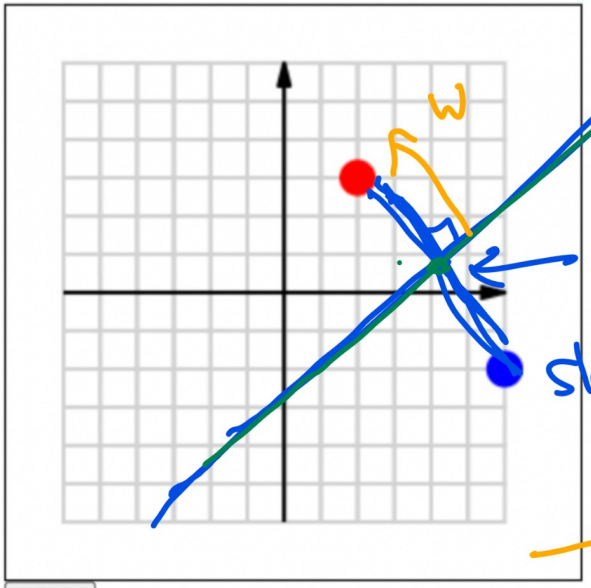
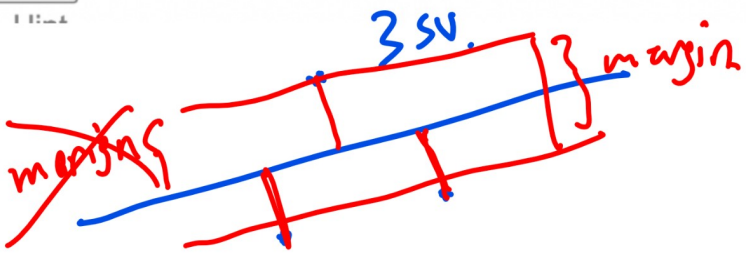
Question 4

• [6 points] A linear SVM (Support Vector Machine) with with weights w_1, w_2, b is trained on the following data set: $x_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $y_1 = 0$ and $x_2 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$, $y_2 = 1$. The attributes (i.e. features) are two dimensional (x_{i1}, x_{i2}) and the label y_i is binary. The classification rule is $\hat{y}_i = 1_{\{w_1 x_{i1} + w_2 x_{i2} + b \geq 0\}}$. Assuming $b = -27$, what is (w_1, w_2) ? The drawing is not graded.



Clear



$$\begin{cases} x_2 = 0.8 x_1 + (\frac{1}{2} - 3.2) \\ 1/2 = 0.8 \cdot 4 + ? \end{cases}$$

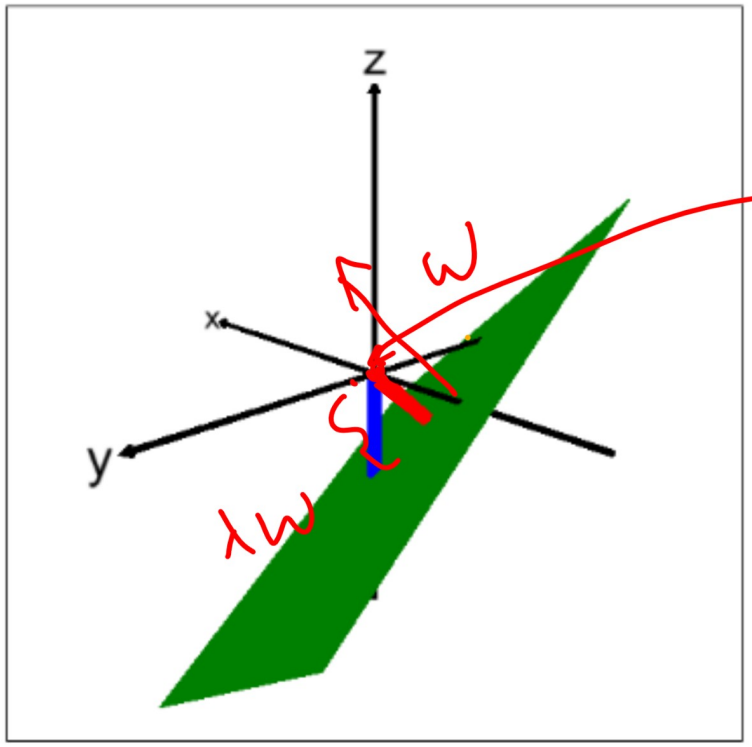
$$0.8 x_1 - x_2 - 2.7 = 0 \qquad b = -27$$

