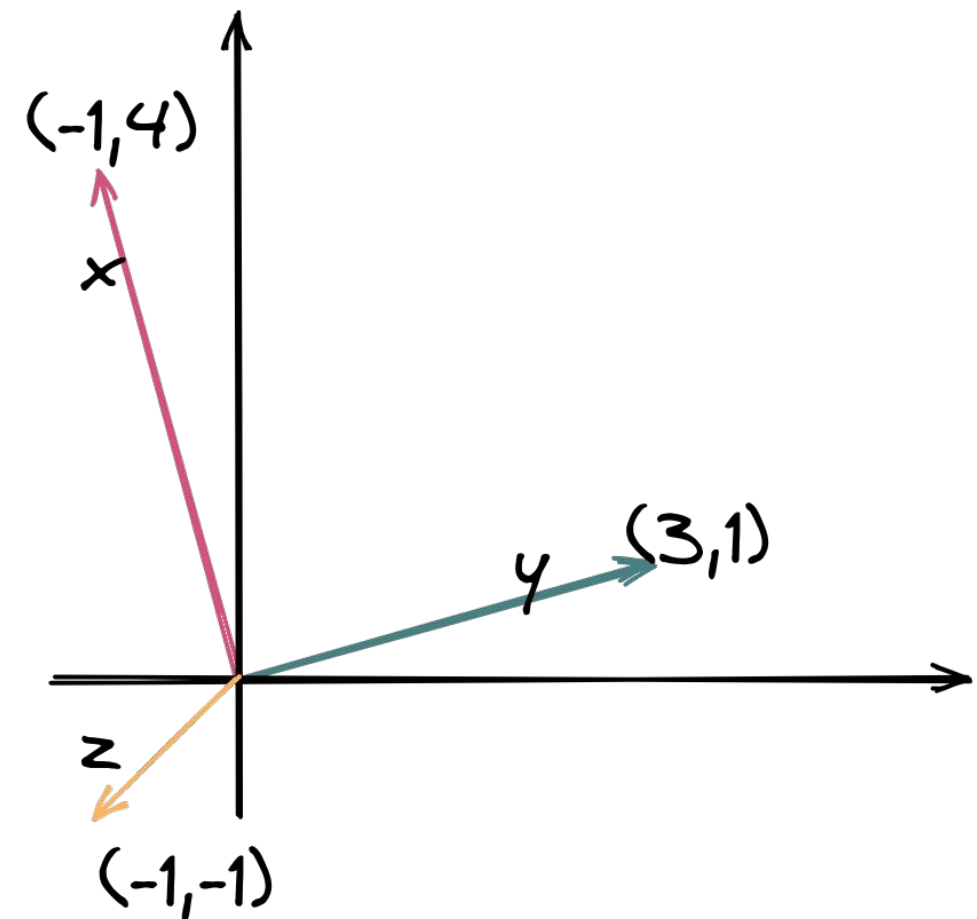
A vector can be seen as a point in space or a directed line segment with a magnitude (length) and direction.

Scalar values are defined only by magnitude.

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

$$X^T = (-1, 4)$$

"transpose"



# Vector arithmetic

Let  $\mathbf{a} = (a_1, \dots, a_n)^T$  and  $\mathbf{b} = (b_1, \dots, b_n)^T$  be two vectors.

Let vector  $\mathbf{z} = \mathbf{a} + \mathbf{b}$

$$\mathbf{z} = (a_1 + b_1, \dots, a_n + b_n)^T$$

Examples:

$$\mathbf{a} = (3, 7)^T; \mathbf{b} = (2, -3)^T; \Rightarrow \mathbf{z} = (5, 4)^T$$

