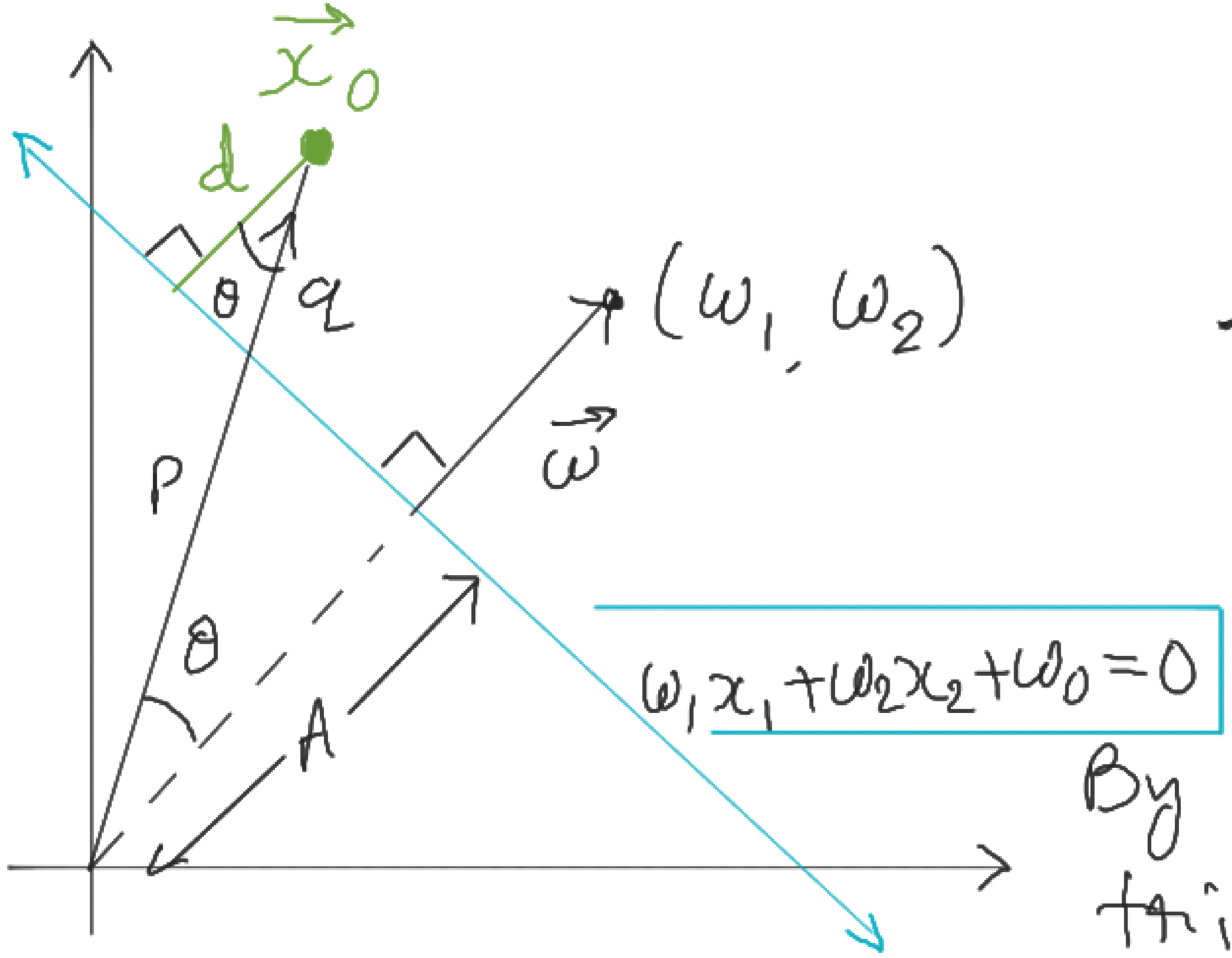
☆ Another formula of distance of a point from a line



$$\|\vec{x}_0\| = p + q$$

$$\therefore \quad q = \|\vec{x}_0\| - p$$

$$\therefore \quad \boxed{q = \|\vec{x}_0\| - \frac{A}{\cos\theta}} \text{ —— } (\mathrm{I})$$

By applying trigonometry on upper triangle, $\cos\theta = \dfrac{d}{q} \Rightarrow d = q\cos\theta$

