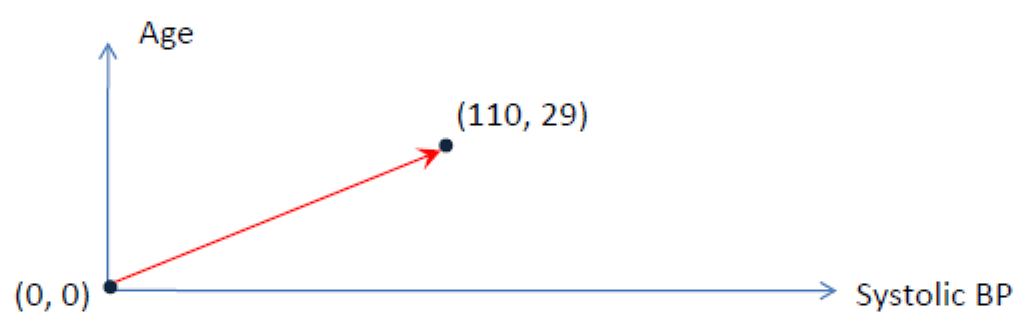
Understanding Vectors

How to represent samples geometrically? Vectors in n-dimensional space ($\mathbb{R}^n$)

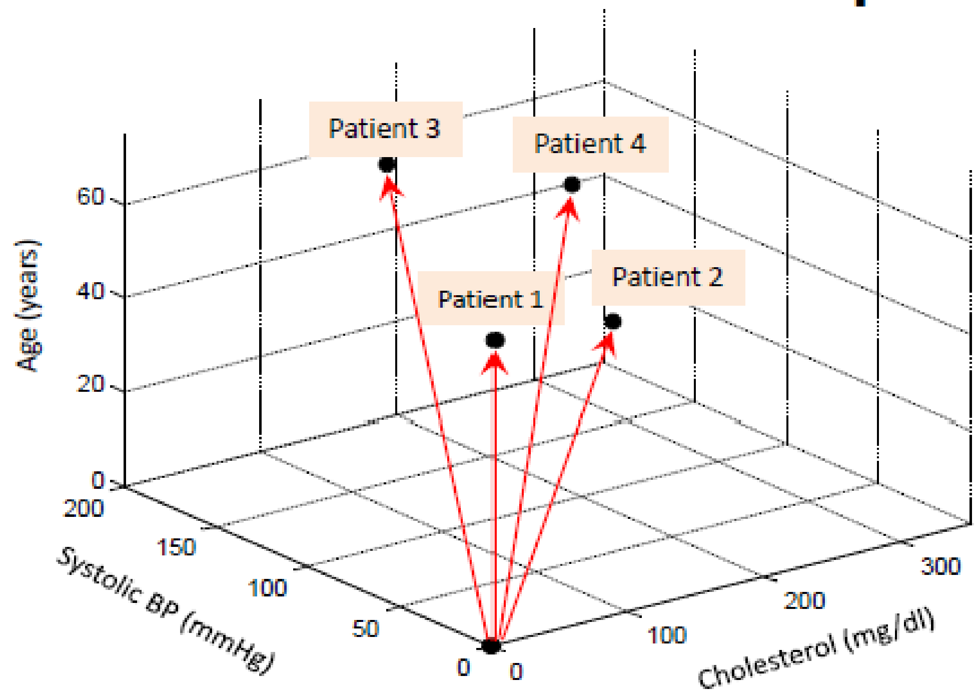
- Assume that a sample/patient is described by *n* characteristics ("features" or "variables")
- **Representation:** Every sample/patient is a vector in $\mathbb{R}^n$ with tail at point with 0 coordinates and arrow-head at point with the feature values.
- **Example:** Consider a patient described by 2 features:

  Systolic BP = 110 and *Age* = 29.

  This patient can be represented as a vector in $\mathbb{R}^2$:

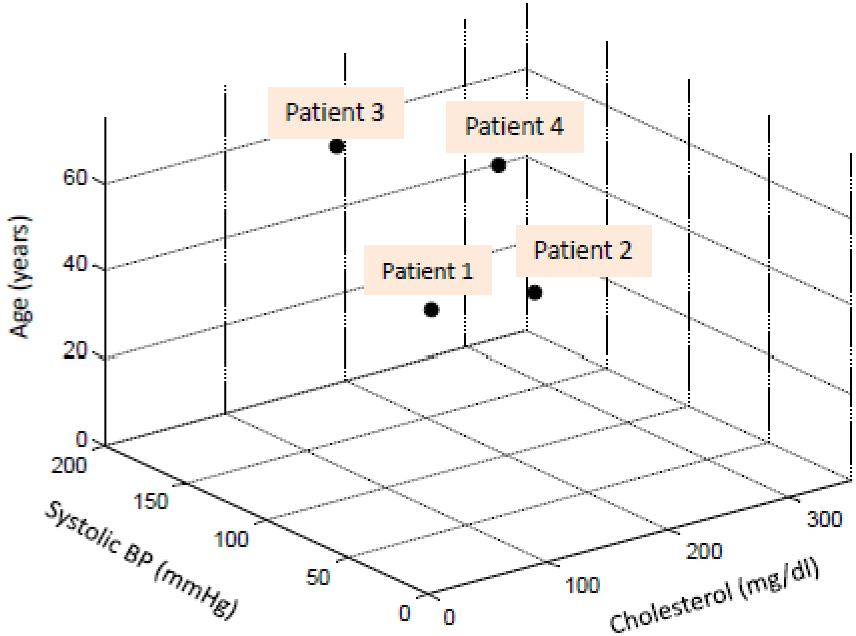


How to represent samples geometrically? Vectors in n-dimensional space ($\mathbb{R}^n$)



Patient id	Cholesterol (mg/dl)	Systolic BP (mmHg)	Age (years)	Tail of the vector	Arrow-head of the vector
1	150	110	35	(0,0,0)	(150, 110, 35)
2	250	120	30	(0,0,0)	(250, 120, 30)
3	140	160	65	(0,0,0)	(140, 160, 65)
4	300	180	45	(0,0,0)	(300, 180, 45)

How to represent samples geometrically? Vectors in n-dimensional space ($\mathbb{R}^n$)



Since we assume that the tail of each vector is at point with 0 coordinates, we will also depict vectors as points (where the arrow-head is pointing).

Basic operation on vectors in $\mathbb{R}^n$

1. Multiplication by a scalar

Consider a vector $\vec{a} = (a_1, a_2, \ldots, a_n)$ and a scalar c

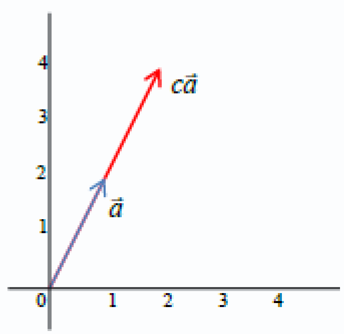
Define: $c\vec{a} = (ca_1, ca_2, \ldots, ca_n)$

When you multiply a vector by a scalar, you "stretch" it in the same or opposite direction depending on whether the scalar is positive or negative.

$\vec{a} = (1,2)$

$c = 2$

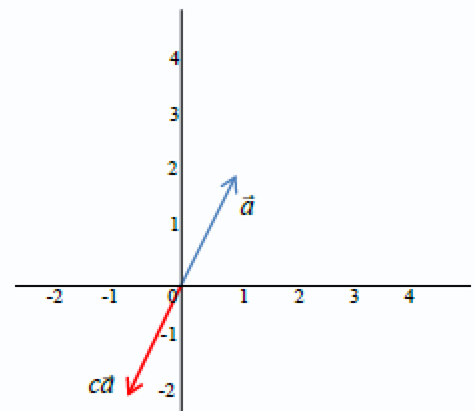
$\vec{c}a = (2,4)$



$\vec{a} = (1,2)$

$c = -1$

$\vec{c}a = (-1,-2)$



Basic operation on vectors in $\mathbb{R}^n$

2. Addition

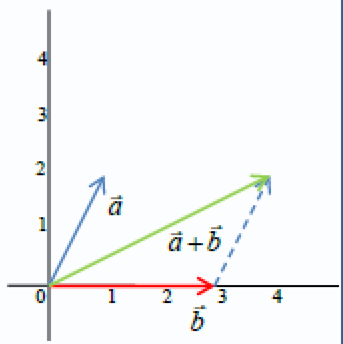
Consider vectors $\vec{a} = (a_1, a_2, \ldots, a_n)$ and $\vec{b} = (b_1, b_2, \ldots, b_n)$

Define: $\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$

$\vec{a} = (1,2)$

$\vec{b} = (3,0)$

$\vec{a} + \vec{b} = (4,2)$



Recall addition of forces in classical mechanics.