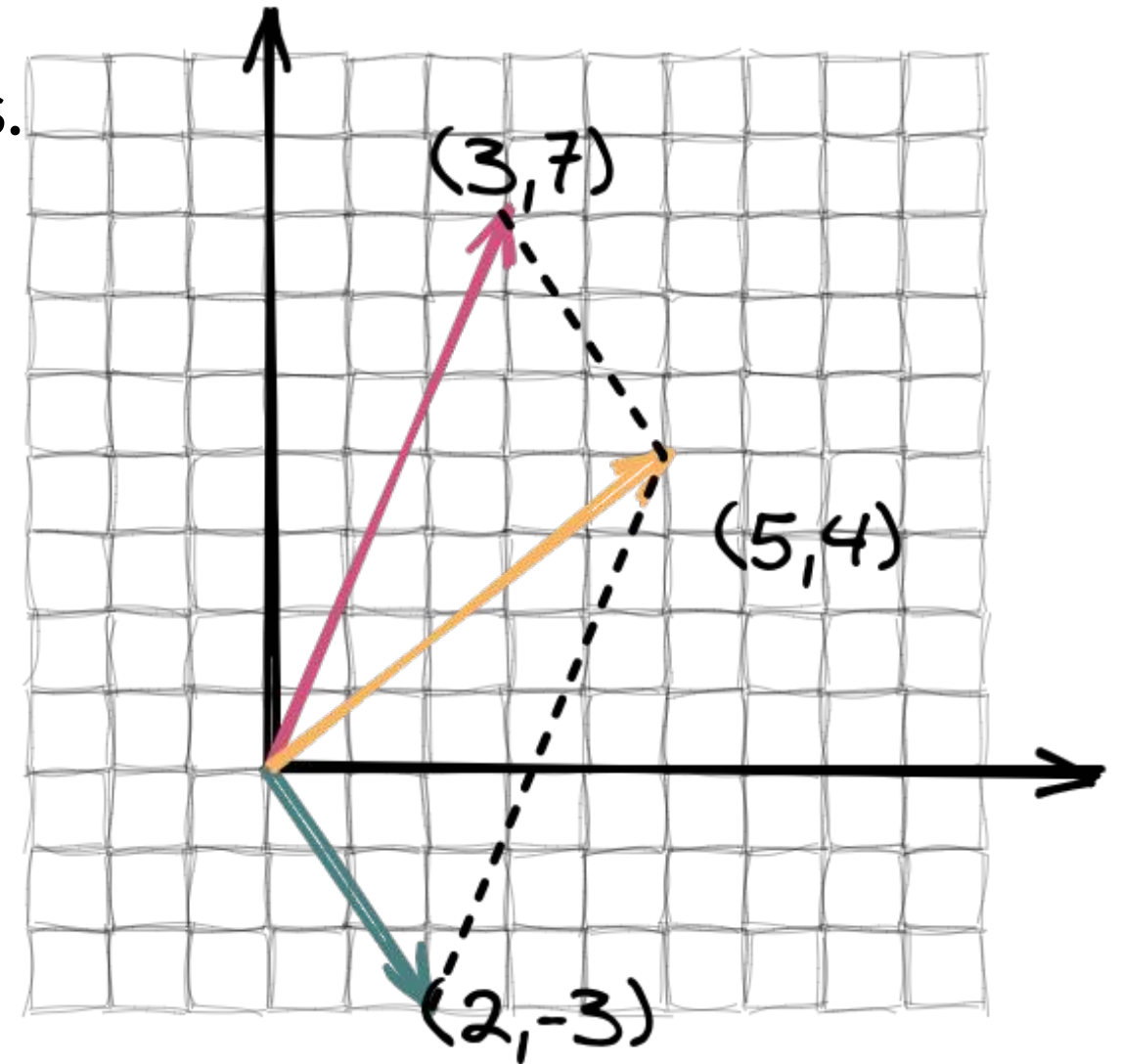
$$\mathbf{a} = (3, 2, 1)^T; \mathbf{b} = (1, 2, 0)^T; \Rightarrow \mathbf{z} = (4, 4, 1)^T$$

Let  $k$  be scalar,  $\mathbf{a} = (a_1, \dots, a_n)^T$

Let vector  $\mathbf{z} = k \times \mathbf{a}, \Rightarrow \mathbf{z} = (k \times a_1, \dots, k \times a_n)^T$

Example:

$$\mathbf{a} = (3, 7)^T, k = 3 \Rightarrow \mathbf{z} = (9, 21)^T$$



# Vector arithmetic

Let  $\mathbf{a} = (a_1, \dots, a_n)^T$  and  $\mathbf{b} = (b_1, \dots, b_n)^T$  be two vectors.

