

Q1 Prove Equation (6) in Lecture 5, i.e. $F1 = \frac{TP}{TP + \frac{FN+FP}{2}}$

$F1$ is harmonic mean of precision & recall

$$\text{harmonic mean} = \frac{2xy}{x+y}$$

$$\text{Let } x = \text{precision} = \frac{TP}{TP+FP} \quad ; \quad y = \text{recall} = \frac{TP}{TP+FN}$$

$$\text{Then } F1 = \frac{(2)(\frac{TP}{TP+FP})(\frac{TP}{TP+FN})}{(\frac{TP}{TP+FP} + \frac{TP}{TP+FN})} - \textcircled{1}$$

$$- \textcircled{2}$$

$$\textcircled{1} : (2)(\frac{TP}{TP+FP})(\frac{TP}{TP+FN}) = \frac{2(TP)^2}{(TP+FP)(TP+FN)}$$

$$\textcircled{2} : \frac{TP}{TP+FP} + \frac{TP}{TP+FN} = \frac{TP(TP+FN) + TP(TP+FP)}{(TP+FP)(TP+FN)} = \frac{TP^2 + TP \cdot FN + TP^2 + TP \cdot FP}{(TP+FP)(TP+FN)}$$

$$\frac{\textcircled{1}}{\textcircled{2}} = \frac{2(TP)^2}{2(TP)^2 + TP(FN+FP)} - \frac{2TP}{2TP+FN+FP} = \frac{TP}{TP + \frac{FN+FP}{2}} //$$

Q2 Given $s(x) = \omega^T x + b$ is decision function of SVM

a) Show $\omega \perp$ every vector on the decision boundary

Decision boundary is defined as $s(x) = \omega^T x + b = 0$

Let x_1 and x_2 be two points on the decision boundary

Then $\omega^T x_1 + b = 0$ -① subtract ② from ①

$$\omega^T x_2 + b = 0 - \textcircled{2} \quad \omega^T (x_1 - x_2) = 0$$

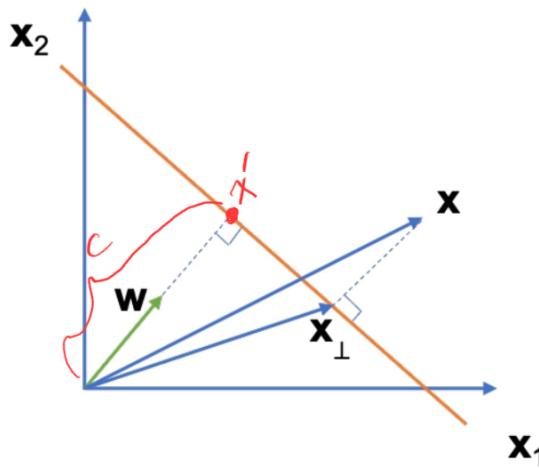
the distance between two

arbitrary points on the decision boundary
= any vector on the db

$\therefore \omega \perp$ any $(x_1 - x_2)$

$\omega \perp$ every vector lying within the decision boundary.

b) Show distance from origin to decision boundary is $\frac{|b|}{\|w\|_2}$



let x' be a point on the DB, such that that the vector \vec{x}' from the origin to DB = Cw

Given x' is on DB, then $w^T x' + b = 0$

$$w^T Cw + b = 0$$

$$Cw^T w + b = 0 \quad \frac{-b}{\|w\|_2^2}$$

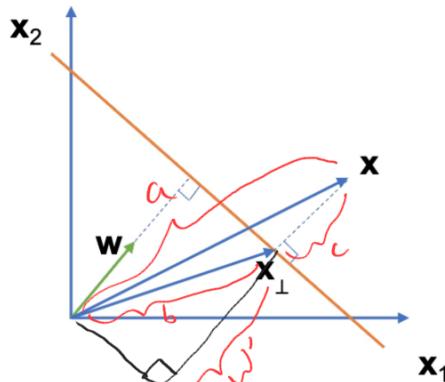
$$C\|w\|_2^2 = -b \quad C = \frac{-b}{\|w\|_2^2}$$

$$\|x'\| = \|Cw\| = C\|w\| = \frac{-b}{\|w\|_2^2} \cdot \|w\| = \frac{-b}{\|w\|_2}$$

The absolute distance of $\vec{x}' = \frac{|b|}{\|w\|_2}$

c) Show distance from x to decision boundary is

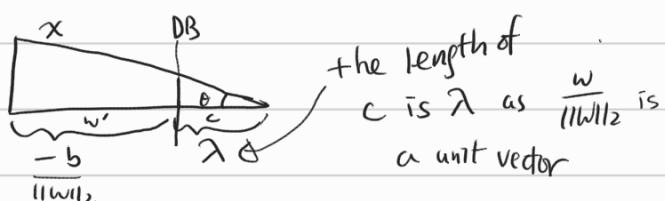
$$\frac{|s(x)|}{\|w\|_2} = \frac{w^T x + b}{\|w\|_2}, \text{ where } x \text{ is arbitrary point.}$$



