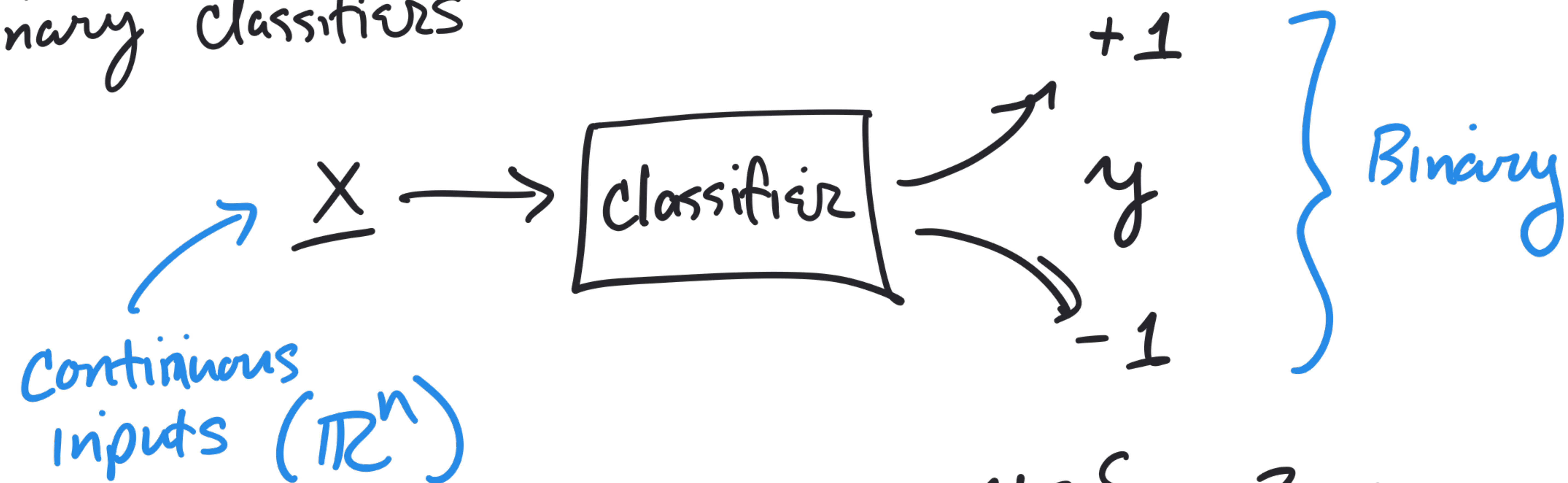
LECTURE 16

Classification +

Support Vector Machines

Classification

- Finding a "classifier"
- Binary classifiers



$y \in \{-1, 1\}$ is an arbitrary notation

$\{0, 1\}, \{-\infty, \infty\}, \dots$

Example Classifiers

1. 1D, Binary classifier
   Diagnosing Diabetes mellitus (type II)

   $\Rightarrow$ Fasting blood glucose level ≥ 126 mg/dl

   $\underline{x} \in \mathbb{R}^1 \longrightarrow \boxed{x_1 \geq \underline{126}}$ $\rightarrow$ +1 diabetes / $\rightarrow$ −1 no diabetes