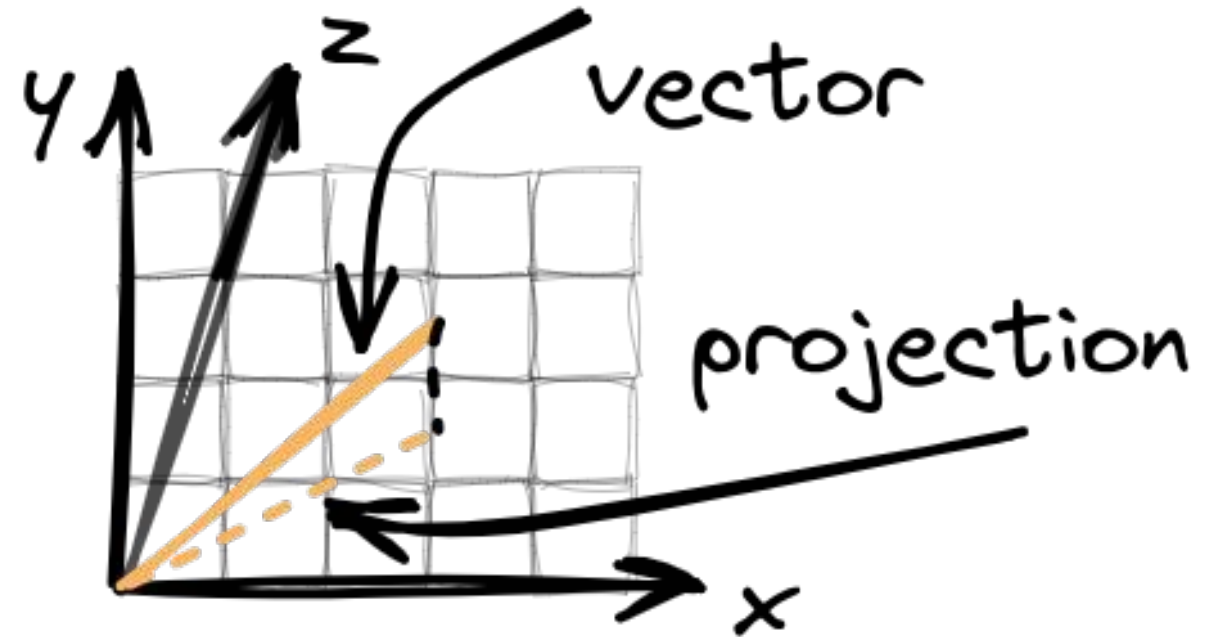
Let vector  $\mathbf{z} = \mathbf{a} \times \mathbf{b}$

$$\mathbf{z} = \sum (a_1 \times b_1, \dots, a_n \times b_n)^T$$

Examples:

$$\mathbf{a} = (3, 7)^T; \mathbf{b} = (2, -3)^T; \Rightarrow \mathbf{z} = \sum (6, -21)^T = -15$$

$$\mathbf{a} = (3, 2, 1)^T; \mathbf{b} = (1, 2, 0)^T; \Rightarrow \mathbf{z} = \sum (3, 4, 0)^T = 7$$



Projection: projection of  $y$  onto  $x$  is a perpendicular line from  $y$  onto  $x$  (meet at point) and the projection vector is the vector to that point.

$$\text{Projection}_{\mathbf{a}}(\mathbf{b}) = ((\mathbf{a} \times \mathbf{b}) \times \mathbf{a}) / (\mathbf{a} \times \mathbf{a})$$

Example:

$$\begin{aligned} \text{Projection}_{(4,3,0)}((25,0,5)) &= (25 \times 4 + 3 \times 0 + 0 \times 5) \times (4, 3, 0) / (4^2 + 3^2 + 0^2) = \\ &= 100 \times (4, 3, 0) / 25 = 4 \times (4, 3, 0) = (16, 12, 0) \end{aligned}$$