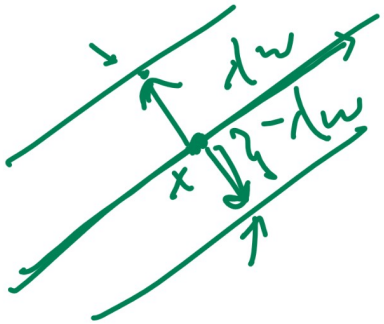
$$8 x_1 - 10 x_2 - 27 = 0$$

MQ2

Question 2

• [2 points] Given a weight vector $w = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$, consider the line (plane) defined by $w^T x = c = 2.22$. Along this line (on the plane), there is a point that is the closest to the origin. How far is that point to the origin in Euclidean distance?





$0 + \lambda w$ is on the plane

$w^T(\lambda w) = 2.22$

$\lambda = \frac{2.22}{w^T w} = \frac{2.22}{\|w\|^2}$

length of $\lambda w \Rightarrow \|\lambda w\| = \lambda \|w\|$

$\frac{2.22}{\|\begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}\|} = \frac{2.22}{\sqrt{4+9+4}} = \dots$

$= \frac{2.22}{\|w\|^2} \|w\| = \frac{2.22}{\|w\|}$

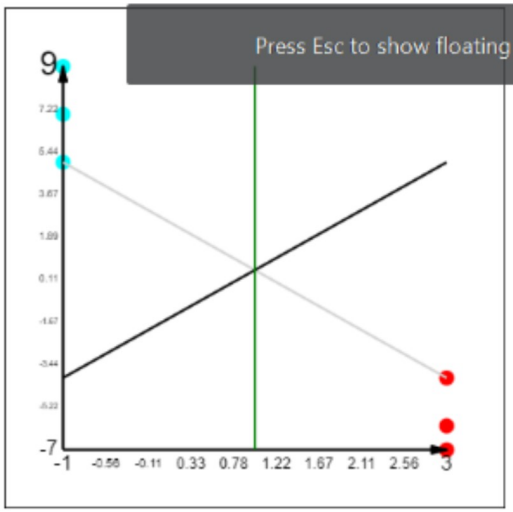
Question 6

M4 Q6

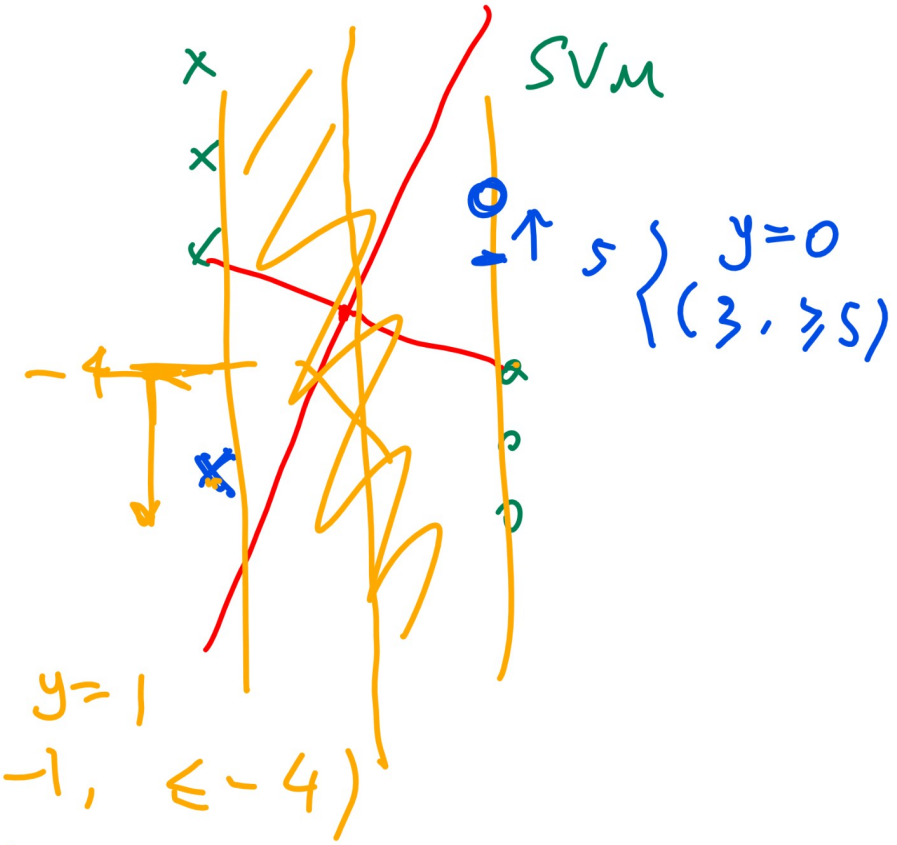
• [4 points] Given the following training set, add one instance $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ with $y = 0$ so that all instances are support vectors for the Hard Margin SVM (Support Vector Machine) trained on the new training set.

