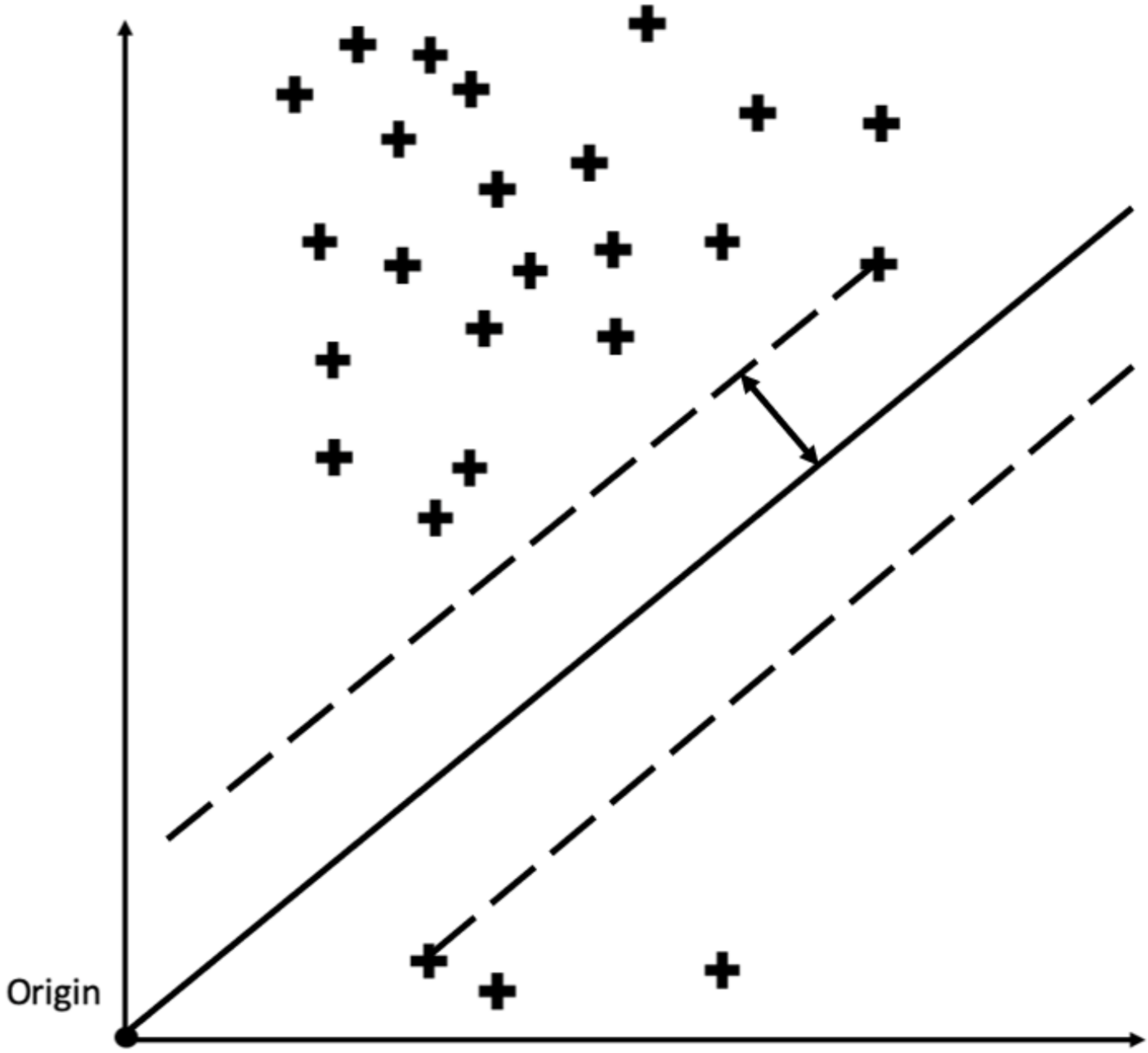
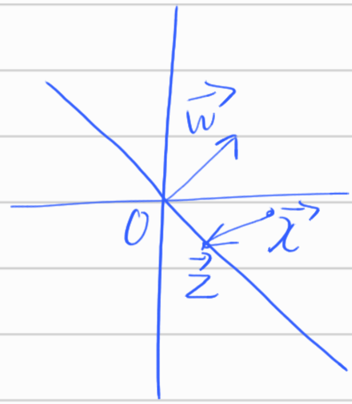
a)



This decision boundary maximize the margin.

