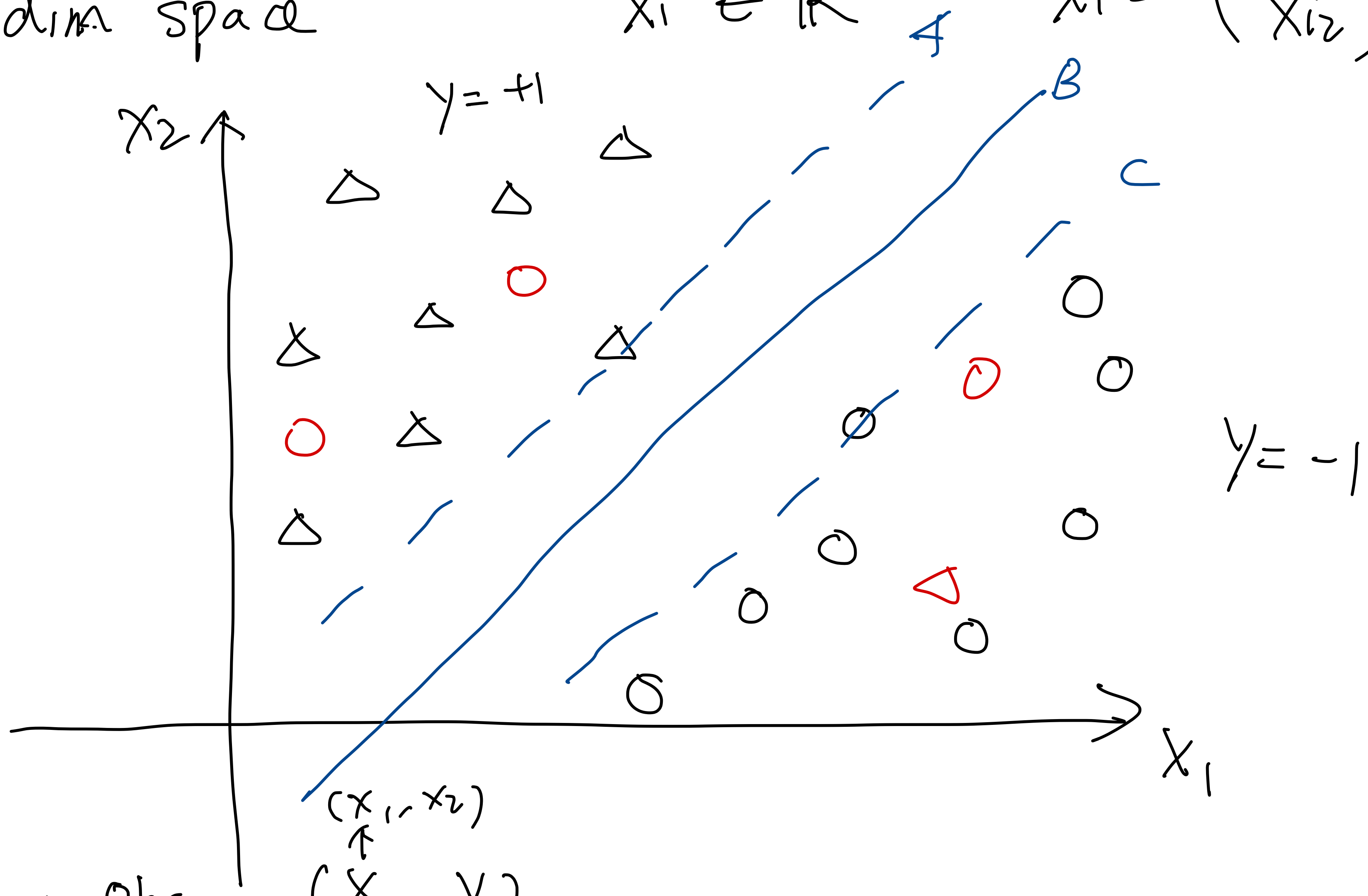
SVM (Support Vector Machine)

Training Data $(x_1, y_1), \cdots, (x_n, y_n)$

Binary Classification $y_i = +1$ or -1

2-dim space $x_i \in \mathbb{R}^2$ $x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix}$



New Obs. (x, y)

Model $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

$$f(x) = \hat{Y} = \begin{cases} +1 & \text{if } f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \geq 0 \\ -1 & \text{if } \beta_0 + \beta_1 x_1 + \beta_2 x_2 < 0 \end{cases}$$

Q: How to find $f(x)$? i.e. $\beta_0, \beta_1, \beta_2$?