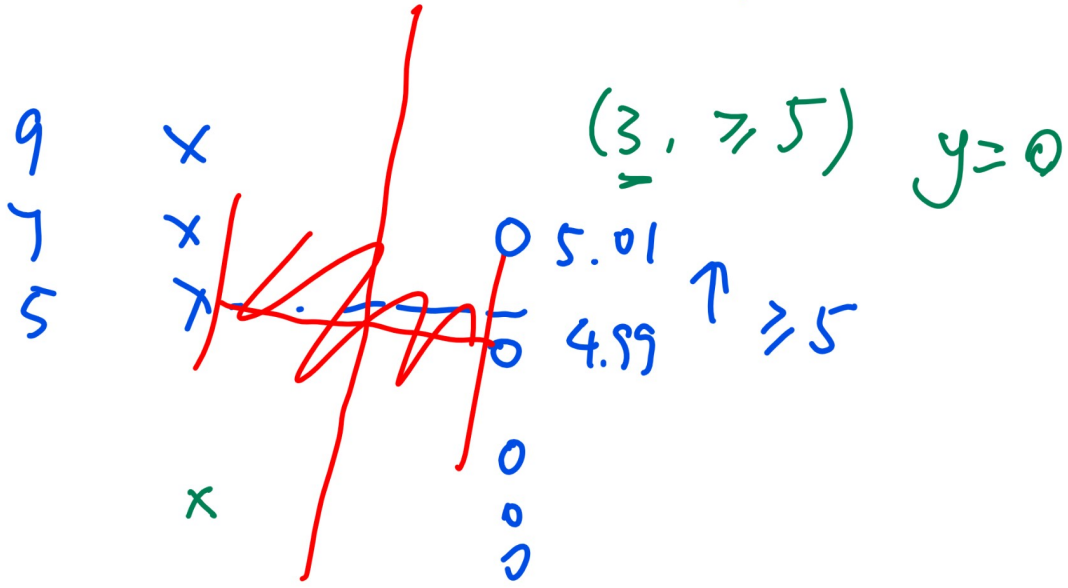
y=1

x_1	x_2	y
3	−6	0
3	−7	0
3	−4	0
−1	5	1
−1	9	1
−1	7	1



Press Esc to show floating meeting controls





m5 Q5

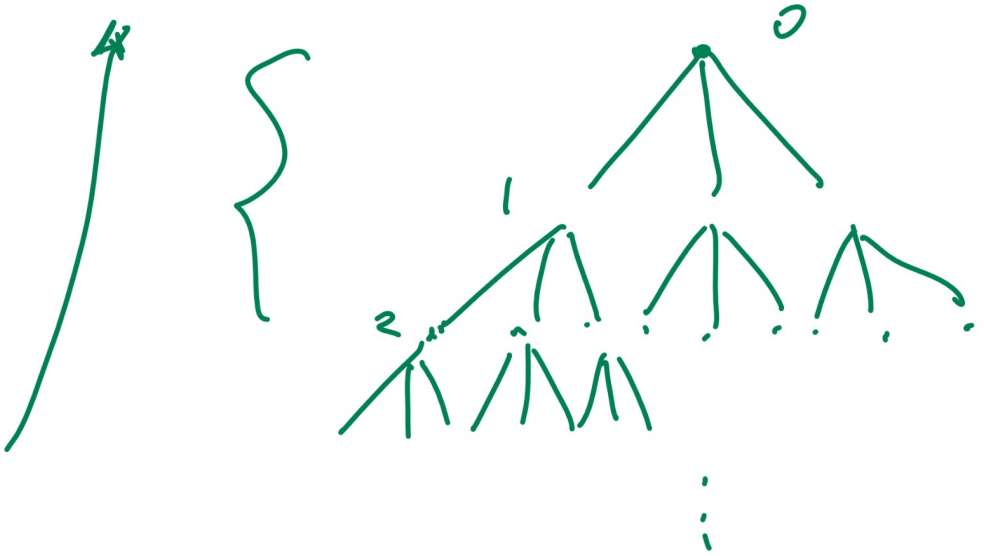
Question 5

• [3 points] A decision tree has depth $d = 5$ (a decision tree where the root is a leaf node has $d = 0$). All its internal node has $b = 3$ children. The tree is also complete, meaning all leaf nodes are at depth d. If we require each leaf node to contain at least 20 training examples, what is the minimum size of the training set?

if depth is 2 :