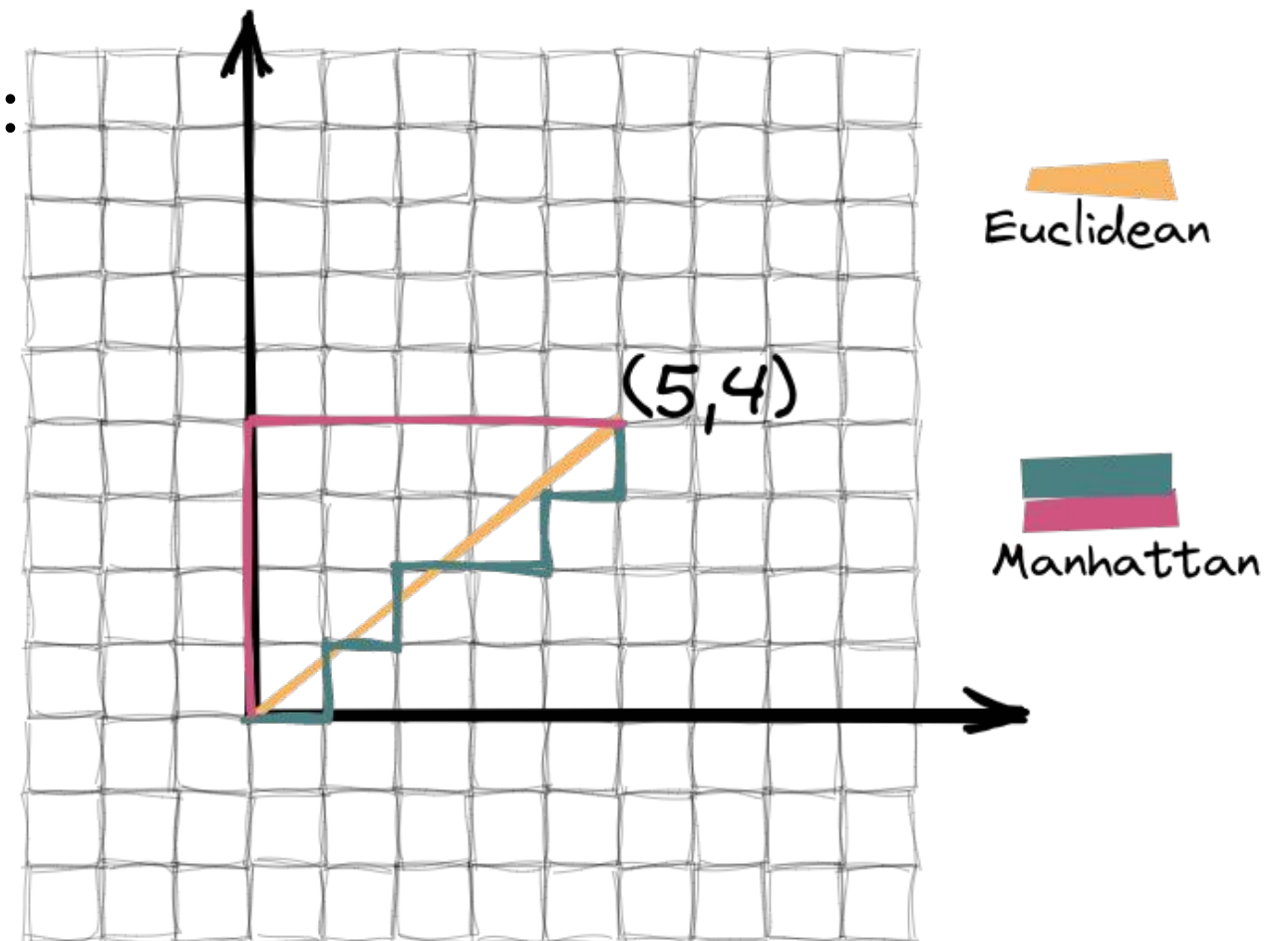
# Norm of a vector

Norm of a vector may be understood as distance:

$$d(x,y) = ||y - x||$$

There are more than one type of distance:

- Euclidean
  - Manhattan
  - Minkowski
- etc





# Matrix arithmetic

Definition: an  $m \times n$  two-dimensional array of values (usually real numbers).

- $m$  rows
- $n$  columns

Matrix referenced by two-element subscript

- 1<sup>st</sup> element in subscript is row
- 2<sup>nd</sup> element in subscript is column

=>  $a_{\text{row, columns}}$

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \bullet & \bullet & \bullet & a_{1,N} \\ \bullet & \bullet & & & & \bullet \\ \bullet & & \bullet & & & \bullet \\ \bullet & & & \bullet & & \bullet \\ a_{N,1} & \bullet & \bullet & \bullet & \bullet & a_{N,N} \end{pmatrix}$$

Vector can be considered as 1D matrix => as vectors

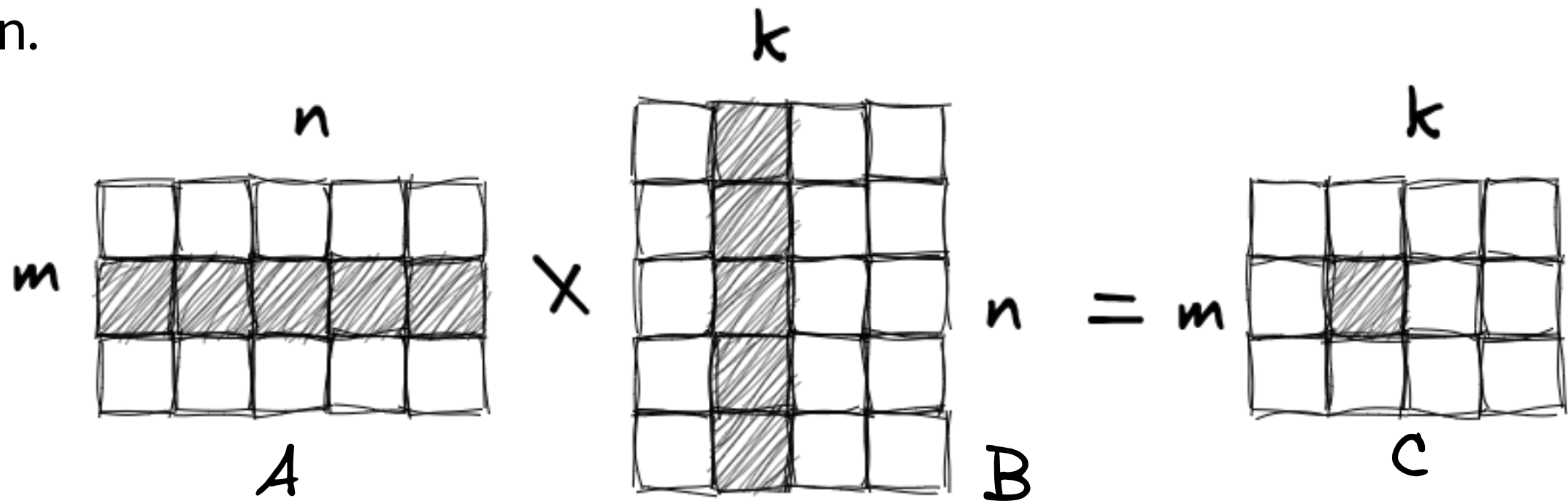
- transpose
- addition
- multiplication by scalar

