Principles and Practices of Data Science

Lecture 6

Melvin Ayala

Lecture 6: Geometry

Sections

1. Norm of a Vector

What is the norm of a vector?

length of the vector = vector norm = vector's magnitude

the length of a vector:

- nonnegative number
- extent of the vector in space

Given an n-dimensional vector,

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

the norm can be calculated as:

$$\|X\|_n = \left(\sum_i \|x_i\|^n \right)^{1/n}$$

A special case is defined when n tends to infinity:

$$\|X\|_\infty = \max_i \|x_i\|$$

most commonly encountered vector norm:

L2-norm = $\|x\|_2 = \sqrt{{x_1}^2 + {x_2}^2 + \ldots + {x_n}^2}$

Some norms for v = (1, 2, 3):

$$L^1 - \text{norm} = \|x\|_1 = \left((x_1)^1 + (x_2)^1 + (x_3)^1\right)^{1/1} = (1^1 + 2^1 + 3^1)^{1/1} = 6$$

$$L^2 - \text{norm} = \|x\|_2 = \left((x_1)^2 + (x_2)^2 + (x_3)^2\right)^{1/2} = (1^2 + 2^2 + 3^2)^{1/2} = 3.742$$

$$L^3 - \text{norm} = \|x\|_3 = \left((x_1)^3 + (x_2)^3 + (x_3)^3\right)^{1/3} = (1^3 + 2^3 + 3^3)^{1/3} = 3.302$$

$$L^4 - \text{norm} = \|x\|_4 = \left((x_1)^4 + (x_2)^4 + (x_3)^4\right)^{1/4} = (1^4 + 2^4 + 3^4)^{1/4} = 3.146$$

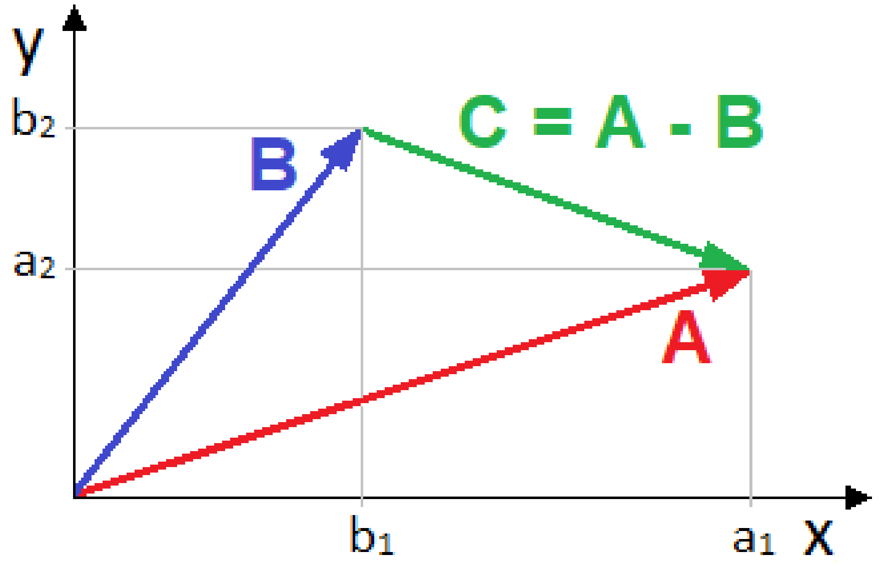
$$L^\infty - \text{norm} = \|x\|_\infty = \left((x_1)^\infty + (x_2)^\infty + (x_3)^\infty\right)^{1/\infty} = (1^\infty + 2^\infty + 3^\infty)^{1/\infty} = 3$$

2. Dot Product of Two Vectors

If A and B are two vectors with a common initial point, then A, B, and C = A – B form a triangle.

By the Law of Cosines: $\|A - B\|^2 = \|A\|^2 + \|B\|^2 - 2\|A\|\,\|B\|\cos\theta$
with θ: angle between the two vectors.

Using the formula for the magnitude of a vector (with 2 dimensions):
$\|A\|^2 = a_1^2 + a_2^2$
$\|B\|^2 = b_1^2 + b_2^2$
$\|A - B\|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2$



we obtain:
$\|A - B\|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2$
$\qquad = a_1^2 + a_2^2 + b_1^2 + b_2^2 - 2\|A\|\,\|B\|\cos\theta$

therefore:
$(a_1 - b_1)^2 + (a_2 - b_2)^2 = a_1^2 + a_2^2 + b_1^2 + b_2^2 - 2\|A\|\,\|B\|\cos\theta$
$a_1^2 - 2a_1b_1^2 + b_1^2 + a_2^2 - 2a_2b_2^2 + b_2^2 - a_1^2 - a_2^2 - b_1^2 - b_2^2 = -2\|A\|\,\|B\|\cos\theta$
…
$a_1b_1 + a_2b_2 = \|A\|\,\|B\|\cos\theta$

What is the dot product of two vectors?