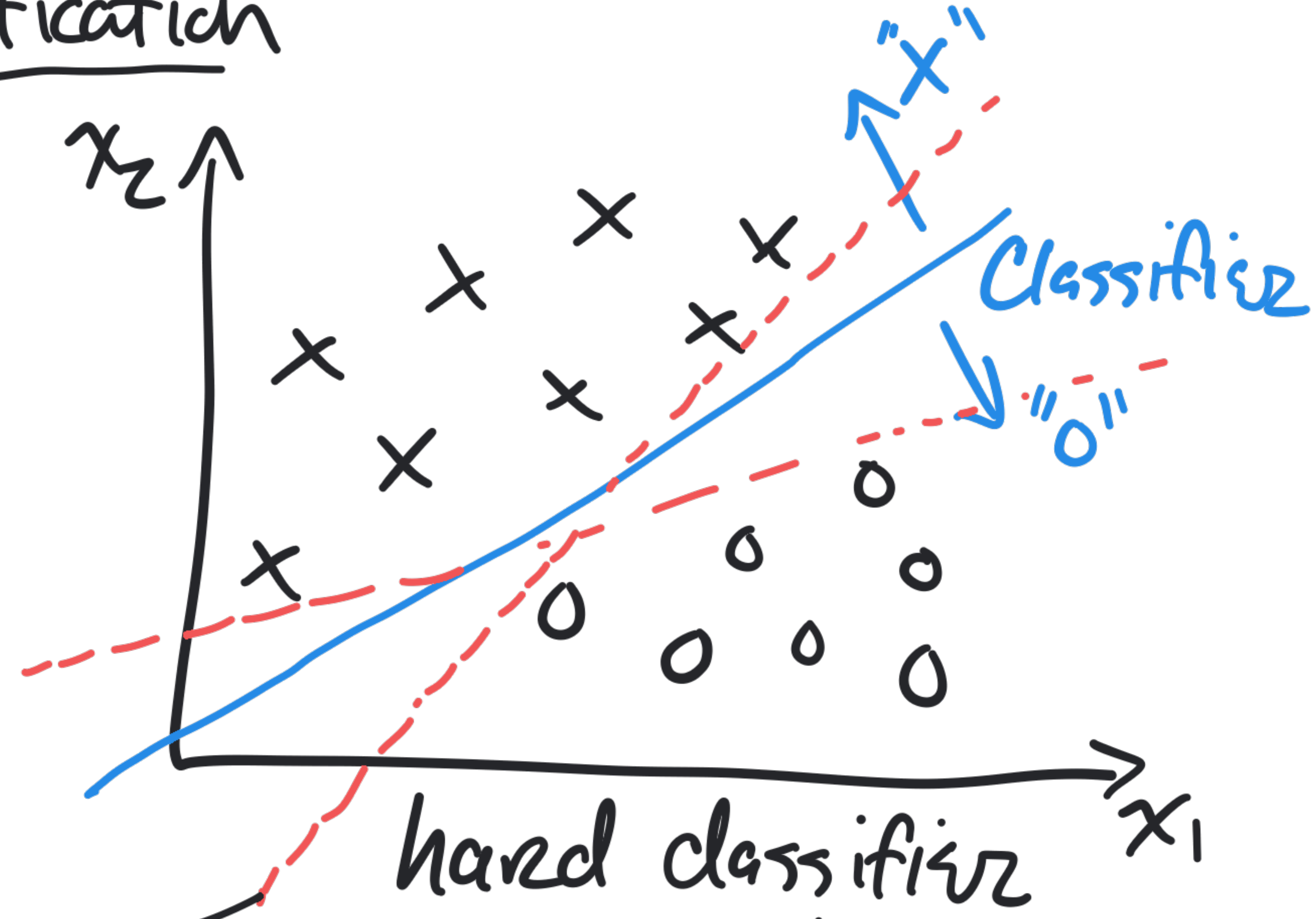
   Blood glucose

2. 2D, Binary Classifier
   Diagnosing obesity from height, weight

   $\underline{x} \in \mathbb{R}^2 = \begin{pmatrix} \text{height [m]} \\ \text{weight [kg]} \end{pmatrix} \longrightarrow \boxed{\frac{w}{h^2} \geq 30}$ $\rightarrow$ +1 obese / $\rightarrow$ −1 normal

Geometric view of Classification

Classify as "x" or "o"