# Matrix multiplication

Matrix-matrix multiplication is defined as the rows by columns multiplication:

$$c_{i,j} = a_{i,1} \times b_{1,j} + \dots + a_{i,n} \times b_{n,j} = \sum a_{i,z} \times b_{z,j}$$

A vector-matrix multiplication just a special case of a matrix-matrix multiplication.

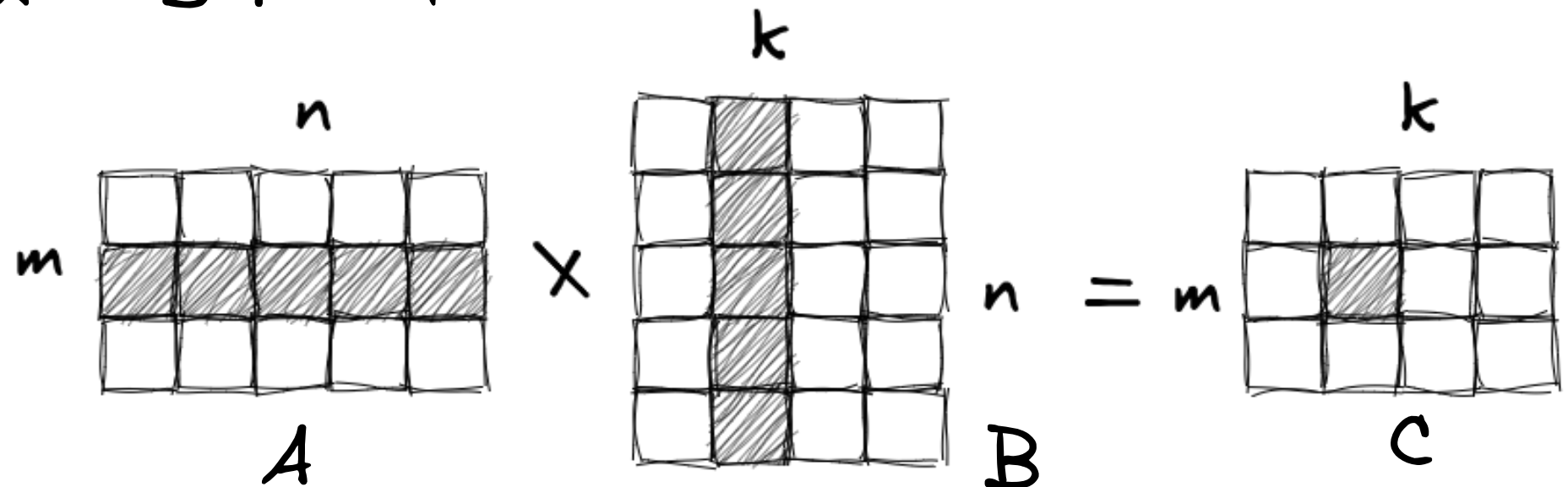


# Matrix multiplication

$$A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix} \quad C = A \times B = \begin{matrix} & c_{i,1} & c_{i,2} & c_{i,3} \\ \begin{matrix} c_1 \\ c_2 \end{matrix} & \begin{bmatrix} 4 & 9 & 12 \\ 9 & 13 & 17 \end{bmatrix} \end{matrix}$$

■  $c_{1,3} = 2 \times 2 + 1 \times 2 + 2 \times 3 = 12$

■  $c_{2,1} = 1 \times 0 + 3 \times 2 + 3 \times 1 = 9$



# Matrix multiplication

Matrix multiplication is associative:

$$A \times (B \times C) = (A \times B) \times C$$

Matrix multiplication is not commutative:

$$A \times B \neq B \times A$$

Matrix transposition rule:

$$(A \times B)^T = B^T \times A^T$$

# Linear transformation

$$A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix} \quad x = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad C = Ax = \begin{bmatrix} 2 \times a + 1 \times b + 2 \times c \\ 1 \times a + 3 \times b + 3 \times c \end{bmatrix}$$