- is the sum of the products of corresponding components.
- is the product of their magnitudes, times the cosine of the angle between them.
- in a vector space, it is a way to multiply vectors together, with the result of this multiplication being a scalar.
- the dot product of a vector with itself is the square of its magnitude.

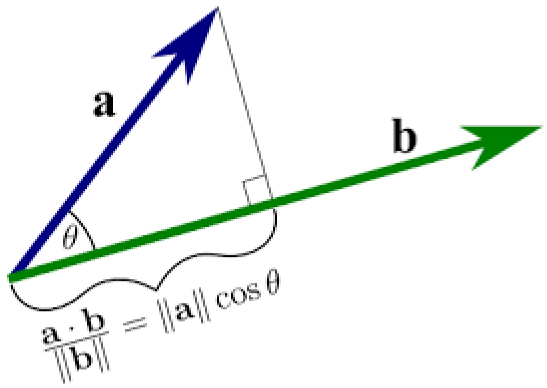
Definition:

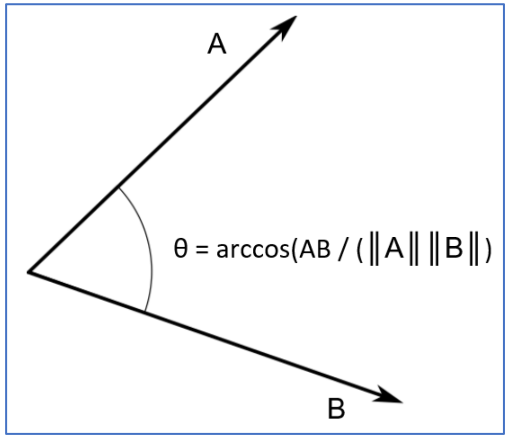
The dot product of two vectors A and B is defined as:

$A\ B = a_1b_1 + a_2b_2$

$A\ B = \|A\|\ \|B\| \cos(\theta)$

where θ is the angle between A and B.





Note: Other notations are common, such as <A, B>

An Equivalent Definition of a Dot Product

Definition:

The dot product of two vectors A and B can also be calculated as:

$$A \cdot B = \sum_{i=1}^{n} a_i b_i$$

$$A \cdot B = a_1 b_1 + \; a_2 b_2 + \cdots + \; a_n b_n$$

Example:

What is the dot product of A = [1, 2] and B=[3, 4]?

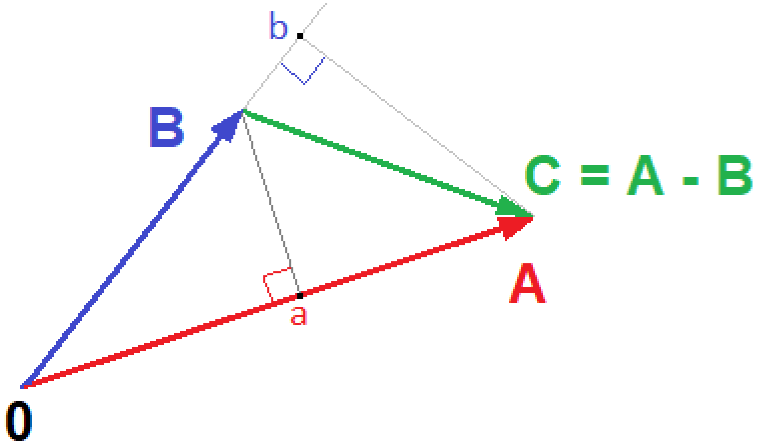
$$A \cdot B = \sum_{i=1}^{2} a_i b_i$$

$$A \cdot B = a_1 b_1 + \; a_2 b_2 = 1(3) + 2(4) = 11$$

Geometric Interpretation of the Dot Product

Geometric Property:

- the dot product of two vectors is commutative (the order does not matter)
- two nonzero vectors are perpendicular, or orthogonal, if and only if their dot product is equal to zero.



$$\overline{0A} * \overline{0a} = \overline{0B} * \overline{0b}$$

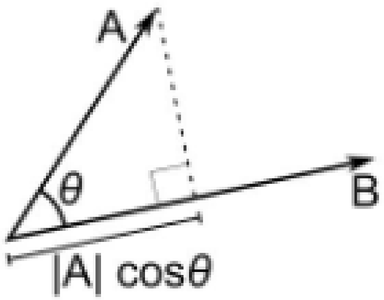
Other Properties:

- multiplying a vector by a constant multiplies its dot product with any other vector by the same constant.
- the dot product of a vector with the zero vector is zero.

Geometric Interpretation of the Dot Product (Cont.)

Geometric interpretation of dot product

A dot product is the magnitude of one vector times the portion of the vector that points in the same direction as that vector (the projection in the direction of the other.)



$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \, \|\mathbf{B}\| \cos\theta$$

3. Cauchy-Schwarz Inequality

What is the Cauchy-Schwarz Inequality?