Principle of SVM: Maximal Margin Classifier

A: $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 1$ Side line

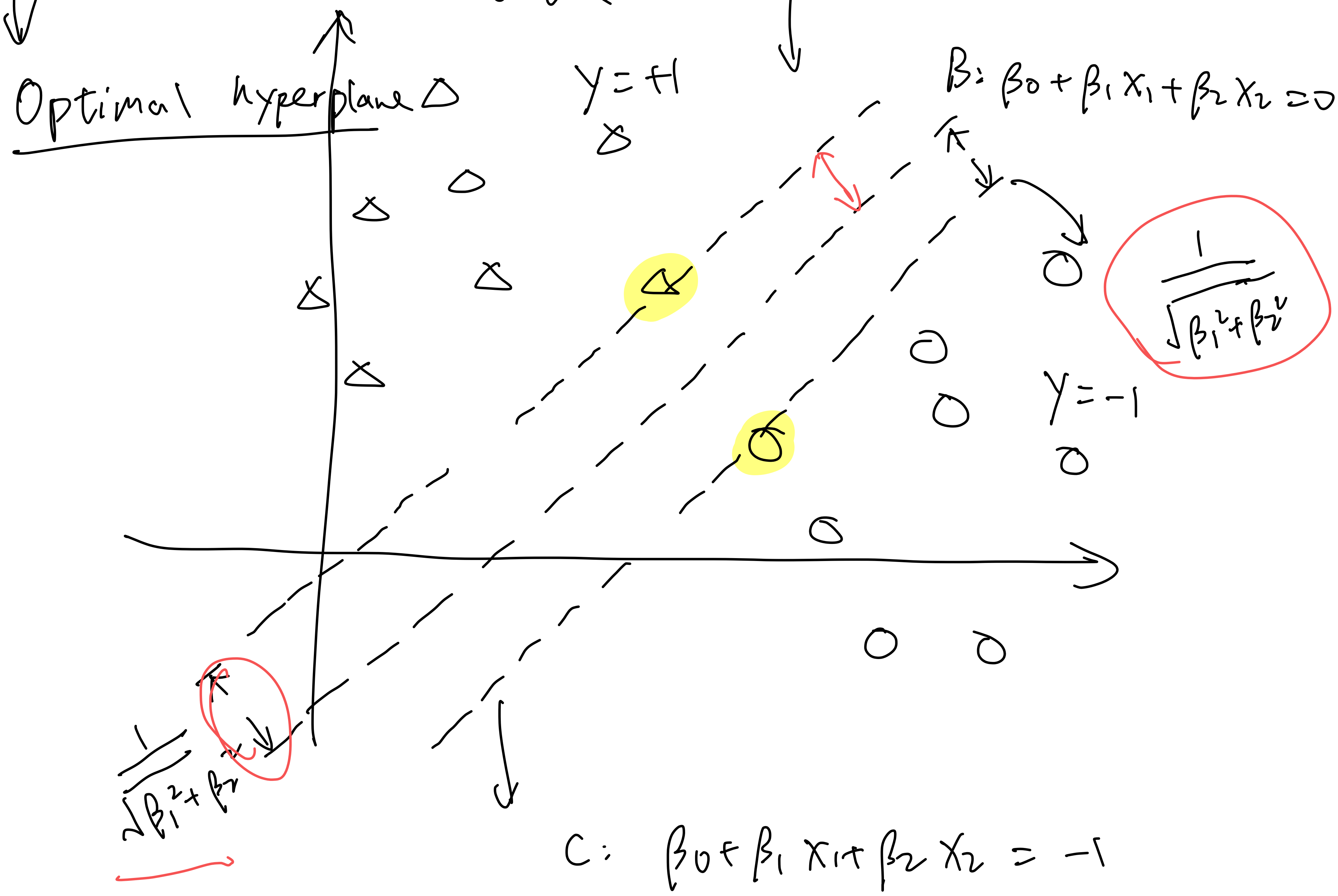
B: $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0$ Mid line

C: $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = -1$ Side line

β_1, β_2 slope

β_0 intercept

A : $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = +1$



Line A , C : pass through the point which is the closest point to Line B

$$d(A, B) = d(B, C) = \frac{1}{\sqrt{\beta_1^2+\beta_2^2}} \rightarrow \text{Margin}$$

Hard-Margin SVM

$$\max \frac{1}{\sqrt{\beta_1^2+\beta_2^2}} \iff \min \ \beta_1^2+\beta_2^2 \quad (*)$$

$$\text{s.t.} \quad y_i(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}) \geq 1$$

$$i = 1, 2, \cdots, n$$

Constrained Opt.