$\mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$T(x) = Ax$$

Function T is a linear transformation, in fact for each vector  $x, y$  and scalar  $c$ :

$$A(x+y) = A(x) + A(y)$$

$$A(cx) = cA(x)$$

Square matrices are important as  $\mathbb{R}^n \rightarrow \mathbb{R}^n$

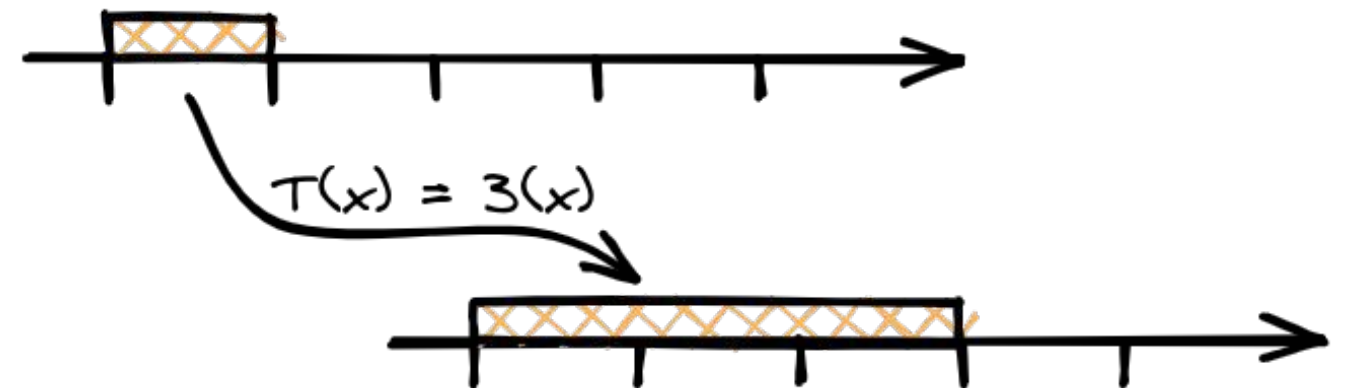
# 1D linear transformation

A one-dimensional linear transformation is a function  $T(\mathbf{x}) = a(\mathbf{x})$  for some scalar  $a$ .

To view the one-dimensional case in the same way we view higher dimensional linear transformations, we can view  $a$  as a  $1 \times 1$  matrix.

Example: one-dimensional linear transformation is the function  $T(\mathbf{x}) = 3(\mathbf{x})$ .

A visualization of this function by its graph:

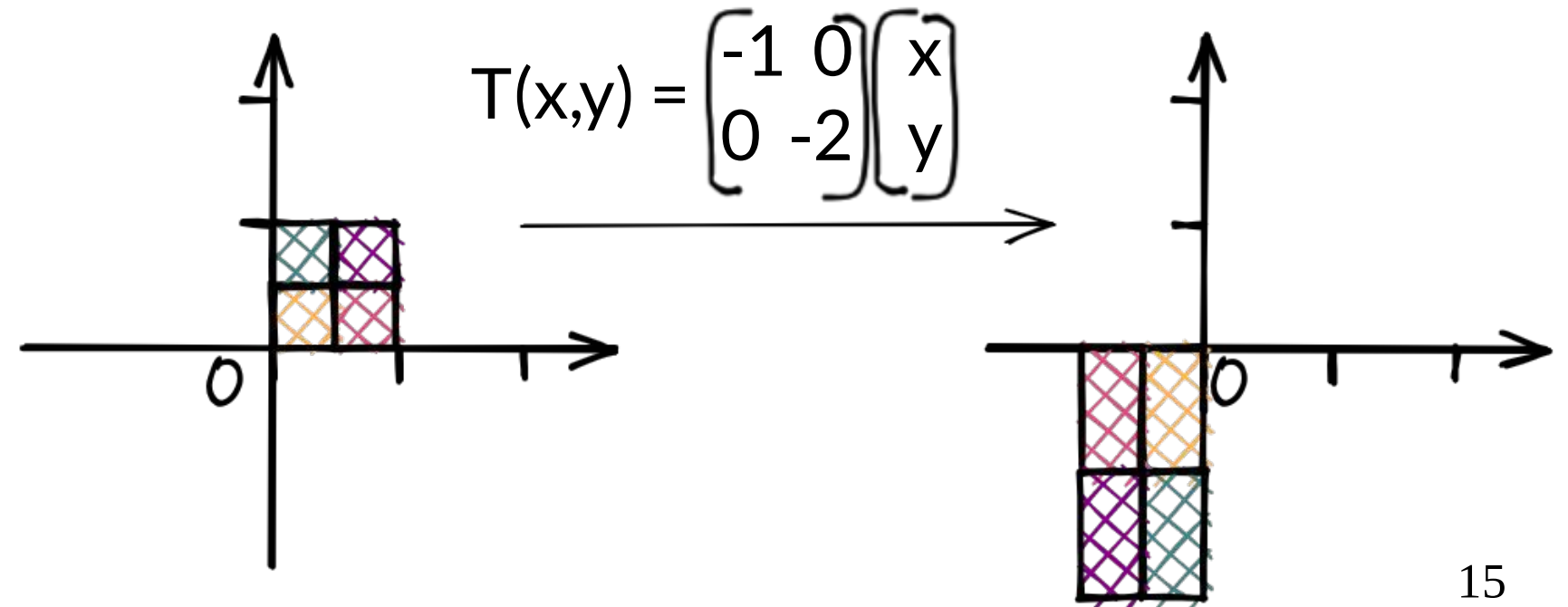


# 2D linear transformation

A two-dimensional linear transformation is a function  $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  of the form:

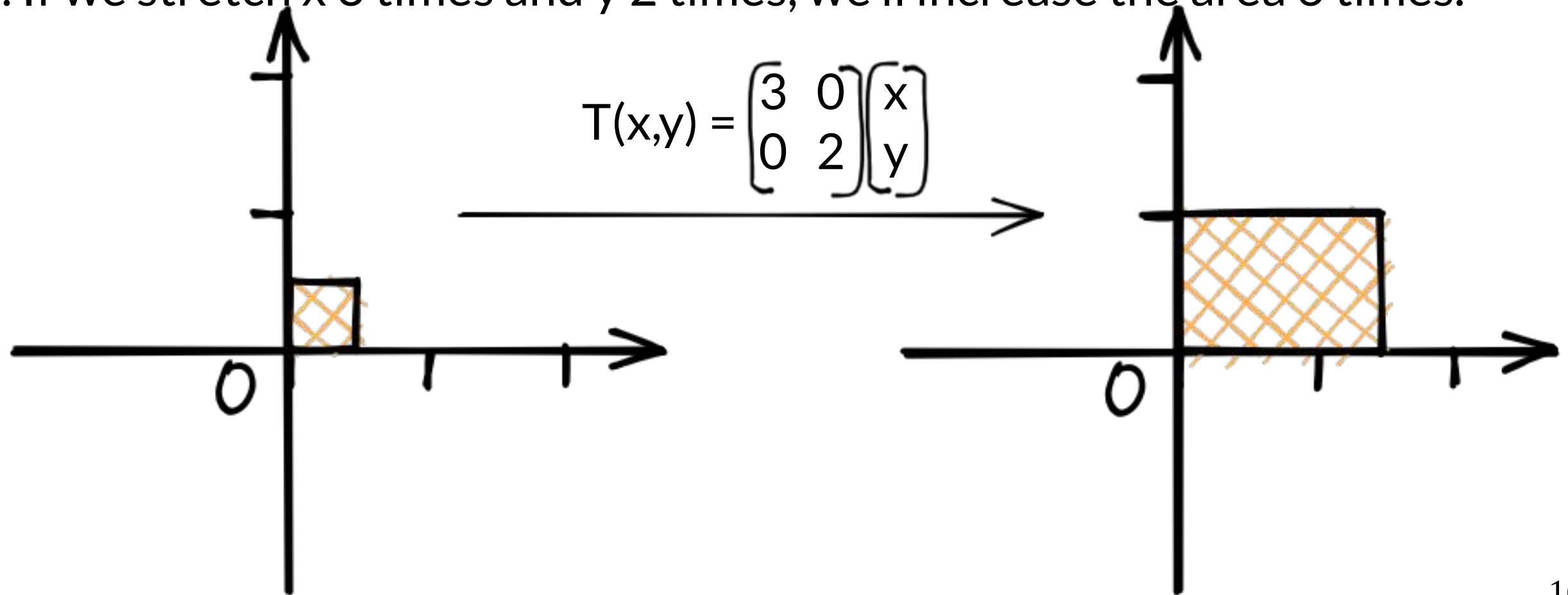
$$T(x,y) = (ax+by, cx+dy) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

We can write this more succinctly as  $T(\mathbf{x}) = A\mathbf{x}$ ,  
where  $\mathbf{x} = [x, y]^T$  and  $A$  is the  $2 \times 2$  matrix.