$$\cos\theta = \frac{A}{p}$$

$$p = \frac{A}{\cos\theta}$$

Substituting q from (I),

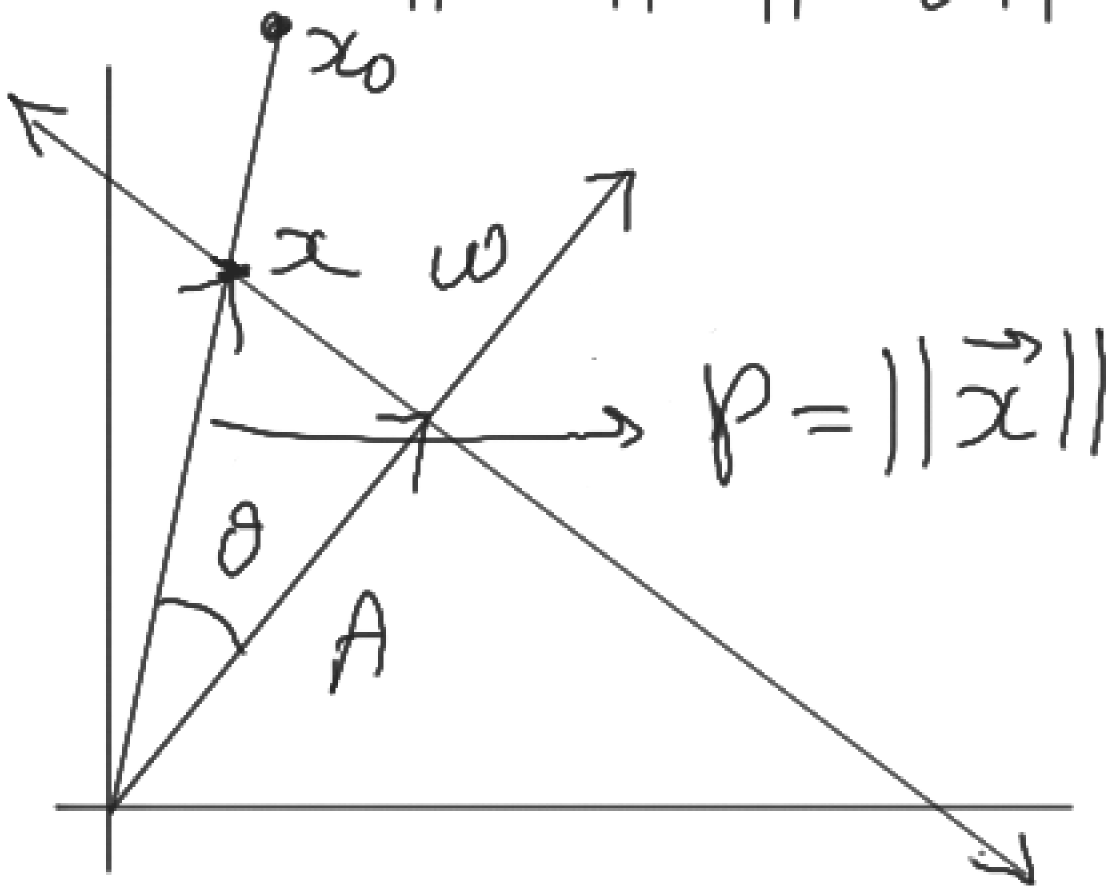
$$d = \left(\|\vec{x}_0\| - \frac{A}{\cos\theta}\right)\cos\theta = \|\vec{x}_0\|\cos\theta - A$$

→ As $\vec{d}$ & $\vec{w}$ shares the same direction, θ can be also considered as an angle between $\vec{x_0}$ & $\vec{w}$ ∴ From the formula of dot product ⇒

$$\cos\theta = \frac{\vec{w}^T \cdot \vec{x_0}}{\|\vec{w}\| \cdot \|\vec{x_0}\|}$$

Putting this in eq^n of 'd'

$$d = \frac{\cancel{\|\vec{x_0}\|}\ \vec{w}^T \cdot \vec{x_0}}{\|\vec{w}\| \cdot \cancel{\|\vec{x_0}\|}} - A \quad \text{——— } \mathrm{II}$$



From this figure,

$$\cos\theta = \frac{A}{\|\vec{x}\|}$$

$$A = \|\vec{x}\| \cos\theta$$

$$\therefore A = \cancel{\|\vec{x}\|} \cdot \frac{\vec{w}^T \cdot \vec{x}}{\|\vec{w}\| \cdot \cancel{\|\vec{x}\|}}$$

