1906534

classmate Date Page

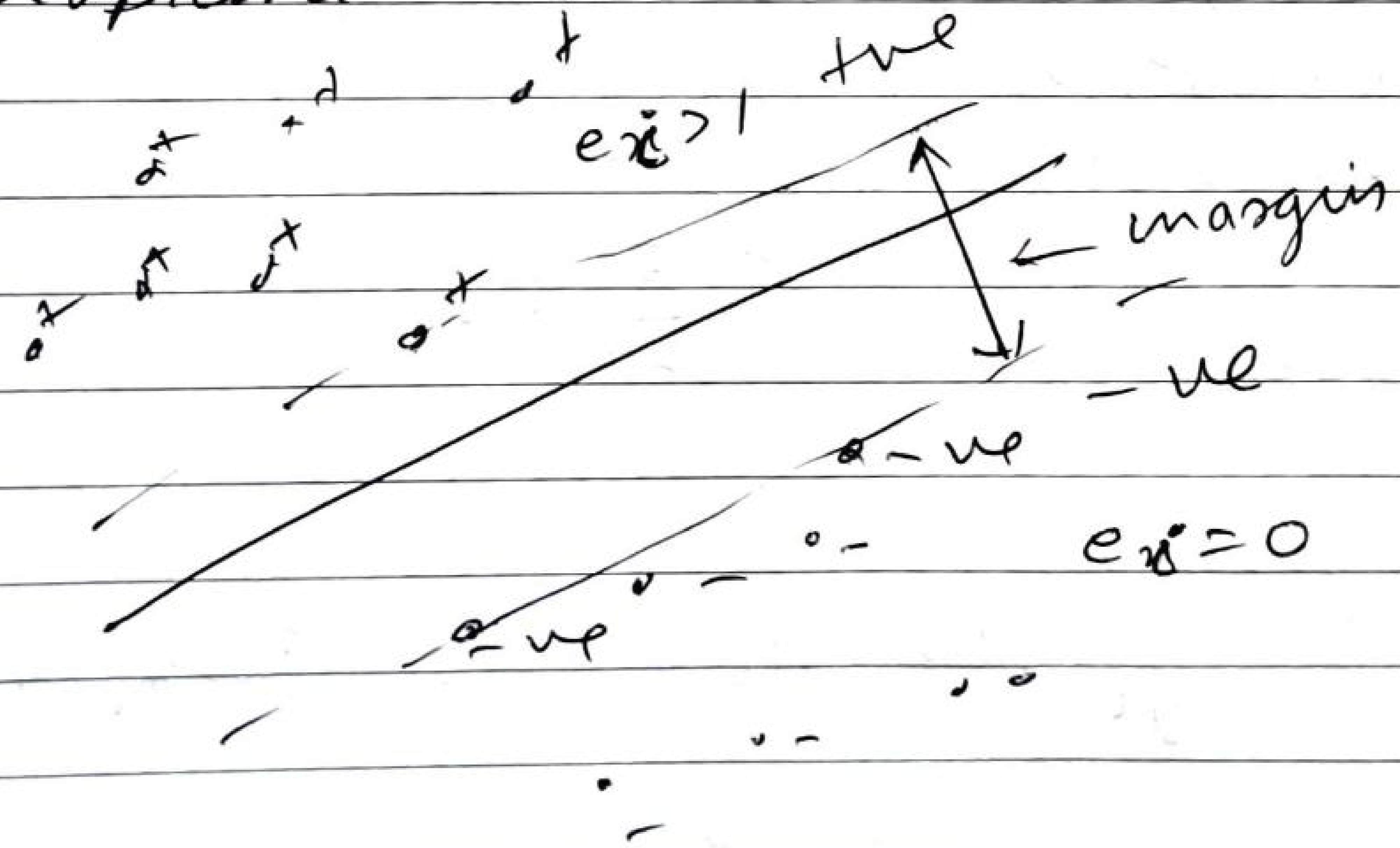
~~③ Introd~~

③ Soft Margin Hyperplane in SVM

In real life problem, the two-class datasets are only rarely linearly separable.

There are two type of derivation:-
→ An instance may lie on the wrong side of the hyperplane and be misclassified.

→ An instance may be on the right side but may lie in the margin [i.e. Not sufficiently away from the hyperplane.



In such cases, we introduce additional variables, ϵ_i called slack variables which store deviation from the margin.

1906534

$E_i = 0$ $\bar{x}_i$ is correctly classified

$0 < E_i < 1$ $\bar{x}_i$ is correctly classified But is in the margin

$E_{x_i} > 1$ $\bar{x}_i$ is misclassified.

So SVM problem can be reformulated as follows:

Given a two class linearly seperable dataset of N points of the form:

$(\bar{x}_1, y_1)(\bar{x}_2, y_2), \ldots, (\bar{x}_N, y_N)$

where y_i's are either +1 or -1

find vector $\bar{w}$ and $\bar{E}_i$ and a number b

which minimize

$$\frac{1}{2}||w||^2 + C\sum_{i=1}^{N} E_i$$

Subject to

$$y_i(\bar{w} \cdot \bar{x}_i + b) \geq 1 - E_{x_i} \text{ for } i = 1, 2, \ldots, N$$