- considered one of the most important and widely used inequalities in mathematics.
- states that for all vectors u and v of an inner product space it is true that:

$$AB \leq \|A\| \, \|B\|$$

- It can also be equivalently written as:

$$(AB)^2 \leq \|A\|^2 \, \|B\|^2$$

4. Linear Combinations

Definition

Let A_1, A_2, …, A_n be n matrices having dimension p x q.

A p x q matrix B is a **linear combination** of A_1, A_2, …, A_n if and only if there exist n scalars a_1, a_2, …, a_n called **coefficients** of the linear combination, such that

$$B = a_1A_1 + a_2A_2 + \ldots + a_nB_n$$

This definition can also be extended to vectors.

Example:

Let A_1 and A_2 be 2 x 2 matrices defined as follows:

$$A_1 = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}$$

Let $a_1 = 2$ and $a_2 = -1$ be two scalars. Then, the matrix
$B = a_1A_1 + a_2A_2$

Is a linear combination of A_1 and A_2.
It is computed as follows:

$$\begin{aligned} B &= a_1A_1 + a_2A_2 \\ &= 2\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} - 1\begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2(0) & 2(2) \\ 2(1) & 2(1) \end{bmatrix} + \begin{bmatrix} -1(3) & -1(0) \\ -1(1) & -1(1) \end{bmatrix} \\ &= \begin{bmatrix} 0-3 & 4-0 \\ 2-1 & 2-1 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 4 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Linear Combination of Vectors

- defined the same way.

Example:

Let A_1, A_2, and A_3 be 3 x 1 column vectors defined as follows:

$$A_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Let B be another 3 x 1 column vector defined as:

$$B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Is B a linear combination of A_1, A_2, and A_3?

Solution:

$$B \; ?= 1 \cdot A_1 + 1 \cdot A_2 + 0 \cdot A_3$$

$$= 1\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = B$$

Yes, B a linear combination of A_1, A_2, and A_3.

5. Linear Independence

Definition:

- two or more vectors are said to be linearly independent if none of them can be written as a linear combination of the others.
- if at least one of them can be written as a linear combination of the others, then they are said to be linearly dependent.

5.1. Linearly Dependent Vectors

Definition:

Let S be a linear space.

Some vectors $x_1, x_2, \dots, x_n \in S$ are said to be linearly dependent

if and only if

there exist n scalars $a_1, a_2, \dots, a_n$ such that:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$$

and at least one of the n scalars $a_1, a_2, \dots, a_n$ is different from zero.

One of the scalars must be different from zero?

- this requirement is fundamental.
- without this requirement the definition would be trivial: we could always choose

  $a_1 = a_2 = \ldots = a_n = 0$

  and obtain as a result

  $a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0x_1 + 0x_2 + \ldots + 0x_n = 0$

  for any set of n vectors.

- if one of the coefficients of the linear combination is different from zero (suppose, without loss of generality, it is a_1), then we can write:

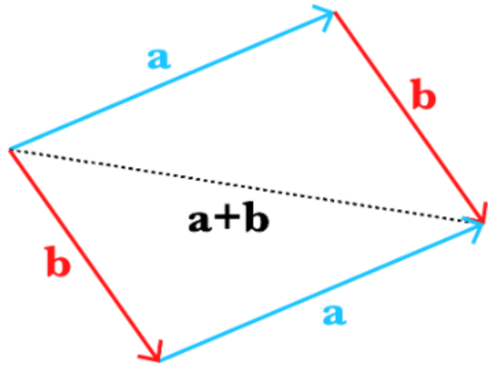
$$x_1 = -\frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_2 \ldots - \frac{a_n}{a_1}x_n$$

That is, x_1 is a linear combination of the vectors $x_2, x_3, \ldots, x_n$.

Geometric Interpretation of Linear Dependence of Vectors

Vector addition a+b

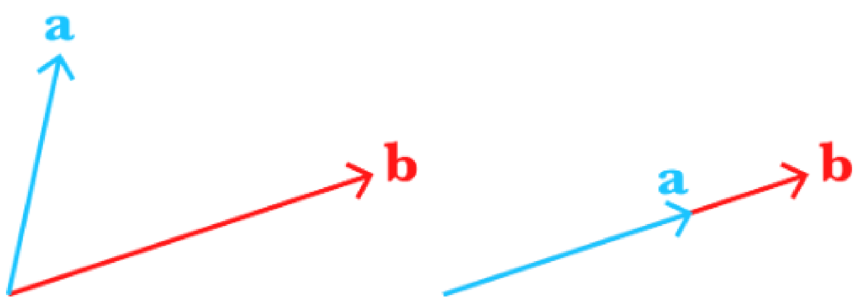
- first draw vector a from the origin
- then draw vector b from the head of a.



Vector addition **a**+**b**

Scale two 2-vectors such that they add to 0=[0,0]?