3^2 leaf · 20 items

$3^2 \cdot 20 = 180$



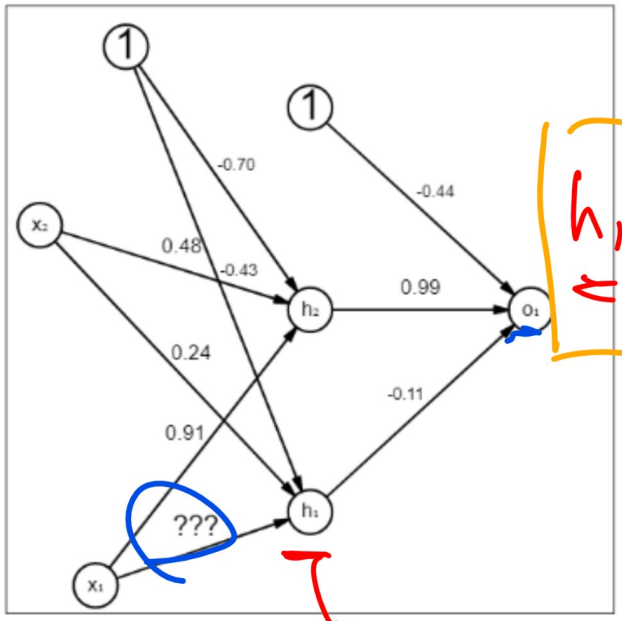
if depth is 5 → $3^5 \cdot 20 = \sim$

Question 3

M323

• [4 points] Fill in the missing weight below so that it computes the following function. All inputs takes value 0 or 1, and the perceptrons are linear threshold units. The first layer weight matrix is $\begin{bmatrix} w & 0.91 \\ 0.24 & 0.48 \end{bmatrix}$, with bias vector $[-0.43 \quad -0.7]$, and the second layer weight vector is $\begin{bmatrix} -0.11 \\ 0.99 \end{bmatrix}$, with bias $[-0.44]$.