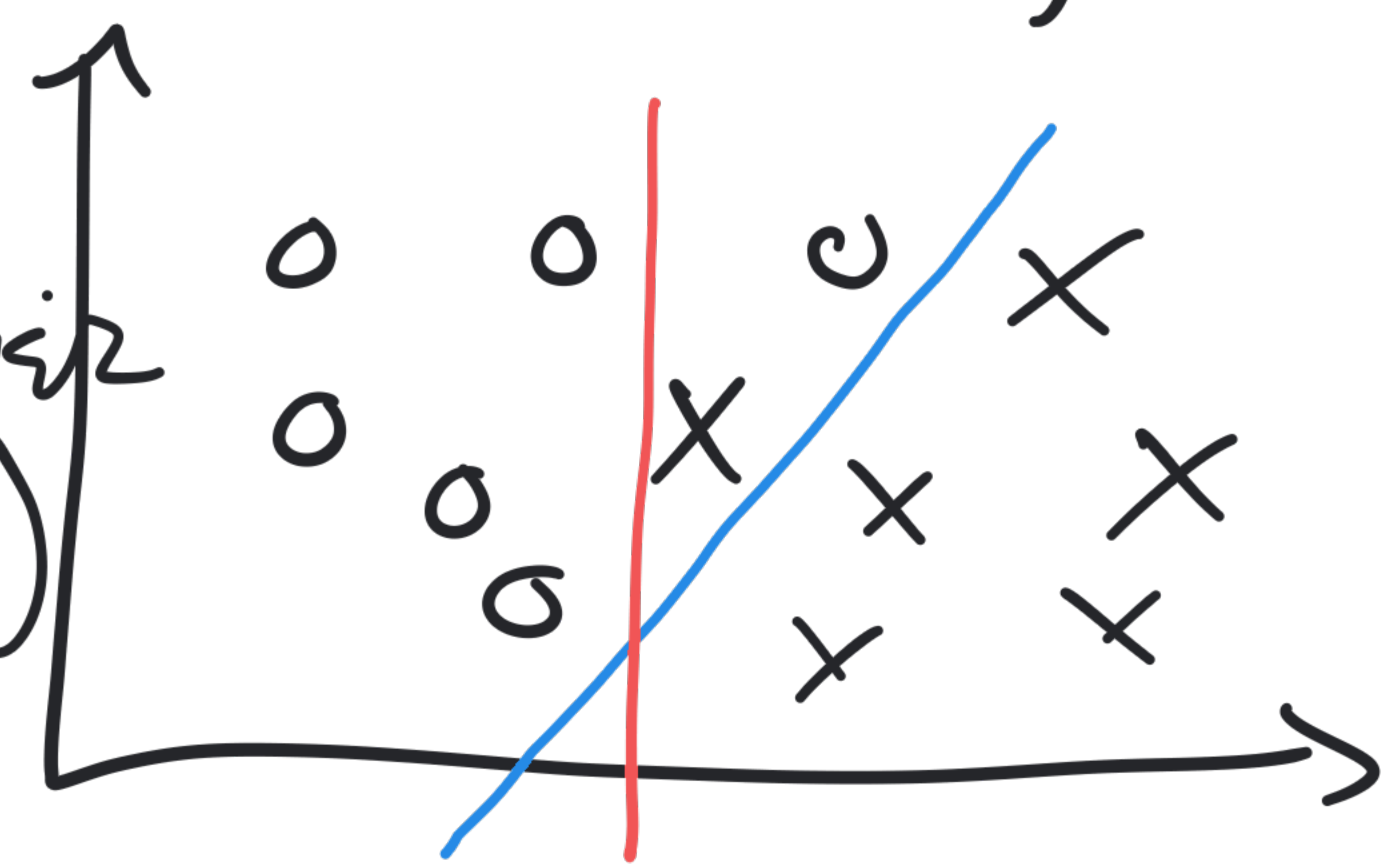
maximize distance around classifier

hard classifier (no exceptions in our data)

1.) minimize # of mis-classifications

2.) max. distance

soft classifier (exceptions in data)



Hard Classifier

Training data $(\underline{x}_1, y_1), (\underline{x}_2, y_2), \ldots, (\underline{x}_k, y_k)$