- they must lie on the same line.
- If they don't lie on the same line, there would be no way to scale and add the vectors such that the computation ends at the origin 0=[0,0].



two linear independent (left) and dependent (right) vectors in 2-dimensions

Example:

Let x_1 and x_2 be two 1-column vectors defined as follows:

$$x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad x_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

they are linearly dependent because: $2\,x_1 - x_2 = 0$

5.2. Linearly Independent Vectors

Definition:

Let S be a linear space.

Some vectors $x_1, x_2, \ldots, x_n \in S$ are said to be linearly independent if they are not linearly dependent.

In the case of linear independence:

$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$

implies that $a_1 = a_2 = \ldots = a_n = 0$

Example:

Let x_1 and x_2 be two 1-column vectors defined as follows:

$x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Are these vectors linearly independent?

Consider a linear combination of these two vectors with coefficients a_1 and a_2:

$a_1x_1 + a_2x_2 = 0$

This is equal to

$$\begin{aligned} a_1x_1 + a_2x_2 &= a_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \end{aligned}$$

Therefore, we obtain

$a_1x_1 + a_2x_2 = 0$ if and only if

$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

that is, if and only if $a_1 = a_2 = 0$.

As a consequence, the two vectors are linearly independent.

Exercises

Exercise 1:

Are the following two vectors linearly independent?

$$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Answer: ?

Exercise 2:

Are the following three vectors linearly independent?

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Answer: ?

5.3. Orthogonality of Vectors

Definitions:

Orthogonal Vectors:

Two 2 vectors are orthogonal if they are perpendicular to each other. i.e., the dot product of the two vectors is zero.

Orthonormal Vectors:

A set of vectors S is orthonormal if every vector in S has magnitude 1 and the set of vectors are mutually orthogonal.

Example:

Are the following vectors orthogonal to each other?

$$x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad x_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Their dot product is 1(2) + 2(4) = 10.

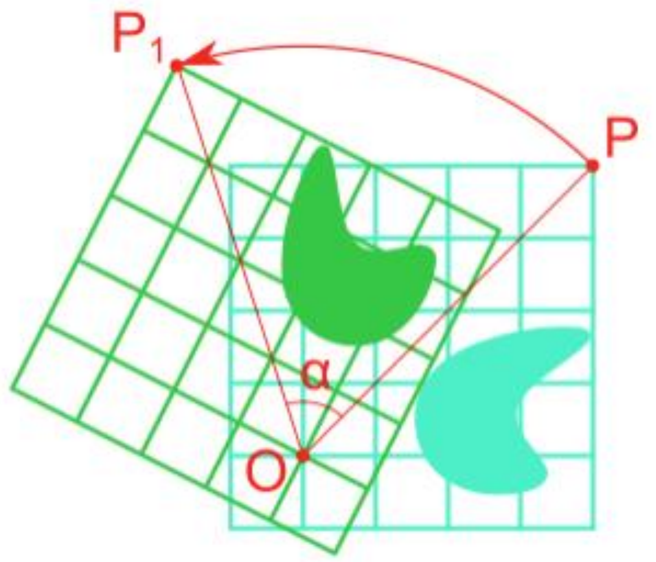
They are not orthogonal.

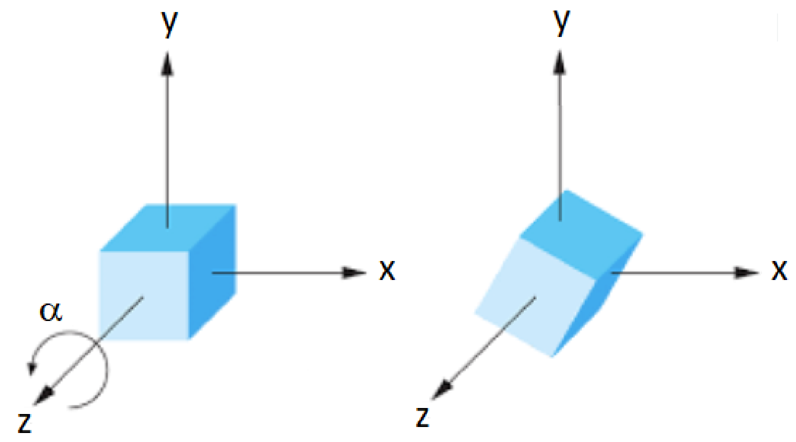
6. Rotation

What is Rotation?

- A mathematical transformation that takes each point in a space and rotates it a certain number of degrees around a given point.
- The result of a rotation is a new figure.
- The result is congruent to the original figure.
- The figure can rotate around any given point.
- Rotation degree:
    - positive → counterclockwise rotation.
    - negative → rotate clockwise rotation.

- Center of rotation can be inside or outside figure.
- Can be represented in 2D and 3D spaces.