b) $\vec{z}$ is an arbitrary point on hyperplane.
The distance to hyperplan from $\vec{x}$ is the projection of $\vec{x}-\vec{z}$ to $\vec{w}$



$$\text{length} = \left|\left\langle \vec{x}-\vec{z}, \frac{\vec{w}}{\|\vec{w}\|}\right\rangle\right| = \frac{1}{\|\vec{w}\|_2}\left|\langle \vec{x}-\vec{z}, \vec{w}\rangle\right|$$
$$= \frac{1}{\|\vec{w}\|_2}\left|\vec{w}^T\vec{x} - \vec{w}^T\vec{z}\right|$$

$\vec{z}$ is on the hyperplane, so $\vec{w}^T\vec{z} = 0$

So distance $= \dfrac{|\vec{w}^T\vec{x}|}{\|\vec{w}\|_2}$

margin is distance from the closest points to decision boundary to the decision boundary, so

$$m = \min_i \frac{|\vec{w}^T x_i|}{\|\vec{w}\|_2}$$

c)

$$\max_{m, \vec{w}} \; m$$

subject to

$$m \geqslant 0$$

$$\frac{|\vec{w}^T x_i|}{\|\vec{w}\|_2} \geq m \quad \forall i$$

$$\vec{w}^T \vec{x}_i > 0 \quad \forall i \text{ such that } y_i = 1$$

and $$m = \min_i \frac{|\vec{w}^T \vec{x}_i|}{\|\vec{w}\|_2}$$

d) $$\max_{m,\vec{w}} m$$

subject to

$m \geq 0$

$\cancel{\frac{|\vec{w}^T x_i|}{\|\vec{w}\|_2} \geq m \quad \forall i}$ since $m = \min_i \frac{|\vec{w}^T \vec{x}_i|}{\|\vec{w}\|_2}$ is the minimum value

$\vec{w}^T \vec{x}_i > 0 \quad \forall i$ such that $y_i = 1$

and $m = \min_i \frac{|\vec{w}^T \vec{x}_i|}{\|\vec{w}\|_2} \Rightarrow \frac{|\vec{w}^T \vec{x}_i|}{\|\vec{w}\|_2} \geq m \quad \forall i$

$$\max_{m,\vec{w}} m$$

subject to

$m \geq 0$

$\vec{w}^T \vec{x}_i > 0 \quad \forall i$ such that $y_i = 1$

$\frac{|\vec{w}^T x_i|}{\|\vec{w}\|_2} \geq m \quad \forall i \Rightarrow \vec{w}^T \vec{x}_i \geq m\|\vec{w}\|_2$ or $\vec{w}^T \vec{x}_i \leq -m\|\vec{w}\|_2 \quad \forall i$

$$\max_{m,\vec{w}} m$$

subject to

$m \geq 0$

$\cancel{\vec{w}^T \vec{x}_i > 0 \quad \forall i \text{ such that } y_i = 1}$ since $\vec{w}^T \vec{x}_i \geq m\|\vec{w}\|_2$, and $m\|\vec{w}\|_2 \geq 0$

$\vec{w}^T \vec{x}_i \geq m\|\vec{w}\|_2 \quad \forall i$ such that $y_i = 1$

$\vec{w}^T \vec{x}_i \leq -m\|\vec{w}\|_2 \quad \forall i$ such that $y_i = -1$

$$\max_{m,\vec{w}} \; m$$

subject to

$m \geq 0$

$\vec{w}^T \vec{x}_i \geq m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = 1$

$\vec{w}^T \vec{x}_i \leq -m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = -1$

$$\max_{m,\vec{w}} \; m$$

subject to

$m \geq 0$

$\vec{w}^T \vec{x}_i \geq m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = 1$

$-1(\vec{w}^T \vec{x}_i) \geq m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = -1$

$$\max_{m,\vec{w}} \; m$$

subject to

$m \geq 0$

$y_i(\vec{w}^T \vec{x}_i) \geq m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = 1$