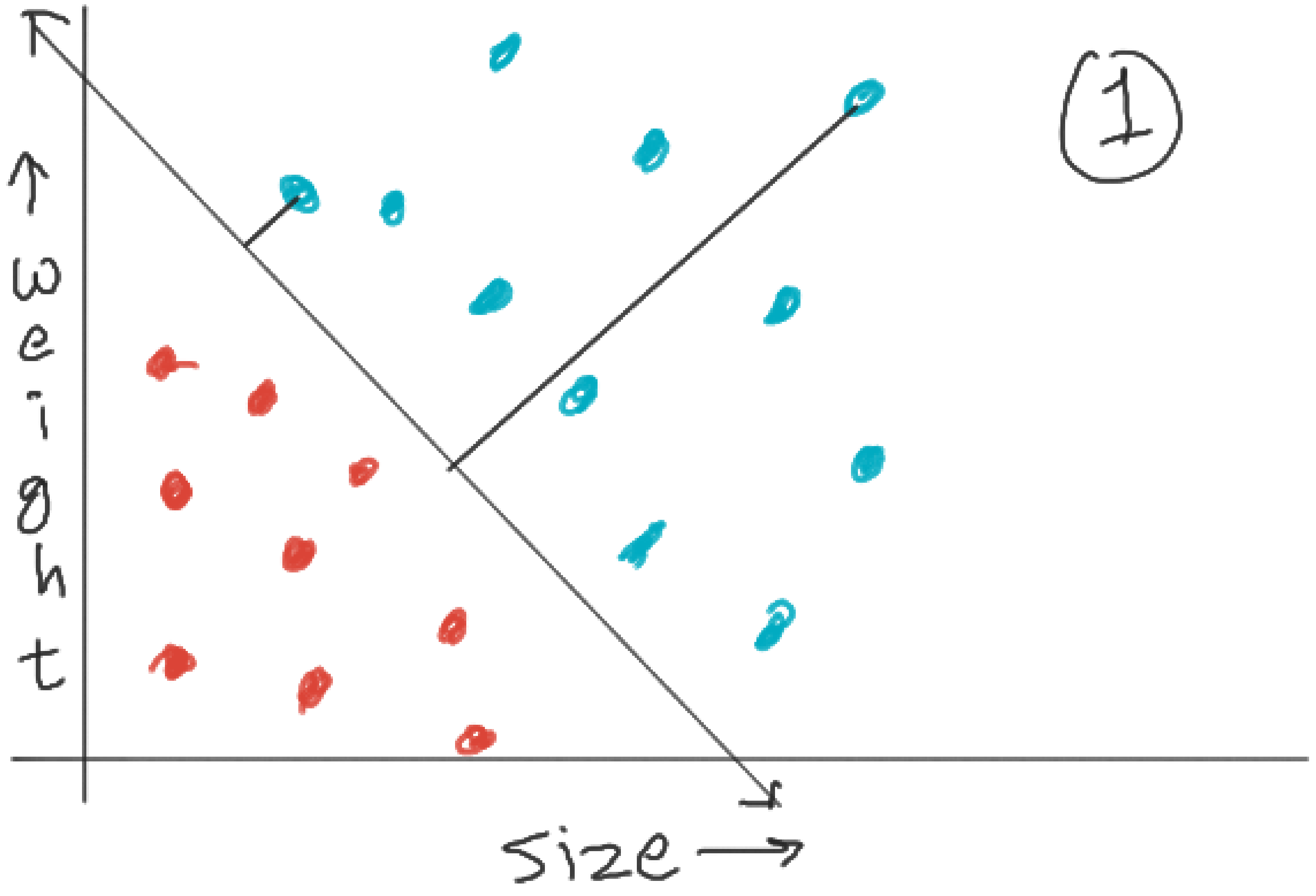
$$\therefore A = \frac{-w_0}{\|\vec{w}\|}$$