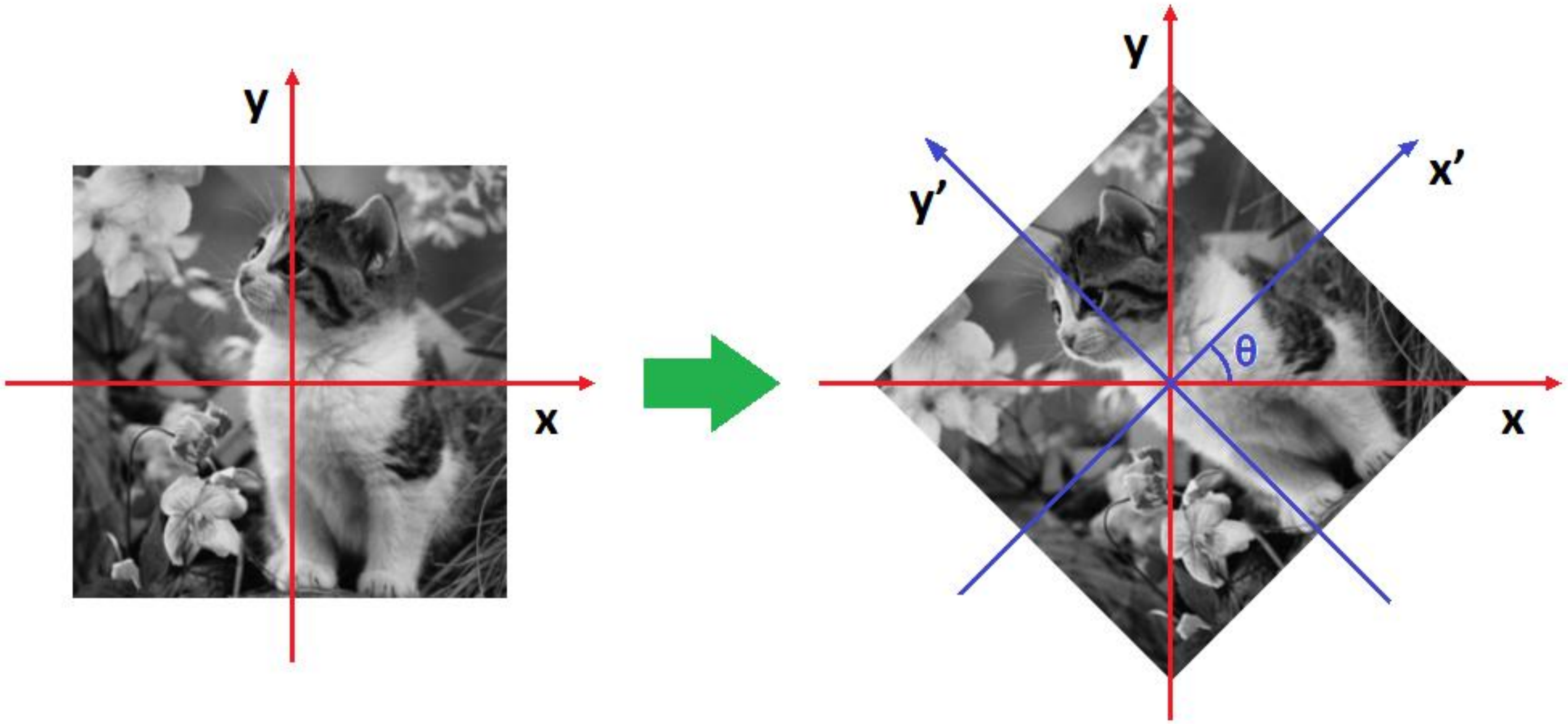




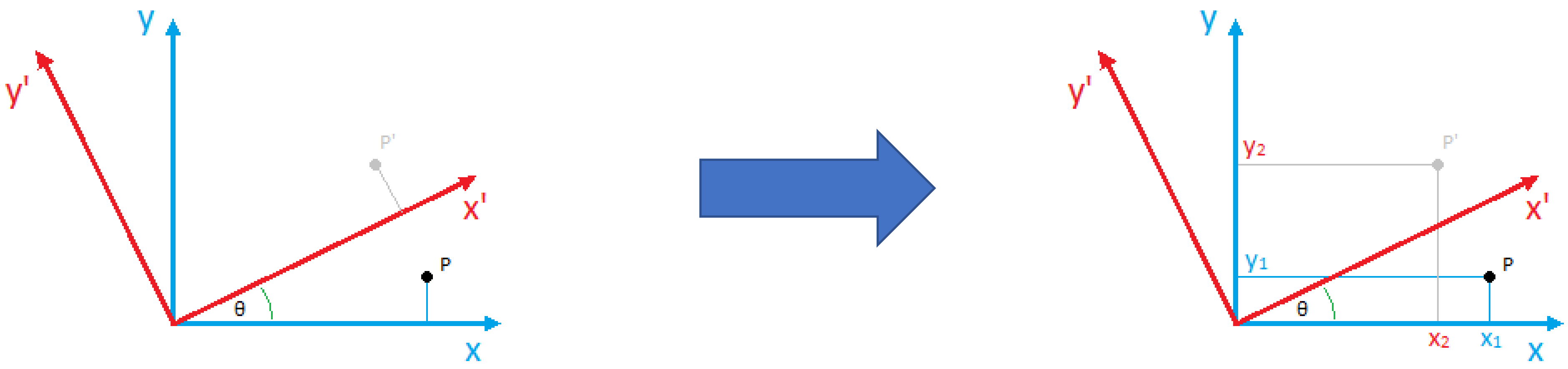
Rotating an Image

Rotating an image involve different steps:

- choosing center of rotation
- choosing angle of rotation
- translating the image if necessary
- performing rotation
- when reading raw image and displaying output image, reassign vertical image axis.