x_1	x_2	y or o_1
0	0	0
0	1	0
1	0	1
1	1	1



h_1, h_2

$j=1, j=2$

h_1: 0 or 1

h_2: 0, 1, 1, 1

$h_2 = \mathbb{1}\{0.91 x_1 + 0.48 x_2 - 0.7 \geq 0\}$

$w^T x + b = 0 \Rightarrow h_2 = 1$

$\mathbb{1}\{-0.7 \geq 0\} = 0$ False

$\mathbb{1}\{0.48 - 0.7 \geq 0\} = 0$ False

$h_1 = \mathbb{1}\{w x_1 + 0.24 x_2 - 0.43 \geq 0\}$

$o_1 = \mathbb{1}\{-0.11 h_1 + 0.99 h_2 - 0.44 \geq 0\}$

0, 1

0

1

0

1

if $w x_1 + 0.24 x_2 - 0.43 \geq 0 \Rightarrow h_1 = 1$

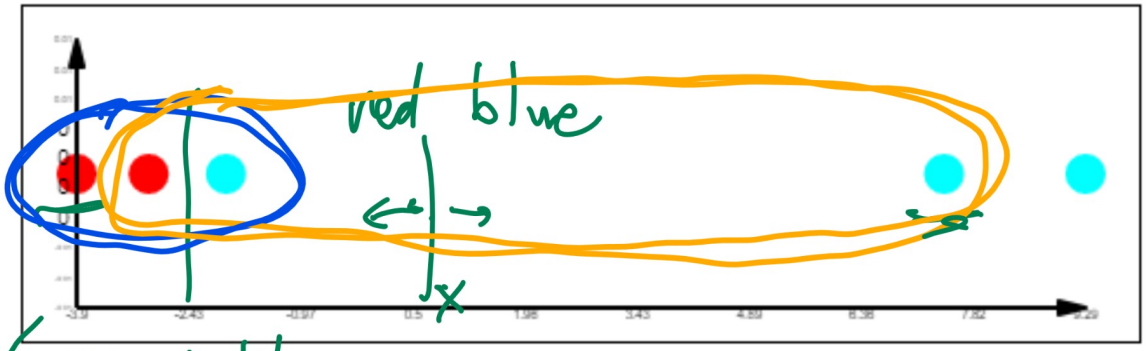
$< 0 \Rightarrow h_1 = 0$

M5Q9

Question 9

5 → 6

• [4 points] You are given a training set of five points and their 2-class classifications (+ or -): $(-3.9, +)$, $(-2.96, +)$, $(-1.95, -)$, $(7.44, -)$, $(9.29, -)$. What is the decision boundary associated with this training set using 3NN (3 Nearest Neighbor)?



1NN
svm decision boundary

if there are 6 points..

1 1 1 0 0 0

3NN $\Sigma(p_r + p_g)$

if $x - \epsilon \Rightarrow$ red $\Rightarrow$ closer to p1, 2, 3

$\Rightarrow$ closer to p1 vs p4

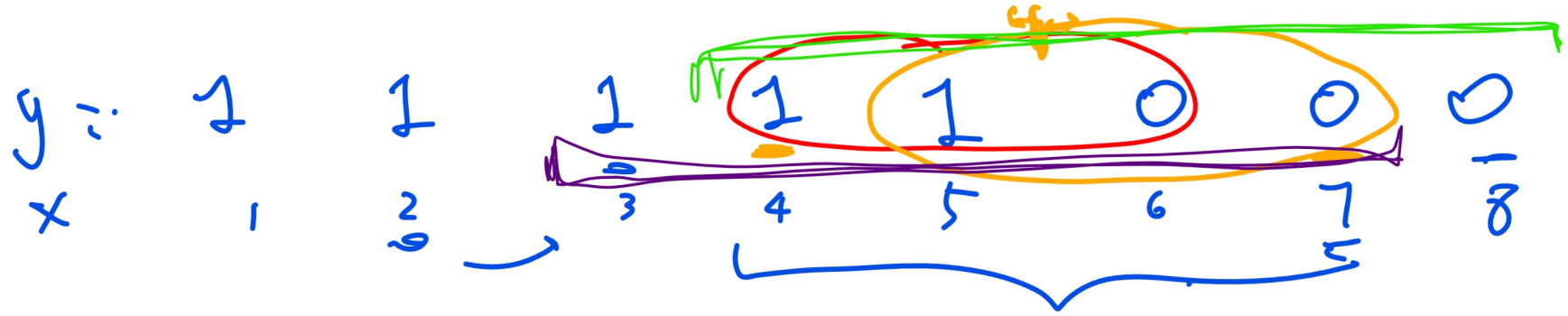
if $x + \epsilon \Rightarrow$ blue $\Rightarrow$ closer to p2, 3, 4 vs p1