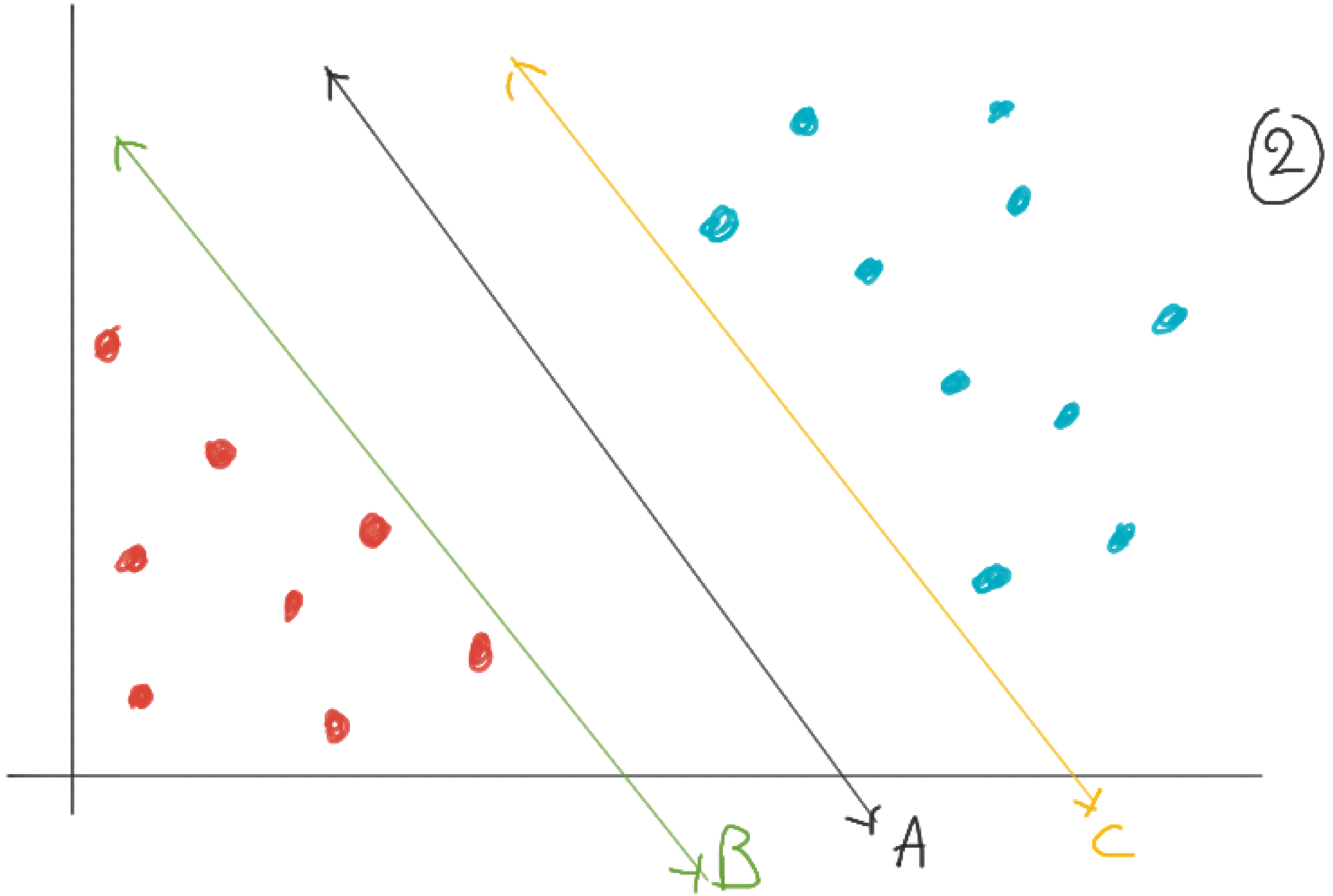
→ Putting this value of 'A' into II,

$$d = \frac{\vec{w}^T \vec{x_0}}{\|\vec{w}\|} + \frac{w_0}{\|\vec{w}\|}$$

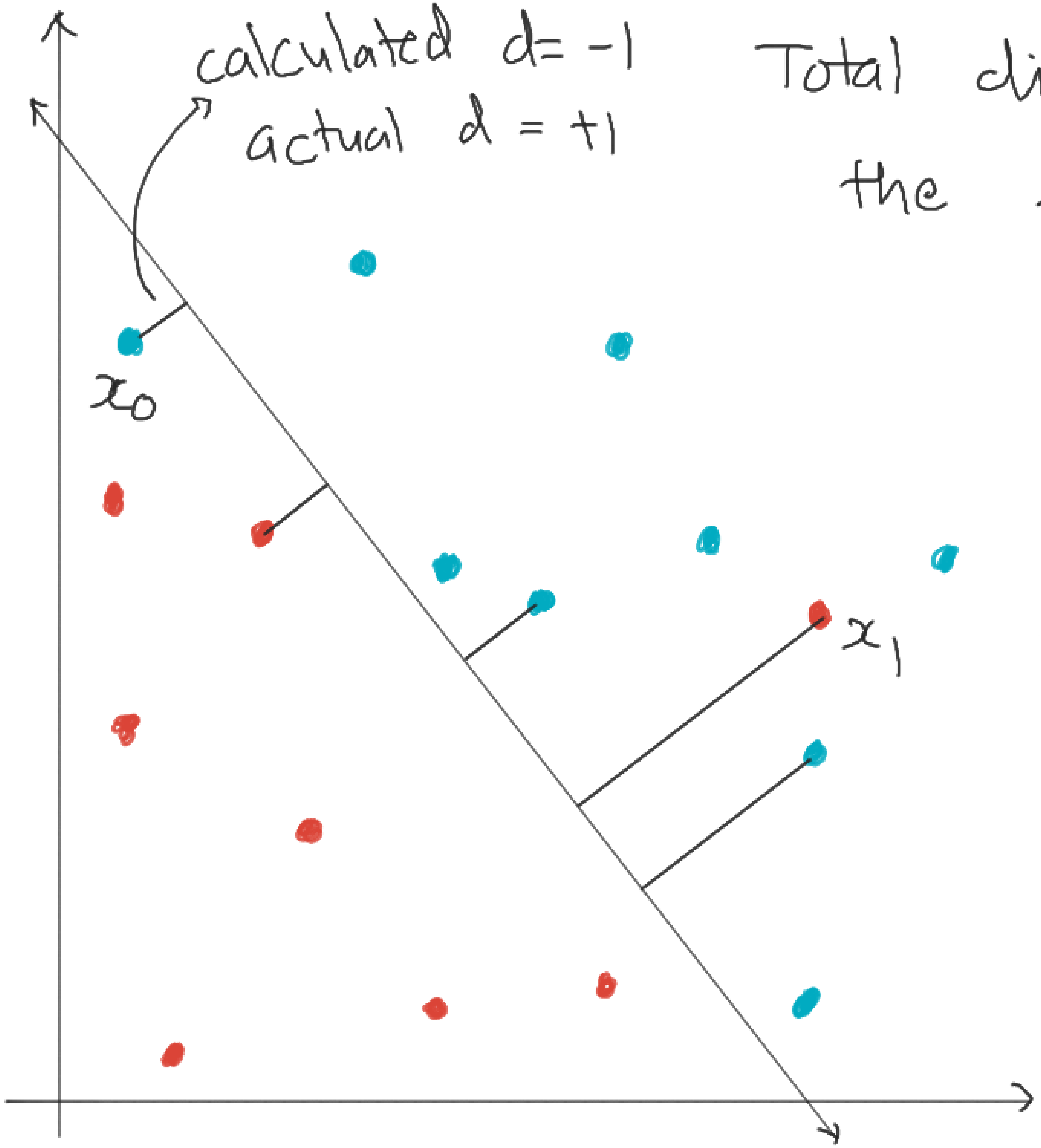
$$\boxed{d = \frac{\vec{w}^T . \vec{x_0} + w_0}{\|\vec{w}\|}}$$ — V.V. Imp. Formula

☆ Loss Minimization





Our goal is to find a classifier whose distance from every point is maximum. In other words, we want to maximize this distance.