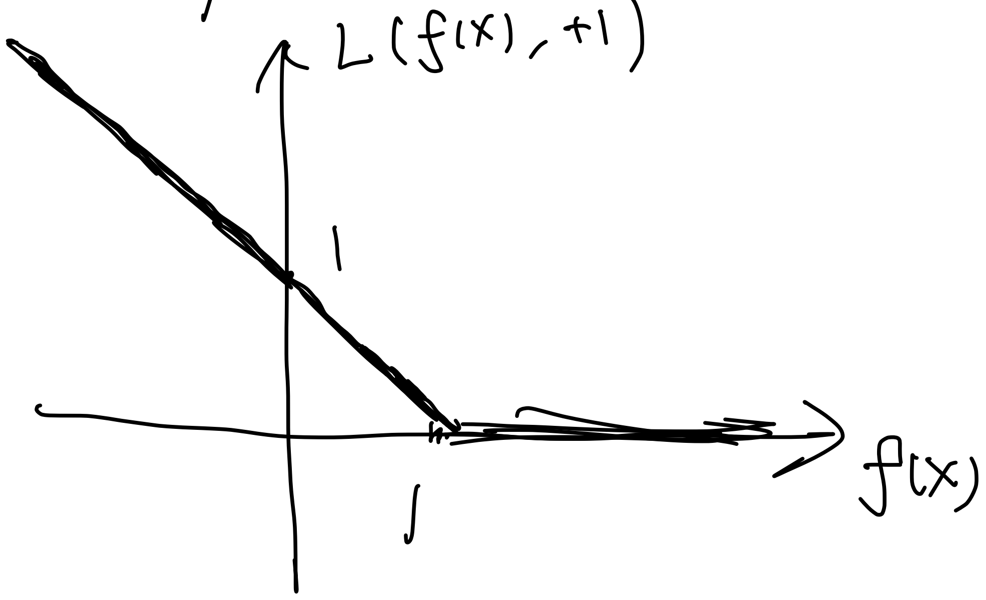
$(*)$ may not have solution when data is not linearly separable

Soft-Margin SVM

Hinge Loss $\quad f(x) = \beta_0+\beta_1 x_1+\beta_2 x_2 \ , \ y = +1, -1$

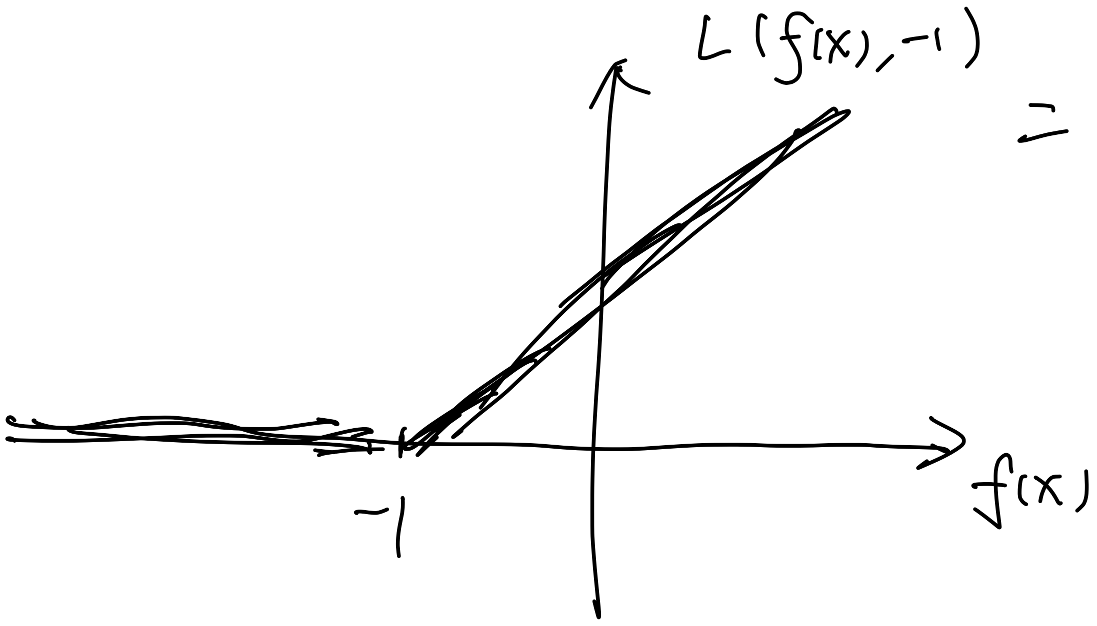
$$L(f(x), y) = \max(0, \ 1 - y \cdot f(x))$$

I. $y = +1$, $\quad L(f(x), y) = \max(0, \ 1 - f(x))$



$$= \begin{cases} 0, & f(x) \geq 1 \\ 1 - f(x), & f(x) < 1 \end{cases}$$

II $y=-1$, $L(f(x), y) = \max(0, 1+f(x))$



$$= \begin{cases} 0, & f(x) \leq -1 \\ 1+f(x), & f(x) > -1 \end{cases}$$

Soft-Margin SVM [λ is a hyperparameter]

$$\min_{\beta_0, \beta_1, \beta_2} \sum_{i=1}^{n} \max\left(0, y_i \cdot (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})\right) + \lambda \cdot \left(\beta_1^2 + \beta_2^2\right) \quad (**)$$

(**) is a convex problem with the solvable optimal solution