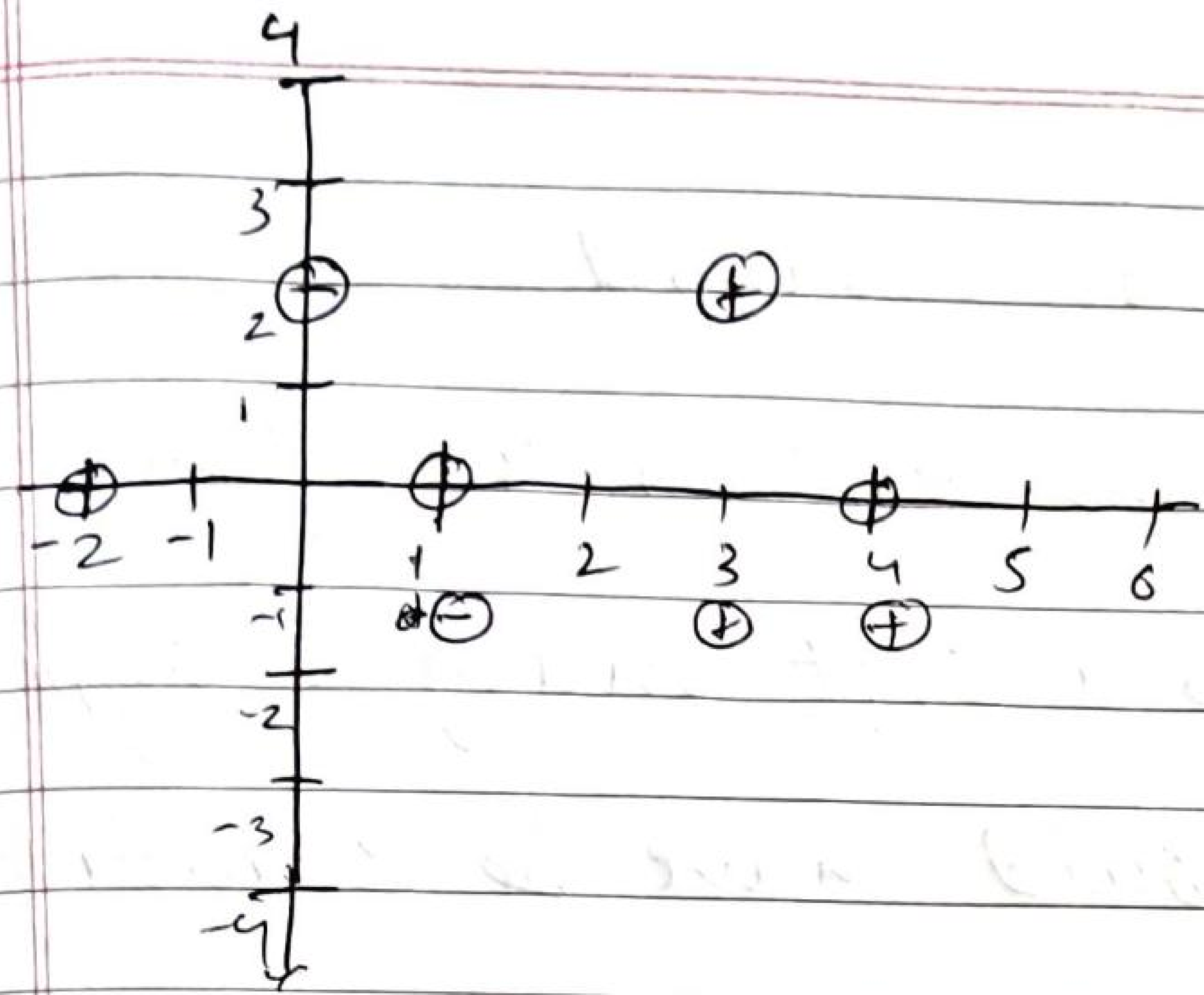
$$E_{x_i} \geq 0 \text{ for } i = 1, 2, \ldots N$$

The hyperplane given by the equation

$\bar{w} \cdot \bar{x} + b = 0$ with the value of $\bar{w}$ and b obtained as solution of the reformulated problem, is called soft margin hyperplane for the SVM problem.

1906534

(b)



Support Vectors

$S_1 = (1,0)$, $S_2 = (3,2)$

$\therefore S_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\tilde{S}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$S_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$; $\tilde{S}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

$\Rightarrow$ 1 is bias input

now, we have

$\alpha_1 \tilde{S}_1 \cdot \tilde{S}_1 + \alpha_2 \tilde{S}_2 \tilde{S}_1 = -1$ — (I)

$\alpha_1 \tilde{S}_1 \cdot \tilde{S}_2 + \alpha_2 \tilde{S}_2 \cdot \tilde{S}_2 = +1$ — (II)

for (I),
$$\alpha_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = -1$$

$$\Rightarrow 2\alpha_1 + 4\alpha_2 = -1 \quad \text{— (III)}$$

for (II),
$$\alpha_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = +1$$

$$\Rightarrow 4\alpha_1 + 14\alpha_2 = +1 \quad \text{— iv)}$$

Solving (III) & (iv) we get $\alpha_1 = -1.5$, $\alpha_2 = 0.5$

now, weight vector $(\bar{w}) = \sum_i \alpha_i \tilde{S}_i = -1.5 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$+ 0.5 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

Scanned by TapScanner

1906537

$$\Rightarrow \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \rightarrow \text{bias input}$$

$\therefore$ our hyperplane eq^n is $y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} x - 1$

and this is a soft margin hyperplane as these would have be 2 additional support vectors but we have ignored them.