# Determinant

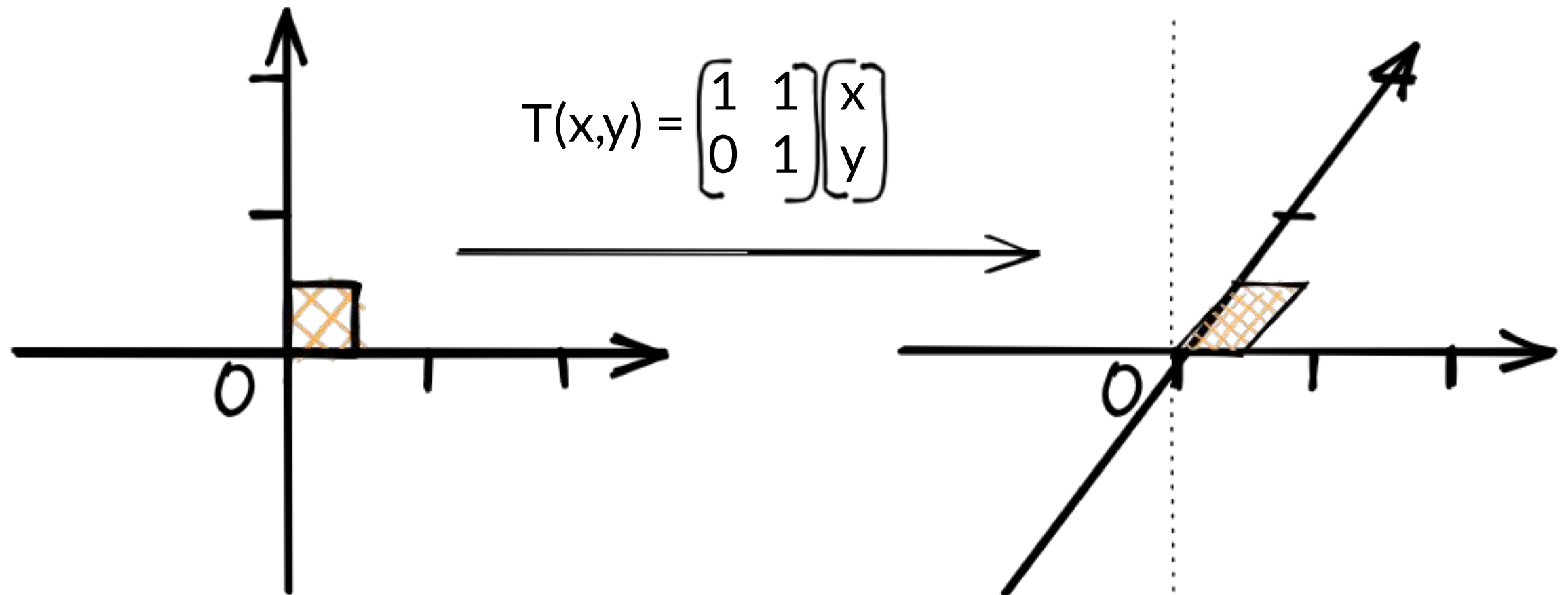
During linear transformations, we perform stretching and squishing some of the dimensions. It would be valuable to determine how our item's area has changed. If we stretch  $x$  3 times and  $y$  2 times, we'll increase the area 6 times.





# Determinant

If we don't change the x and y values, no matter how we tilt our item, it's area won't change.



# Determinant

The scaling factor, by which the linear transformation changes items area is called determinant.

$$T(x,y) = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\det(A) = 6$$

$$T(x,y) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\det(A) = 1$$

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

$$\det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = a \times \det\left(\begin{bmatrix} e & f \\ h & i \end{bmatrix}\right) - b \times \det\left(\begin{bmatrix} d & f \\ g & i \end{bmatrix}\right) + c \times \det\left(\begin{bmatrix} d & e \\ g & h \end{bmatrix}\right)$$

# Determinant

In general, in any dimension  $n$ , the determinant is a scalar value that can be computed from the elements of a square matrix and encodes certain properties of the linear transformation described by the matrix.

The determinant of a matrix  $\mathbf{A}$  is denoted  $\det(\mathbf{A})$ .

A determinant of a square matrix  $A$  can positive, negative or zero.

- Positive determinants denote transformations having a positive area, volume or hyper-volume.
- Negative determinants denote transformations having a negative area, volume or hyper-volume.
- Zero determinants denote transformations having no area, no volume or no hyper-volume

# Matrix inversion

The inverse of a number  $a$  is such that  $a \times a^{-1} = 1$

For example the inverse of 10 is 0.1, as  $10 \times 0.1 = 1$

The inverse of 5 is 0.2, and the inverse of 0.01 is 100.

The inverse of a matrix  $\mathbf{A}$  is that matrix  $\mathbf{A}^{-1}$  such that  $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Sometimes there is no inverse at all. In this case we say that the matrix  $\mathbf{A}$  is not invertible.

A square matrix that is not invertible is called singular.

A square matrix is singular if and only if its determinant is 0.