$\Rightarrow$ close to p4

midpoint

p1 p4

$x = \frac{1}{2}(-3.9 + 7.44)$



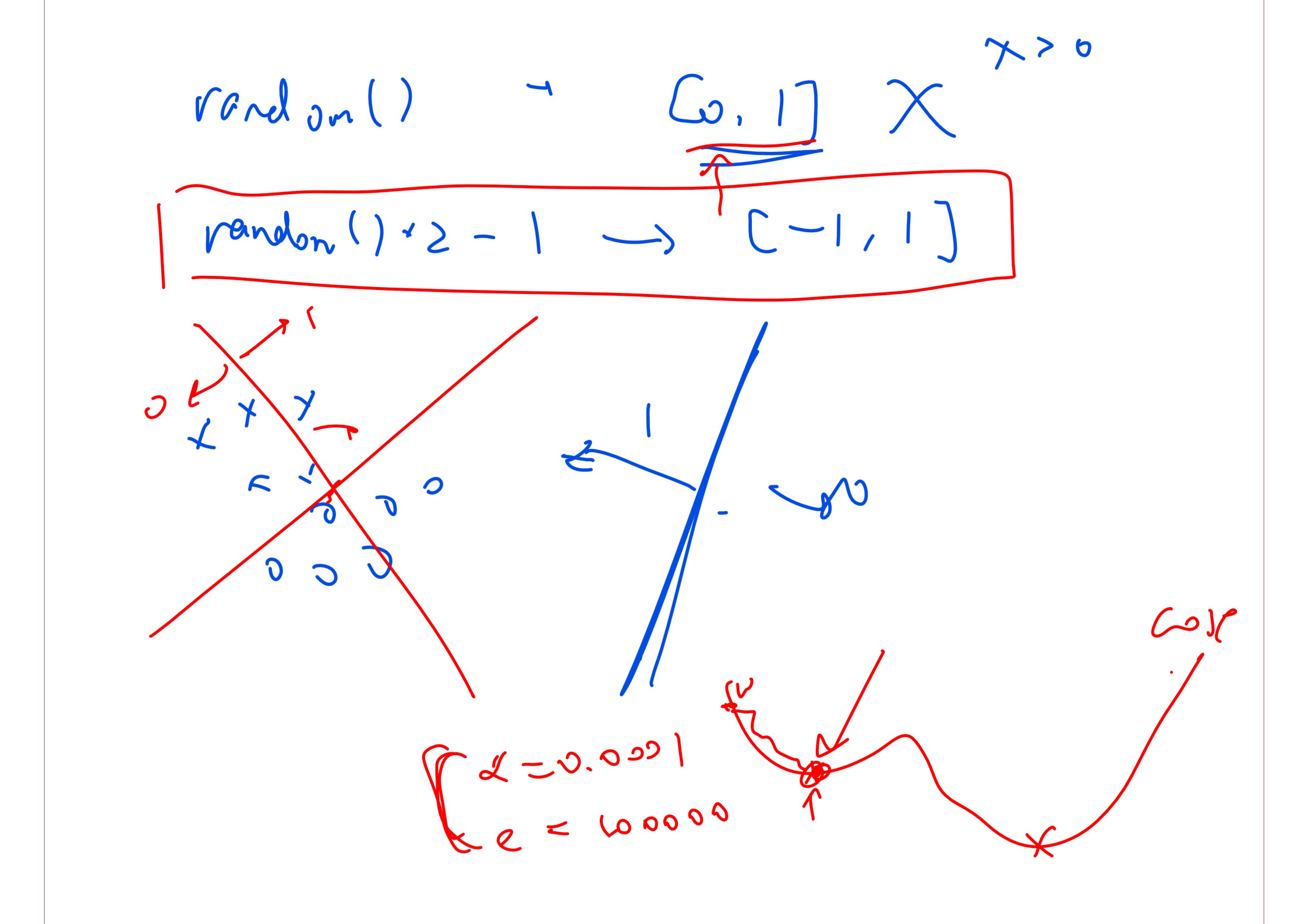
1 NN decision boundary = $\frac{1}{2}(P_5 + P_6) = 5.5$

3 NN ——— = $\frac{1}{2}(P_4 + P_7) = 5.5$

5 NN = $\frac{1}{2}(P_3 + P_8) = 5.5$

P1 $\boxed{w^T x + b}$

$[w \; b] \begin{bmatrix} x \\ 1 \end{bmatrix}$ ← $(x \; 1)$ 가로


random()
[0, 1]
X
X > 0
random() * 2 - 1
[-1, 1]
α = 0.0001
e = 100000
cost