multi-dimensional, continuous input

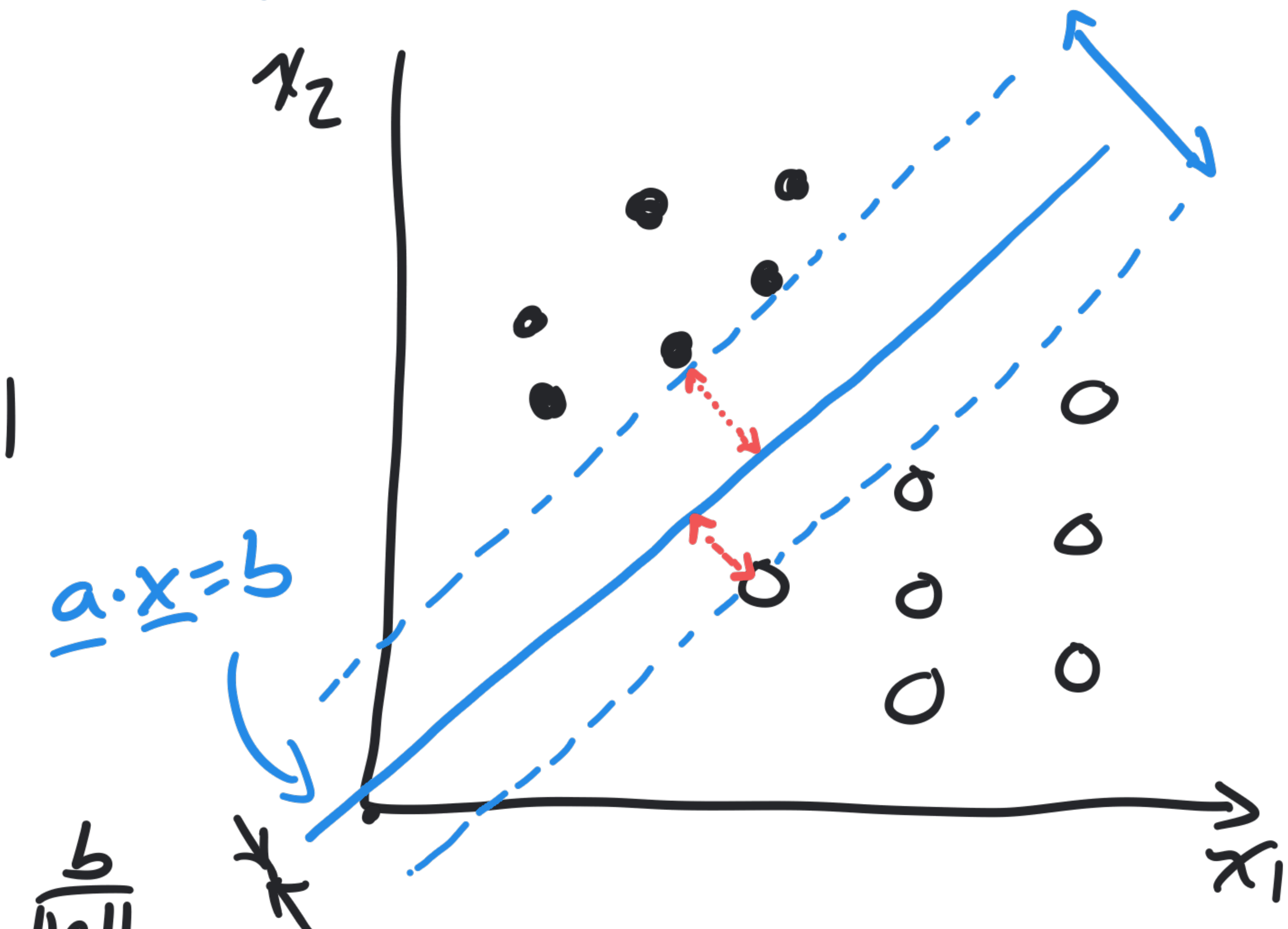
binary output $\{-1, +1\}$

max distance b/w points
s.t.
all +1 above line
all −1 below line

$$\underline{a} \cdot \underline{x} = b + 1 \quad \forall\, y = +1$$

$$\hat{\underline{a}} \cdot \underline{x} = \underbrace{\frac{b}{\|\underline{a}\|}}_{\text{d from origin to line}} + \underbrace{\frac{1}{\|\underline{a}\|}}_{\text{d from line to nearest +1 pt.}}$$