Total distance of all the points from the line can be given as

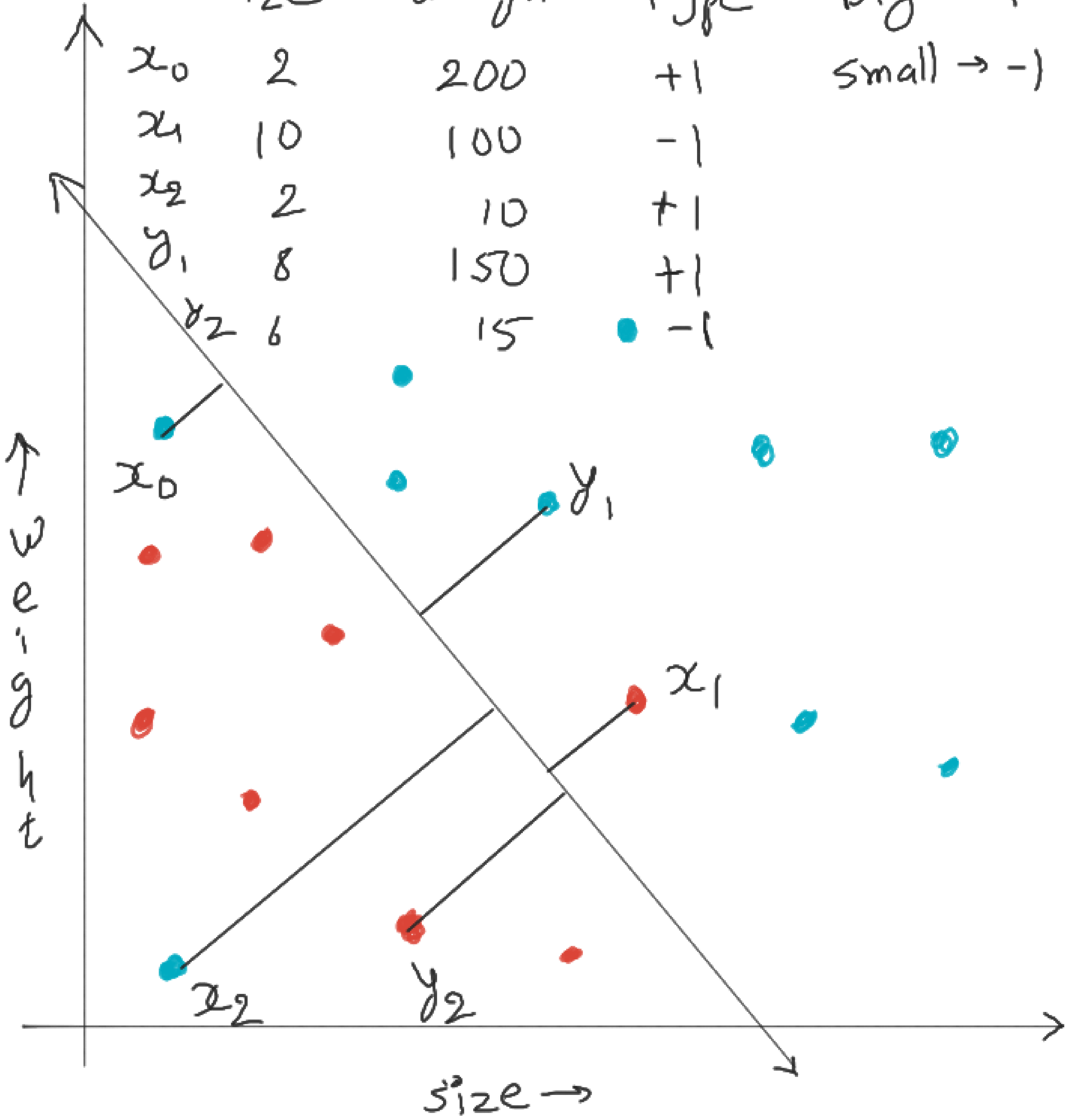
$$\sum_{i=0}^{n} \frac{\vec{\omega}^T . \vec{x}_i + w_0}{\|\vec{\omega}\|}$$

There are two problems with this logic:

1. The positive and negative distances may cancel out each other.
2. We are not penalizing mis-classified points.

	size	weight	type
x_0	2	200	+1
x_1	10	100	-1
x_2	2	10	+1
y_1	8	150	+1
y_2	6	15	-1

Big → +1
small → -1

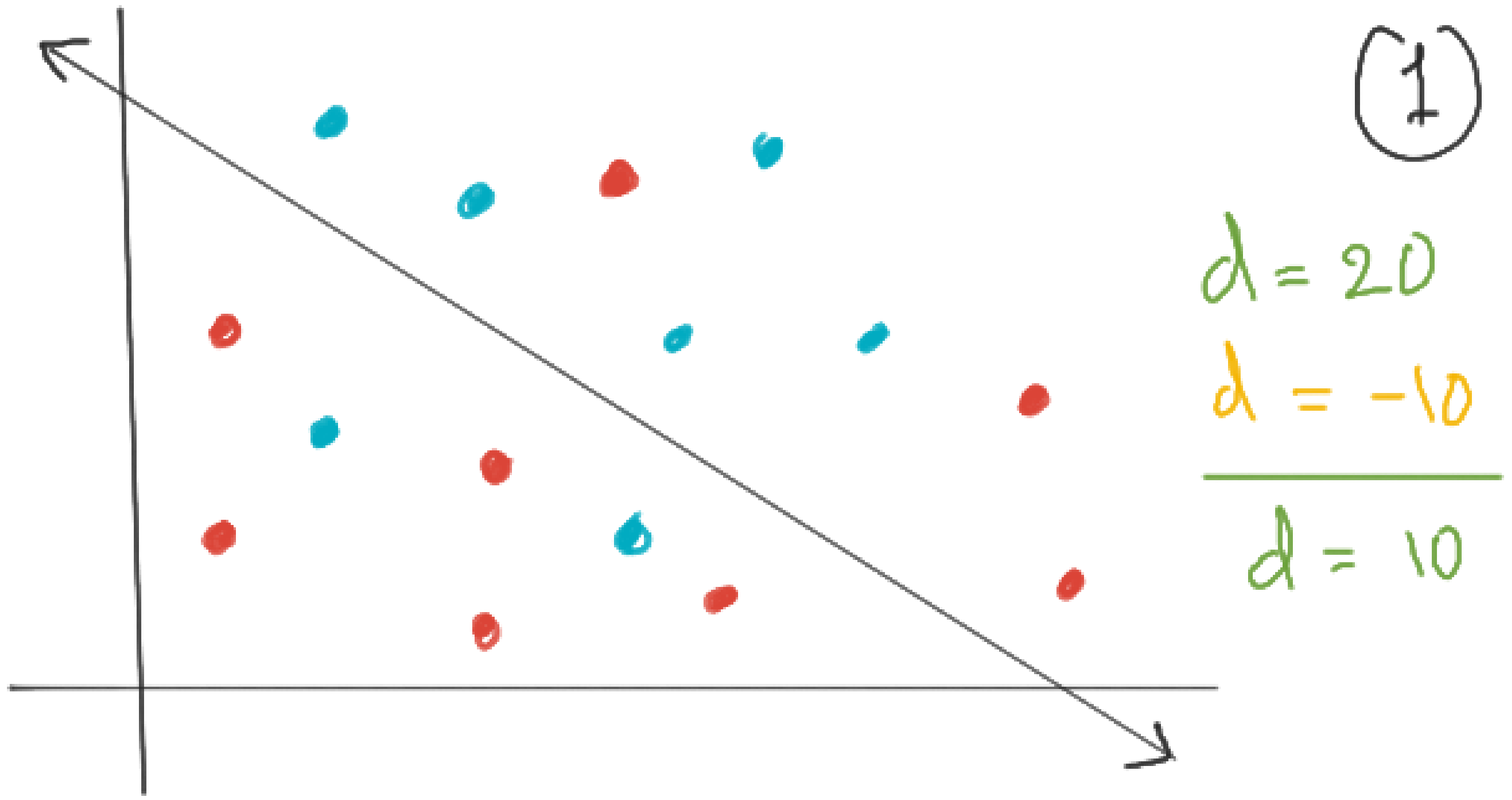


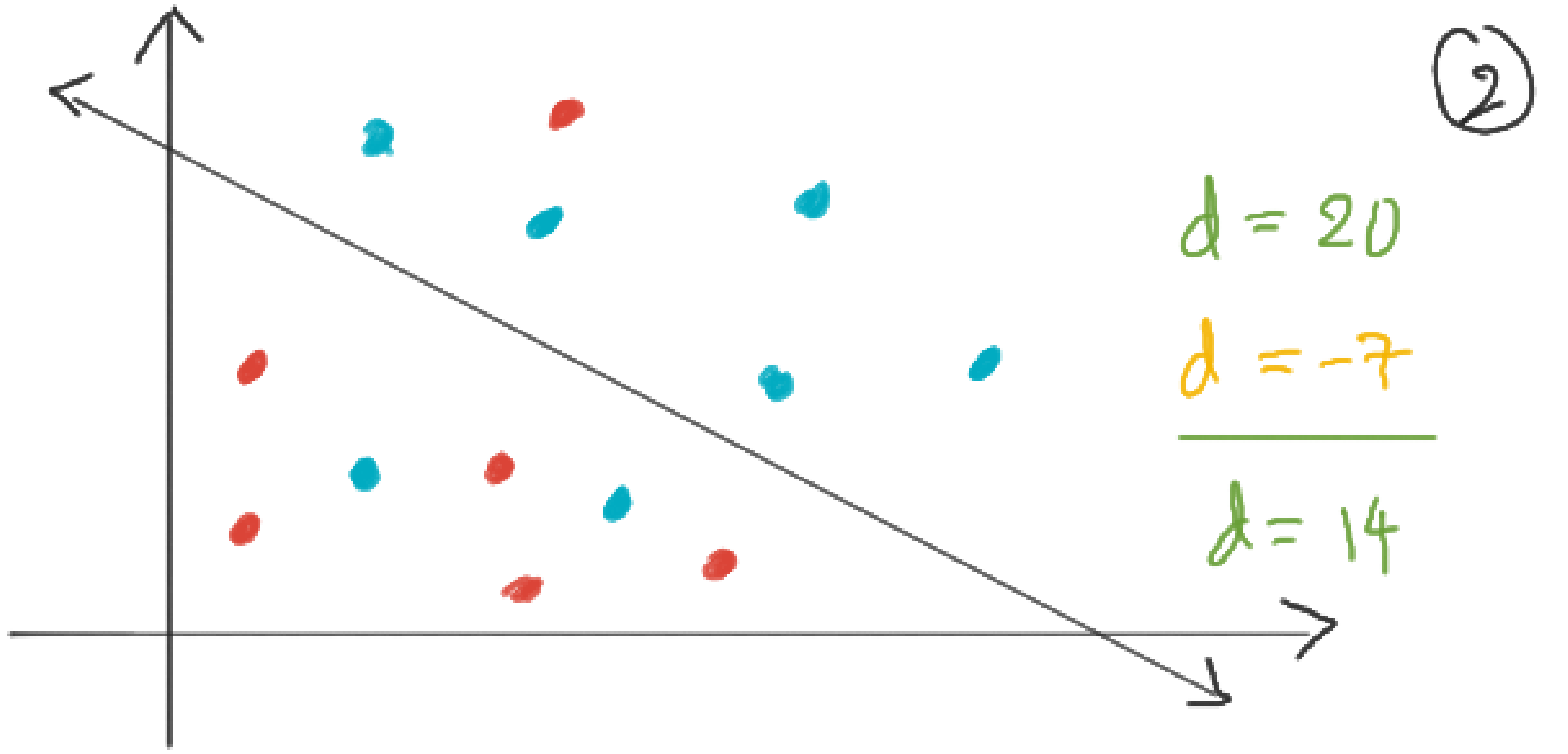
☆ Distances (calculated) of these points are:

point	d_{calc}
x_0	-1
x_1	2
x_2	-5
y_1	3
y_2	-3

☆ what if I multiply label of each point with d_{calc} ?

Point	d_{calc}	label	d_{calc} * label
x_0	-1	+1	-1
x_1	2	-1	-2
x_2	-5	+1	-5
y_1	3	+1	3
y_2	-3	-1	3





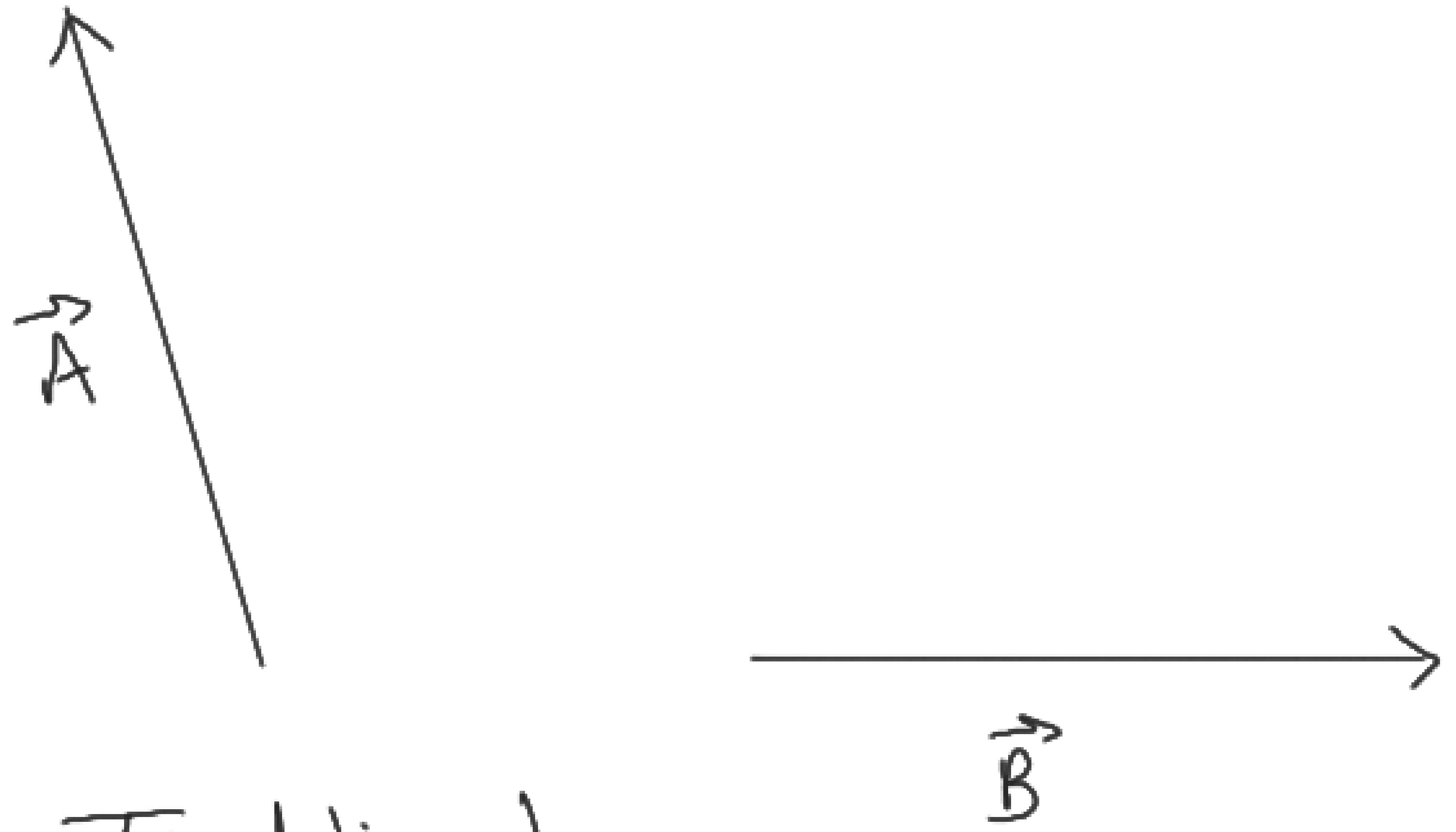
☆ Updated formula of Gain Function

$$\boxed{\sigma(\vec{x}, \vec{\omega}, \omega_0, \vec{y}) = \sum_{i=1}^{n} \left(\frac{\vec{\omega}^T \cdot \vec{x}_i + \omega_0}{\|\vec{\omega}\|} \right) \cdot y_i}$$