From the diagram, we see $a = b + c$

$$c = a - b = x \cdot \frac{w}{\|w\|_2}$$

we found the distance of w' in $2b$ which is $\frac{-b}{\|w\|_2}$



$$\frac{\|w' + c\|}{\|x\|} = \cos \theta$$

$$\|w' + c\| = \frac{-b}{\|w\|_2} + \lambda = \cos \theta \|x\|$$

$$= \frac{w^T x}{\|w\|_2 \|x\|} \|x\| = \frac{w^T x}{\|w\|_2}$$

$$\therefore \lambda = \frac{w^T x}{\|w\|_2} + \frac{b}{\|w\|_2}$$

Q3 a) What is t_i^2 ? Ans: 1 because t_i is either 1 or -1

b) Show Equation 13: $b = \frac{1}{N_s} \sum_{i \in S} \left(t_i - \sum_{j=1}^N a_j t_j k(x_i, x_j) \right)$

Equation 8: $t_i (w^T x_i + b) = 1 \Rightarrow \sum_{i \in S} t_i (w^T x_i + b) = 1$, i is the index of support vectors

Equation 12: $w = \sum_{i=1}^N a_i t_i x_i = \sum_{j=1}^N a_j t_j x_j$, change notation, j is the index of instances so not to confuse with equation 8

$$\text{Plug (12) into (8)} \quad \sum_{i \in S} t_i \left(\sum_{j=1}^N a_j t_j \underbrace{x_j^T x_i + b}_{x_j^T x_i = K(x_i, x_j)} \right) = 1$$

$$\Rightarrow \sum_{i \in S} t_i \sum_{j=1}^N a_j t_j k(x_i, x_j) + \sum_{i \in S} t_i b = 1$$

$$\Rightarrow \sum_{i \in S} t_i b = 1 - \sum_{i \in S} t_i \sum_{j=1}^N a_j t_j k(x_i, x_j)$$

$$\Rightarrow \sum_{i \in S} t_i t_i b = \sum_{i \in S} \left[t_i - t_i \sum_{j=1}^N a_j t_j k(x_i, x_j) \right] \quad \begin{matrix} \text{multiply both} \\ \text{sides by } t_i \end{matrix}$$

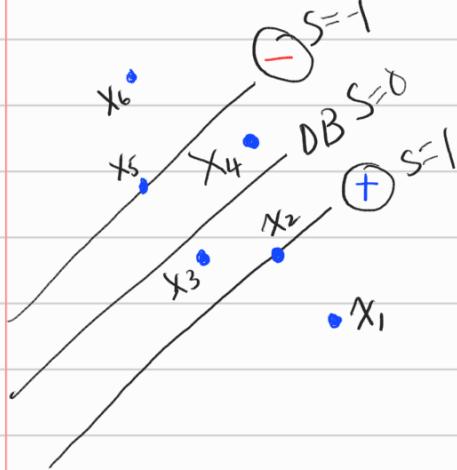
$$\Rightarrow N_s b = \sum_{i \in S} \left[t_i - \sum_{j=1}^N a_j t_j k(x_i, x_j) \right] \quad \begin{matrix} t_i^2 = 1 \end{matrix}$$

$$\Rightarrow b = \frac{1}{N_s} \sum_{i \in S} \left(t_i - \sum_{j=1}^N a_j t_j k(x_i, x_j) \right) //$$

Q4 Show equation 15: $t_i s(x_i) \geq 1 - \varepsilon_i$, $1 \leq i \leq N$ for any pt x_i

Given: ① $\varepsilon_i \left\{ \begin{array}{ll} 0 & \text{if } x_i \text{ on the correct side of the correct margin} \\ |t_i - s(x_i)| & \text{otherwise} \end{array} \right.$

② $t_i s(x_i) \left\{ \begin{array}{ll} 1 & \text{if } x_i \text{ is support vector} \\ > 1 & \text{otherwise} \end{array} \right.$



4 types of points illustrated

in diagram on left =

x_1 — on correct side of correct margin

x_2 — on correct margin

x_3 — classified correctly, but margin violation

x_4, x_5, x_6 — misclassified points

support vector for soft margin case

/ $t_i = 1$ for positive case

	SV	$s(x_i)$	$t_i s(x_i)$	ε_i	$1 - \varepsilon_i$	$LHS \geq RHS?$
x_1	N	> 1	> 1	0	1	✓
x_2	Y	= 1	= 1	0	1	✓
x_3	Y	$0 < ? < 1$	$= 1$	$0 < ? < 1$	$0 < ? < 1$	✓
x_4, x_5, x_6	Y	< 0	= 1	> 1	< 0	✓

L.H.S.
of
equation 15

R.H.S.
of
equation 15

$\therefore t_i s(x_i) \geq 1 - \varepsilon_i$,
for $1 \leq i \leq N$,
for any points x_i