In two dimensions, only a single angle is needed to specify a rotation about the origin.

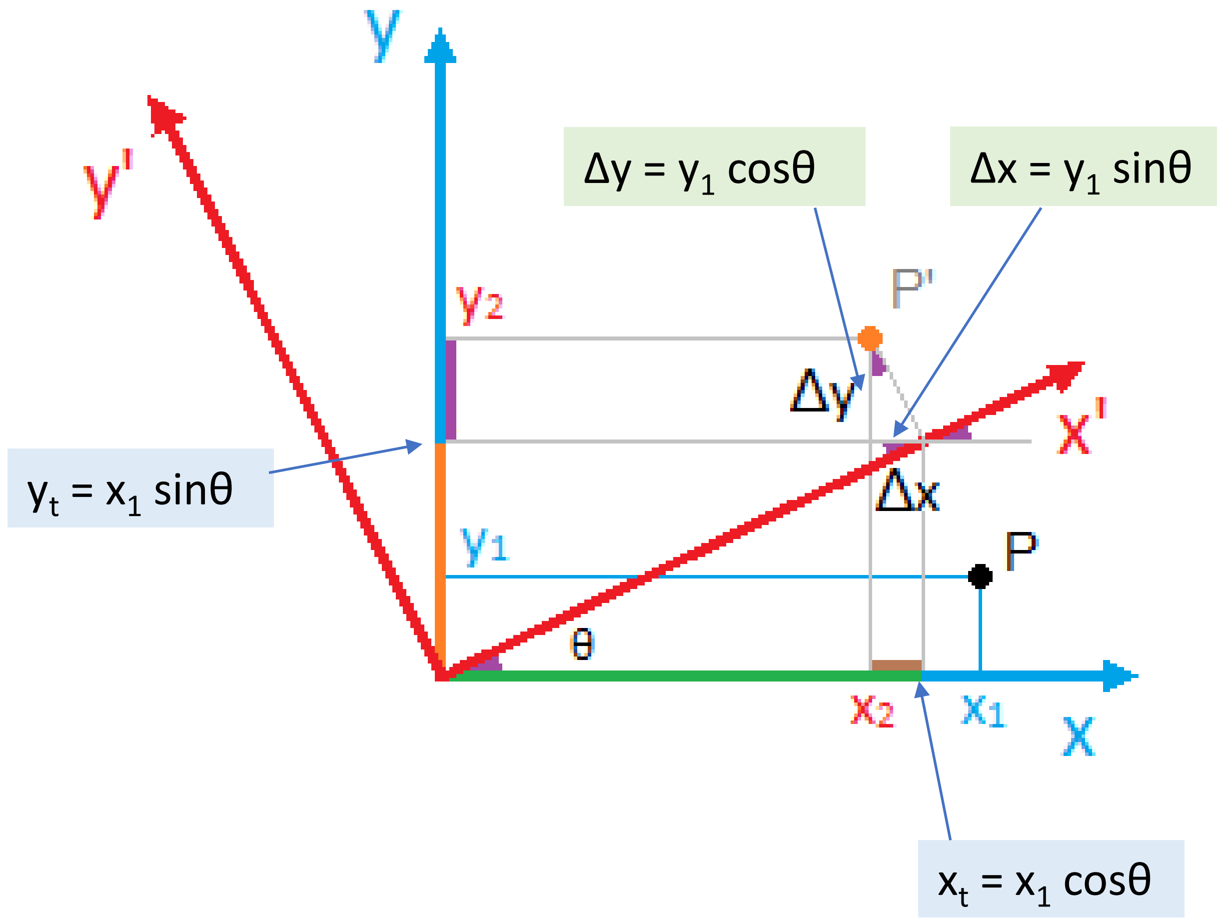


Approach:

An image is of size m by n is composed of (m*n) pixels.
We apply the rotation to each one of those pixels.
The new coordinates $P'(x_2, y_2)$ of a point $P(x_1, y_1)$ that has been rotated θ degrees around (0, 0) need to be calculated.

How do we get them?

How to find the new coordinates? (for the interested reader)



We start by finding x_2.