$y_i(\vec{w}^T \vec{x}_i) \geq m\|\vec{w}\|_2 \quad \forall_i$ such that $y_i = -1$

$$\max_{m, \vec{w}} \; m$$

subject to

$m \geq 0$

$$y_i \left(\vec{w}^T \vec{x}_i \right) \geq m \|\vec{w}\|_2 \quad \forall i$$

---

devide by $|\vec{w}^T \vec{x}_i|$, since $\frac{1}{x}$ is strictly decresing

$$\max_{\vec{w}} \frac{1}{\|\vec{w}\|_2}$$

subject to

$$y_i \left(\vec{w}^T \vec{x}_i \right) \geq 1 \quad \forall i$$

$$\min_{\vec{w}} \|\vec{w}\|_2$$ since $x \to \frac{1}{2} x^2$ is strictly increasing for $x \geq 0$

subject to

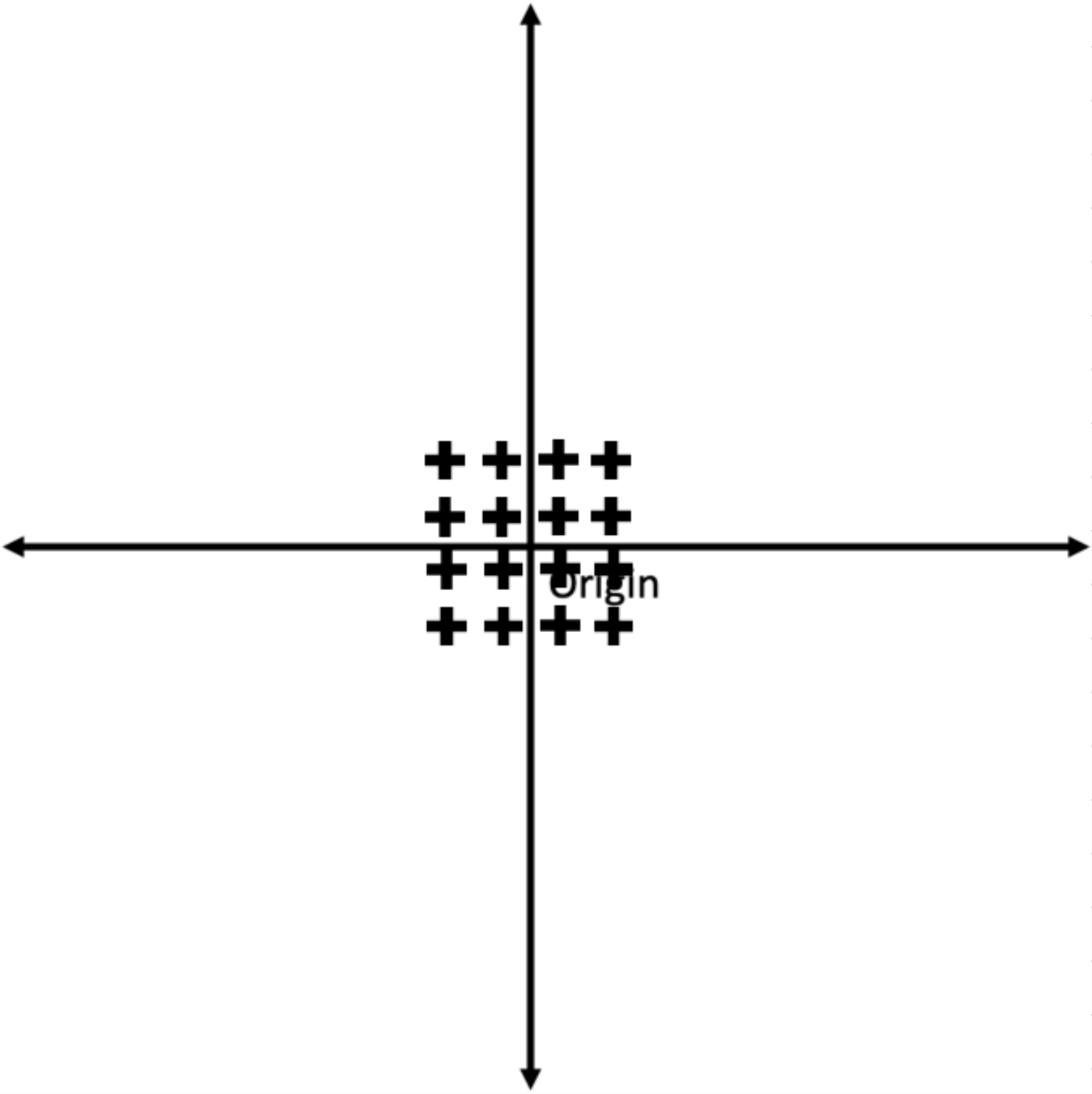
$$y_i \left(\vec{w}^T \vec{x}_i \right) \geq 1 \quad \forall i$$

$$\min_{\vec{w}} \frac{1}{2} \|\vec{w}\|_2^2$$ since $x \to \frac{1}{2} x^2$ is strictly increasing for $x \geq 0$