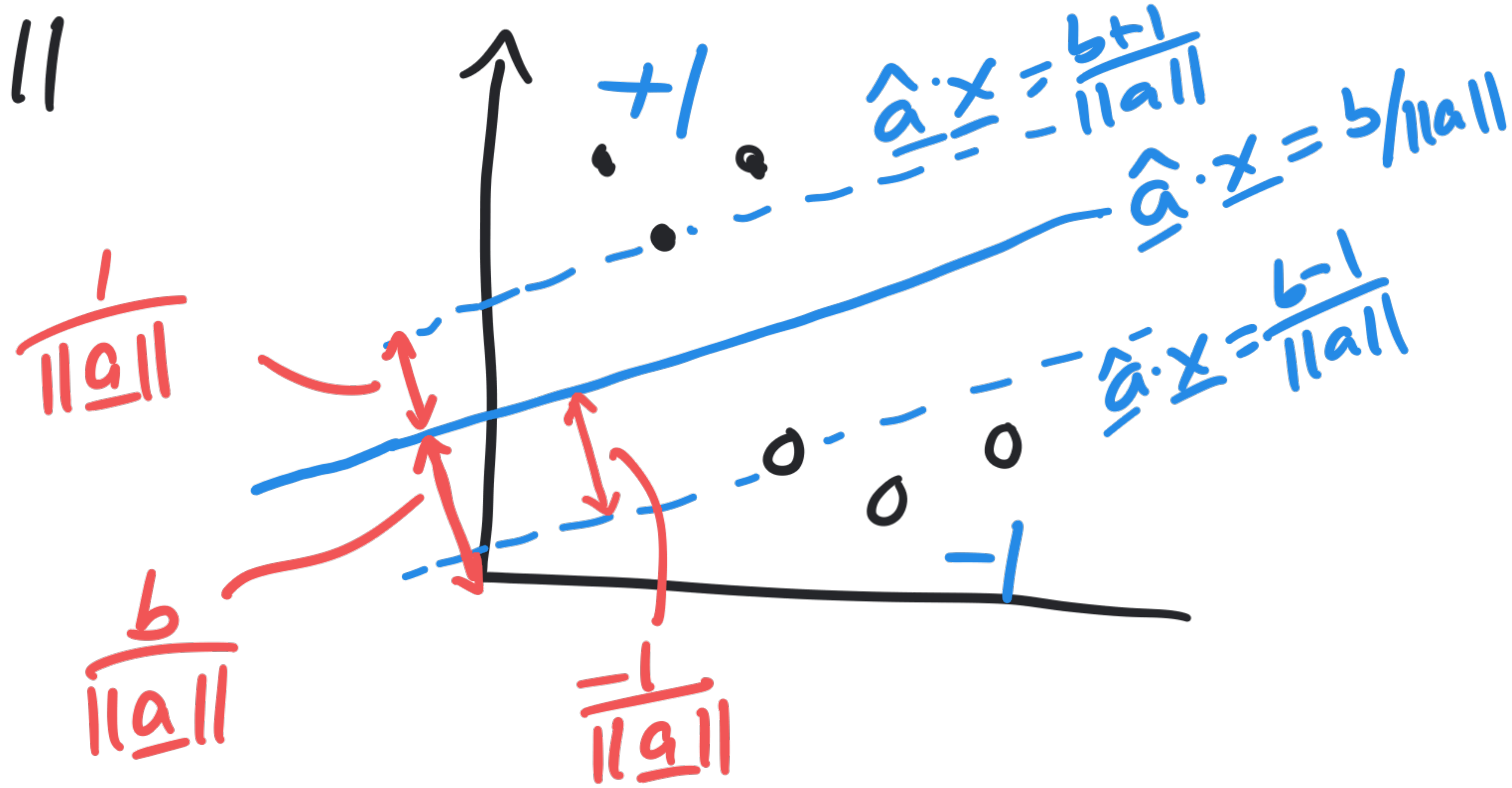
$$\hat{\underline{a}} \cdot \underline{x} = \frac{b}{\|\underline{a}\|} - \frac{1}{\|\underline{a}\|} \quad \forall\, y = -1$$



$\underline{\text{obj}}$: $\max \frac{2}{\|\underline{a}\|} = \min \|\underline{a}\|$