Basic operation on vectors in $\mathbb{R}^n$

3. Subtraction

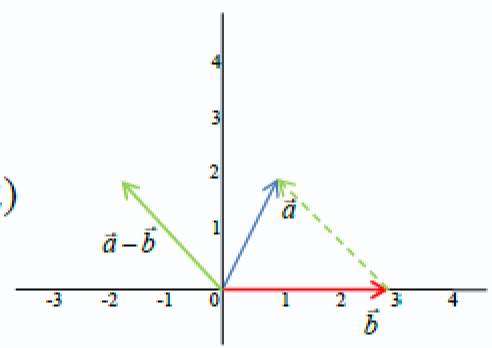
Consider vectors $\vec{a} = (a_1, a_2, \ldots, a_n)$ and $\vec{b} = (b_1, b_2, \ldots, b_n)$

Define: $\vec{a} - \vec{b} = (a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n)$

$\vec{a} = (1,2)$

$\vec{b} = (3,0)$

$\vec{a} - \vec{b} = (-2,2)$



What vector do we need to add to $\vec{b}$ to get $\vec{a}$? I.e., similar to subtraction of real numbers.

Basic operation on vectors in $\mathbb{R}^n$

4. Euclidian length or L2-norm

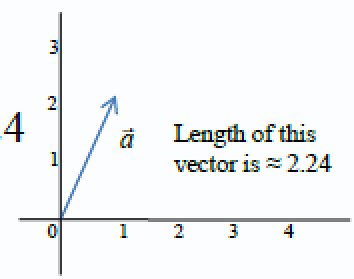
Consider a vector $\vec{a} = (a_1, a_2, \ldots, a_n)$

Define the L2-norm: $\|\vec{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}$

We often denote the L2-norm without subscript, i.e. $\|\vec{a}\|$

$\vec{a} = (1,2)$

$\|\vec{a}\|_2 = \sqrt{5} \approx 2.24$



L2-norm is a typical way to measure length of a vector; other methods to measure length also exist.

Unit Vector

Given the following vector: $\mathbf{q} = < -2, 1 >$,

Find the unit vector **u** of the vector q.

First calculate its magnitude |**q**|:

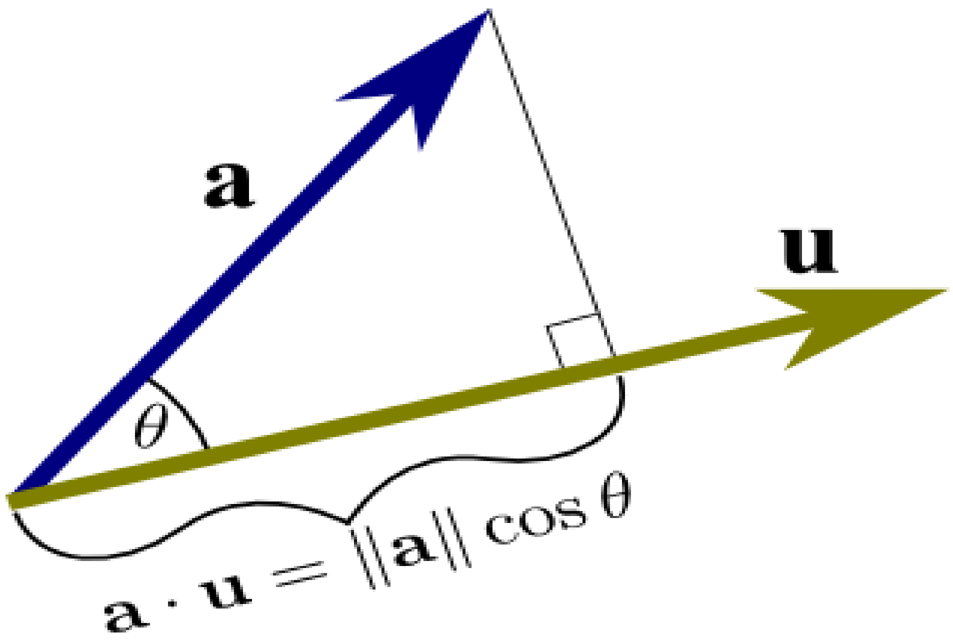
$$|\boldsymbol{q}| = \sqrt{(-2)^2 + 1^2} = \sqrt{5}$$

Now calculate the unit vector of q:

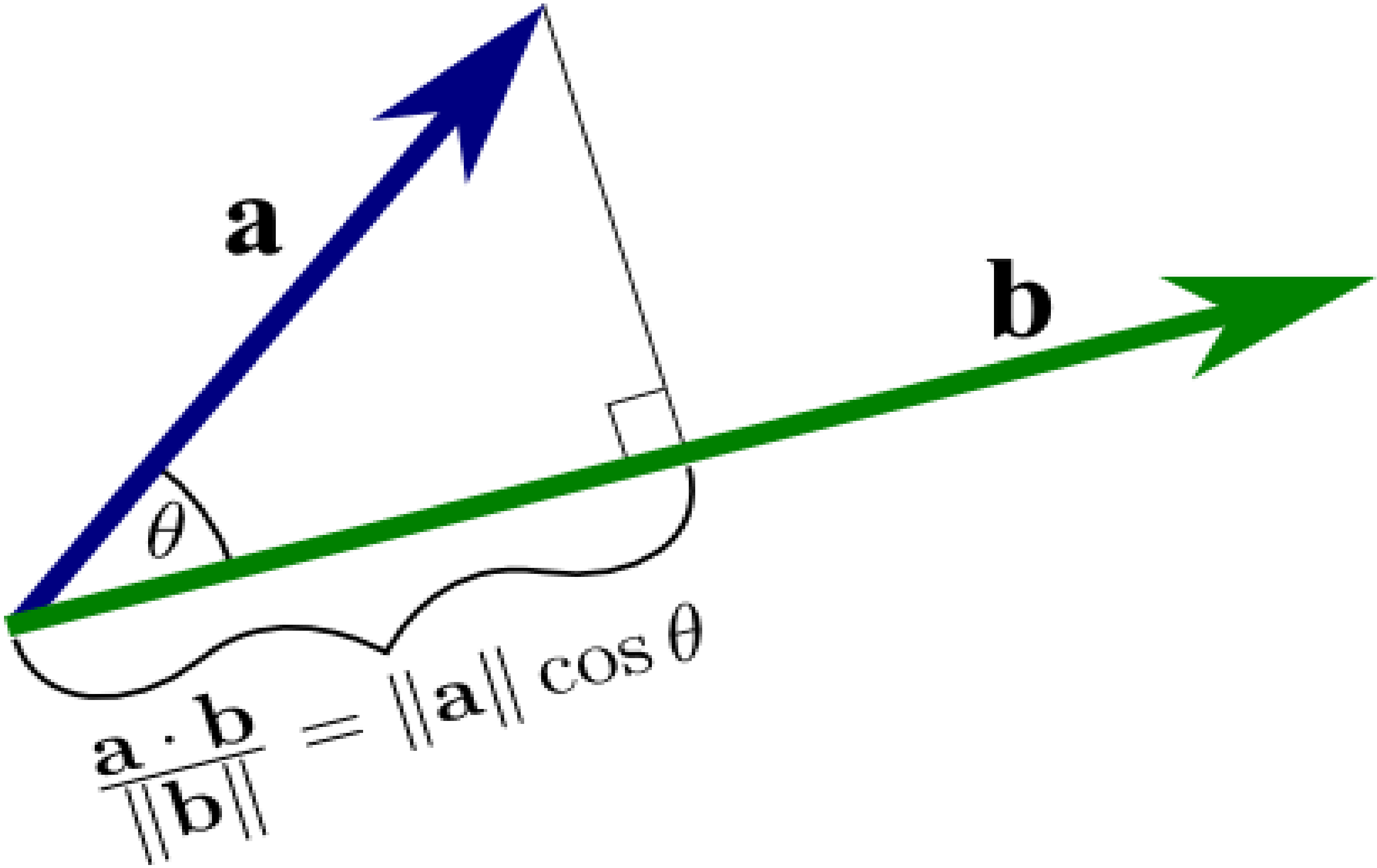
$$\boldsymbol{u} = \frac{< -2, 1 >}{\sqrt{5}} = < -\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} >$$

Dot Product

The dot product of a with unit vector u, denoted as a.u, is defined to be the projection of a in the direction of u:



Dot Product of two vectors



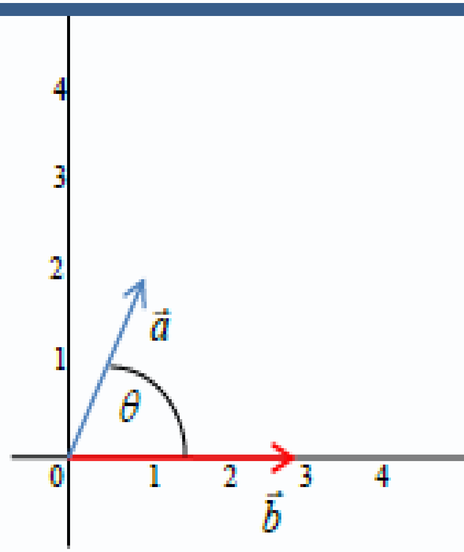
Another perspective of dot product

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

$$\vec{a} = (1,2)$$

$$\vec{b} = (3,0)$$

$$\vec{a} \cdot \vec{b} = 3$$



$$\vec{a} = (0,2)$$

$$\vec{b} = (3,0)$$

$$\vec{a} \cdot \vec{b} = 0$$