subject to

$\vec{w}^T \vec{x}_i \geq 1 \quad \forall i$ since $y_i = 1$

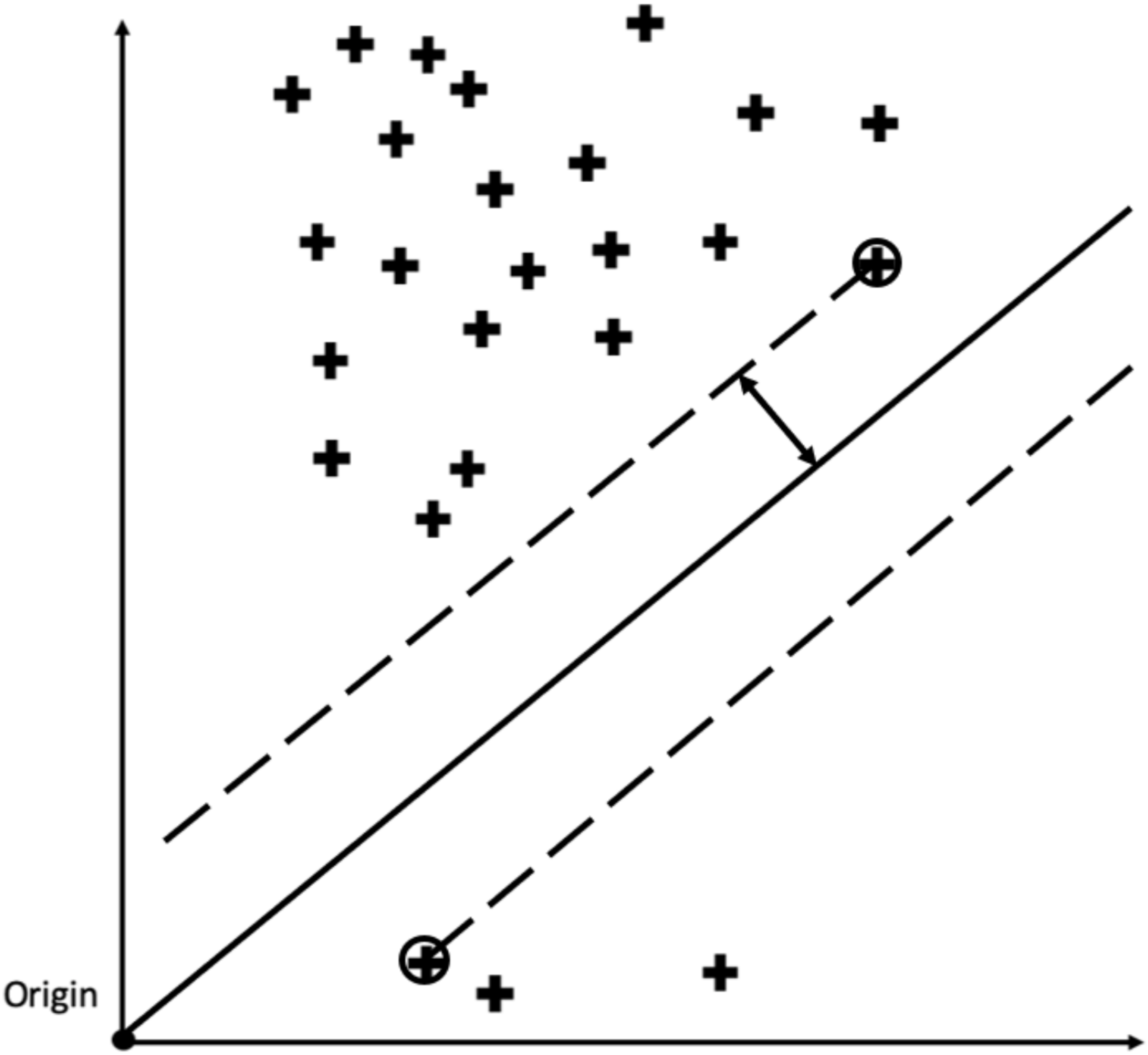
e)



There are no decision boundary such that pass through the origin and all training data points lie on the positive side.

f) Make a neural network, which inputs are the training data, and output is $\vec{w}$ with some hidden layers, and let $\vec{w}^T\vec{x}$ be the loss function, adjust the hyperparamete and train the network until loss is less than 7%.

h)



These two points are lying on the margin, so their dual variables are positive.