$= \min \sum a_i^2$

$\|\underline{a}\| = \sqrt{a_1^2 + \cdots + a_n^2}$



$\underline{a} \cdot \underline{x} \geq b+1 \quad \forall y = +1$

$\underline{a} \cdot \underline{x} \leq b-1 \quad \forall y = -1$

find

$\underline{a} \cdot \underline{x}_i \geq b + y_i \quad \forall y_i = +1$

$\underline{a} \cdot \underline{x}_i \leq b + y_i \quad \forall y_i = -1$

$y_i = 1$

$\underline{a} \cdot \underline{x}_i - b \geq 1$

$\boxed{y_i(\underline{a} \cdot \underline{x}_i - b) \geq 1}$

$y_i = -1$

$-(\underline{a} \cdot \underline{x}_i - b) \geq 1$

$\underline{a} \cdot \underline{x}_i - b \leq -1$

Support Vector Machine problem

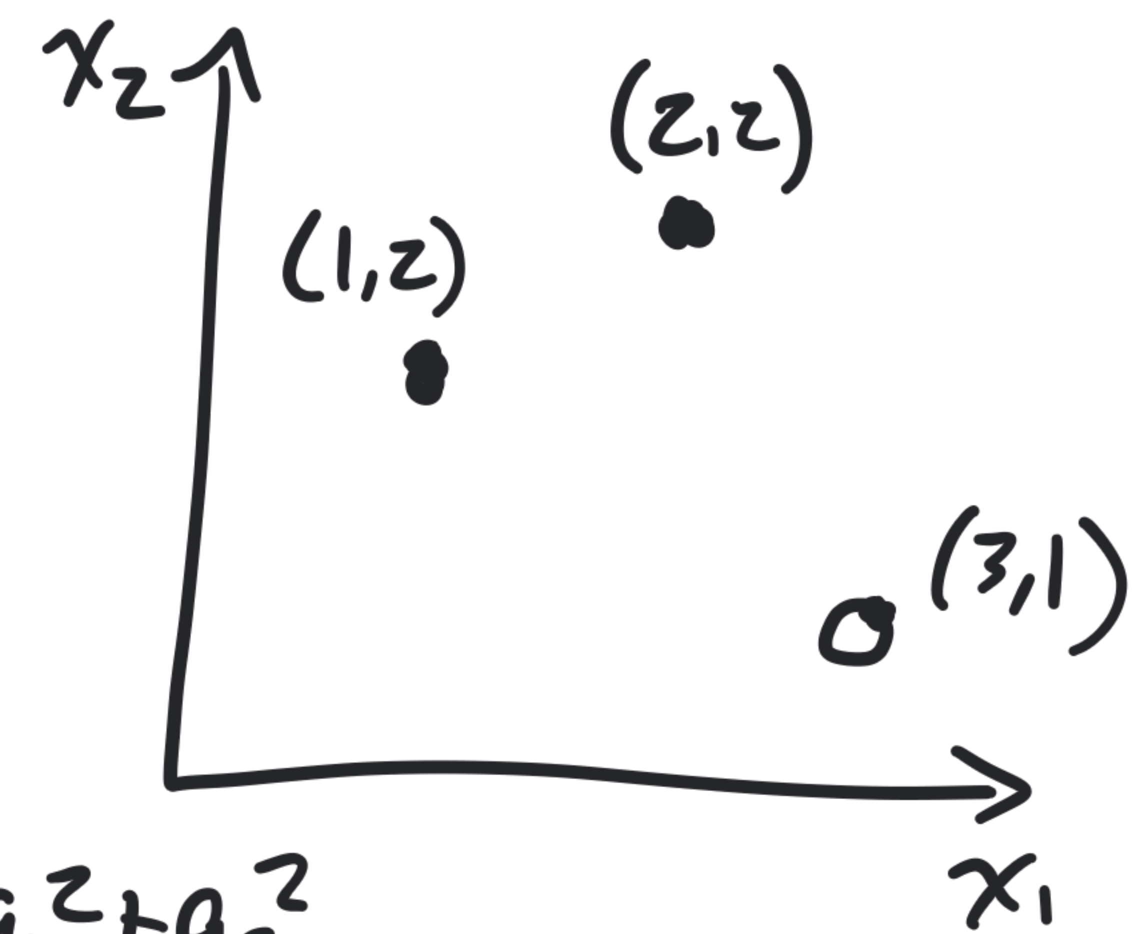
$$\min_{\underline{a}, b} \sum_{i=1}^{n} a_i^2$$ } maximize distance b/w planes

s.t. $y_i(\underline{a} \cdot \underline{x} - b) \geq 1 \quad \forall\, i = 1, \ldots, k$ } Ensure points are on correct side of planes.

$k=3$

$\mathbb{R}^2$
$n=2$

$$\begin{array}{c}
[(1,2), +1] \\
[(2,2), +1] \\
[\underbrace{(3,1)}_{\underline{x}}, \underbrace{-1}_{y}]
\end{array}$$