The orthogonal projection of x' = x_1 onto x is given by:
$x_t = x_1 \cos\theta$

but $x_2 = x_t - \Delta x$

We know that
$\Delta x = y_1 \sin\theta$
$\Delta y = y_1 \cos\theta$

therefore:
$x_2 = x_t - \Delta x = x_1 \cos\theta - y_1 \sin\theta$

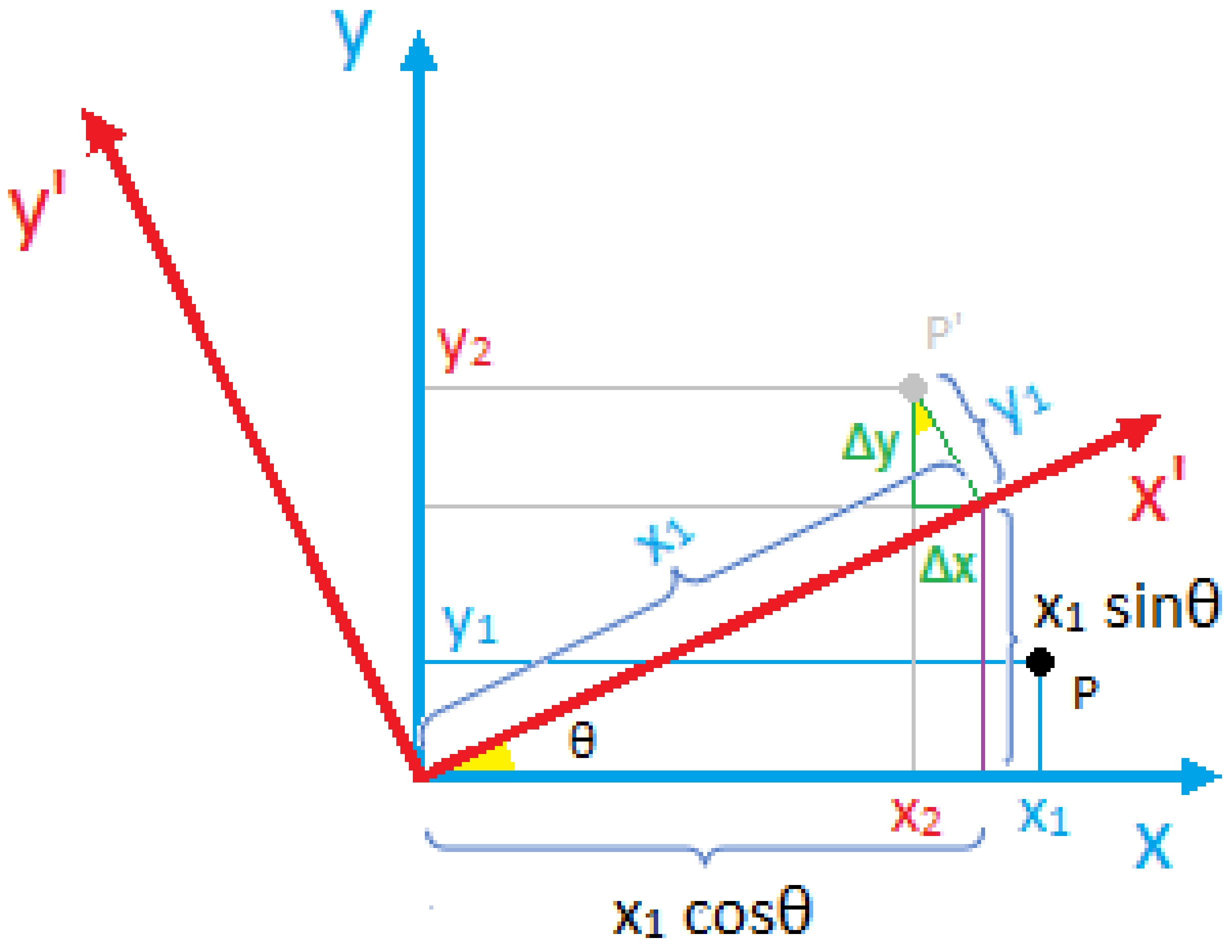
We apply the same reasoning to finding y_2.

The orthogonal projection of x' = x_1 onto y is given by:
$y_t = x_1 \sin\theta$.

but $y_2 = y_t + \Delta y$

therefore:
$y_2 = y_t + \Delta y = x_1 \sin\theta + y_1 \cos\theta$

Another geometric interpretation

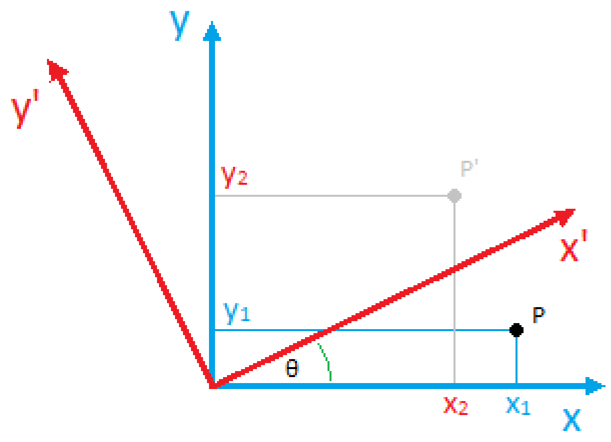


$x_2 = x_1 \cos\theta - \Delta x$
$y_2 = x_1 \sin\theta + \Delta y$

but
$\Delta x = y_1 \sin\theta$
$\Delta y = y_1 \cos\theta$

therefore
$x_2 = x_1 \cos\theta - \Delta x = x_1 \cos\theta - y_1 \sin\theta$
$y_2 = x_1 \sin\theta + \Delta y = x_1 \sin\theta + y_1 \cos\theta$