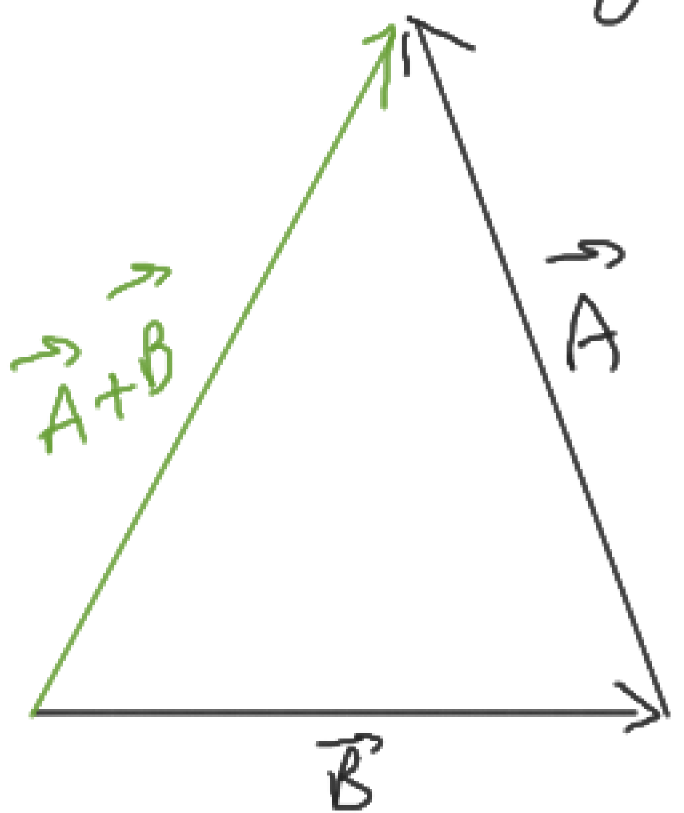
☆ Loss function :

$$\boxed{L(\vec{x}, \vec{\omega}, \omega_0, \vec{y}) = -\sigma(\vec{x}, \vec{\omega}, \omega_0, \vec{y})}$$

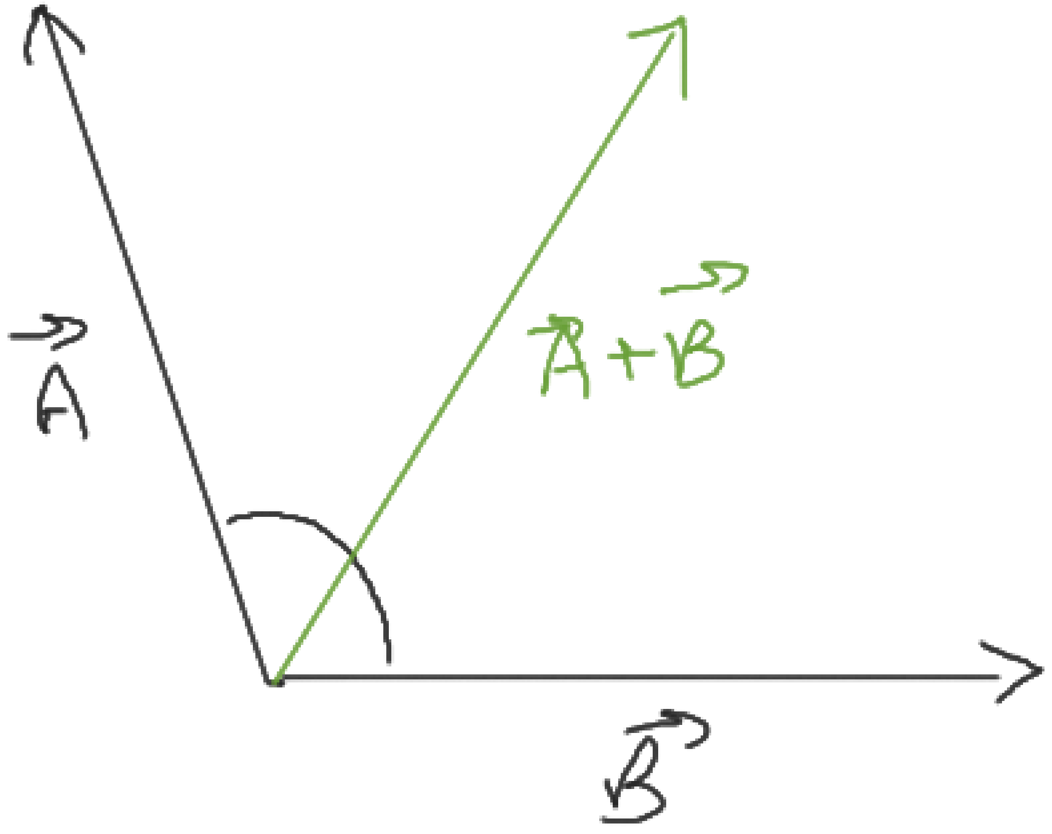
☆ Another approach to visualize vector addition:



→ Traditional way:



→ Alternate way:



☆ Logic of perceptron algorithm

label of blue points is +1

label of red points is -1