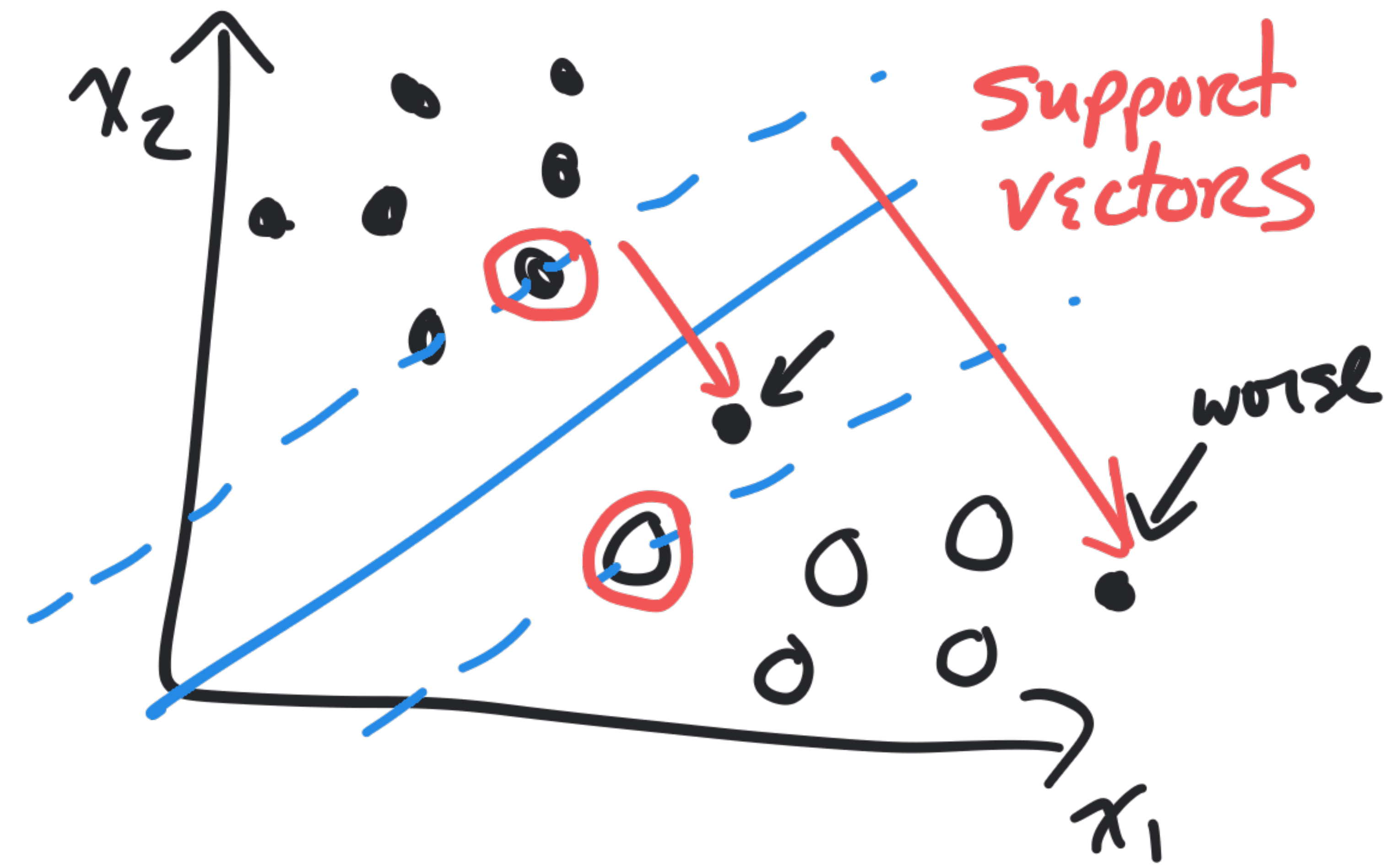
$$\min_{a_1, a_2, b} a_1^2 + a_2^2$$

s.t. $+1(a_1(1) + a_2(2) - b) \geq 1$
$+1(a_1(2) + a_2(2) - b) \geq 1$
$-1(a_1(3) + a_2(1) - b) \geq 1$

$$\min_{a_1, a_2, b} a_1^2 + a_2^2$$

s.t. $a_1 + a_2 - b \geq 1$
$2a_1 + 2a_2 - b \geq 1$
$-3a_1 - a_2 + b \geq 1$

1. Why is it called SVM?



2. What about soft classifiers

For hard, we had constraints

$$y_i(\underline{a}\cdot\underline{x}_i - b) \geq 1$$

If $\underline{x}_i$ is classified correctly,

$$\underline{1 - y_i(\underline{a}\cdot\underline{x}_i - b)} \leq 0$$

For soft classification, minimize this violation

$$\min\left[\frac{1}{n}\underbrace{\sum \max\{0, 1 - y_i(\underline{a}\cdot\underline{x}_i - b)\}}_{\substack{0 \text{ if class. correctly} \\ d>0 \text{ of mis-classification}}}\right] + \underbrace{\lambda \sum a_i^2}_{\substack{\text{separating distance} \\ \text{(like hard class.)}}}$$

After break

Set up + solve SVM problems in Matlab

$\Rightarrow$ HW 5: Classify tumors into subtype given gene expression