We ended up with a system of two equations:



$$x_2 = x_1 \cos\theta - y_1 \sin\theta$$
$$y_2 = x_1 \sin\theta + y_1 \cos\theta$$

which represents the new coordinates $P'(x_2, y_2)$ of a point $P(x_1, y_1)$ that has been rotated θ degrees around (0, 0).

This linear system can be represented as a matrix multiplication tableau:

$$\begin{array}{cc|c} & & x_1 \\ & & y_1 \\ \hline \cos\theta & -\sin\theta & x_2 \\ \sin\theta & \cos\theta & y_2 \end{array}$$

Two-dimensional Rotation Matrix:

- transformation matrix used to perform a rotation in Euclidean space. For example, using the convention below, the matrix.
- to rotate points in the xy plane counterclockwise through an angle θ with respect to the positive x axis about the origin of a two-dimensional Cartesian coordinate system, use the following matrix:

$$R(\theta) = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

- to perform the rotation on a plane point with standard coordinates v = (x, y), it should be written as a column vector, and multiplied by the matrix R:

$$R(x, y, \theta) = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xcos\theta & -ysin\theta \\ xsin\theta & ycos\theta \end{bmatrix}$$

- rotation matrices describe rotations about the origin.

Considerations when rotating images:

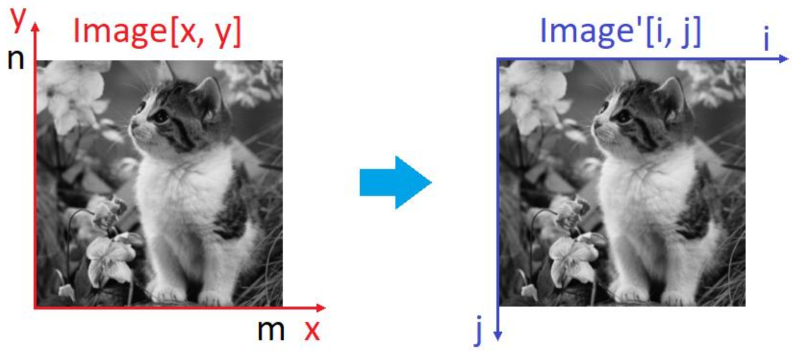
- Image y axis goes down

- A positive rotation (angle) rotates counterclockwise.

- For rotations around another point other than (0,0), replace
- x → (x – xc)
- Y → (y – yc)
- (e.g., the center of the image)

- can produce coordinates which are not integers. To obtain integer positions, results need to be rounded to integer values (row, column). Rounding might leave pixels orphan. To prevent it from happening, rotate in inverse mode.

m cols by n rows



i = x

j = n - i

Image[2, 0] = Image’[2, n]

Image[2, 1] = Image’[2, n-1]

Image[2, 2] = Image’[2, n-2]

...

Considerations when rotating images (Cont.)

- Rounding might produce orphan pixels.

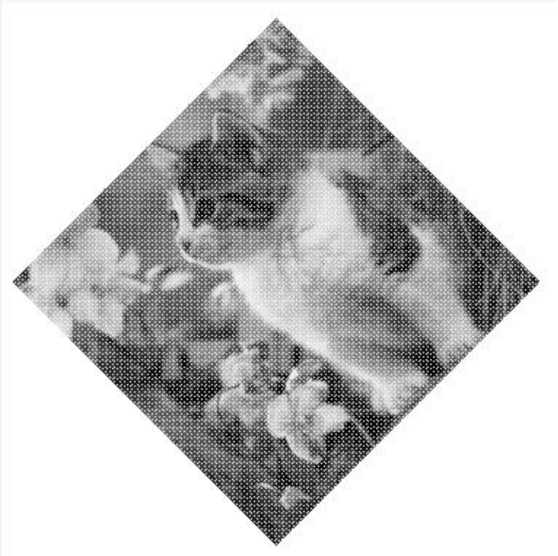
angle = 30°		
		5
		7
0.866	-0.5	0.8
0.5	0.866	8.6

input: original coordinate [i,j] (always integers)

output: coordinate after rotation [i_{new},j_{new}] (not integers)

- To prevent it from happening, rotate in inverse mode.

Image to be rotated =45°



Orphan pixels generated by rounding

Rotation in inverse mode. How?

7. Cosine Similarity

Definition:

- measures the similarity between two vectors of an inner product space.
- measured by the cosine of the angle between two vectors
- determines whether two vectors are pointing in roughly the same direction.
- often used to measure document similarity in text analysis.

End of Lecture