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

F1 is harmonic mean of precision & recall

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

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

Then $F1 = \dfrac{(2)\left(\frac{TP}{TP+FP}\right)\left(\frac{TP}{TP+FN}\right) \text{ — ①}}{\left(\frac{TP}{TP+FP} + \frac{TP}{TP+FN}\right) \text{ — ②}}$

① : $(2)\left(\dfrac{TP}{TP+FP}\right)\left(\dfrac{TP}{TP+FN}\right) = \dfrac{2(TP)^2}{\cancel{(TP+FP)(TP+FN)}}$

② : $\dfrac{TP}{TP+FP} + \dfrac{TP}{TP+FN} = \dfrac{TP(TP+FN) + TP(TP+FP)}{(TP+FP)(TP+FN)} = \dfrac{TP^2 + TP \cdot FN + TP^2 + TP\,FP}{\cancel{(TP+FP)(TP+FN)}}$

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

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

a) show $w \perp$ every vector on the decision boundary

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

Let x_1 and x_2 be two points on the decision boundary

Then $w^T x_1 + b = 0$ — ①   subtract ② from ①

$w^T x_2 + b = 0$ — ② $w^T \cdot (\underbrace{x_1 - x_2}_{}) = 0$

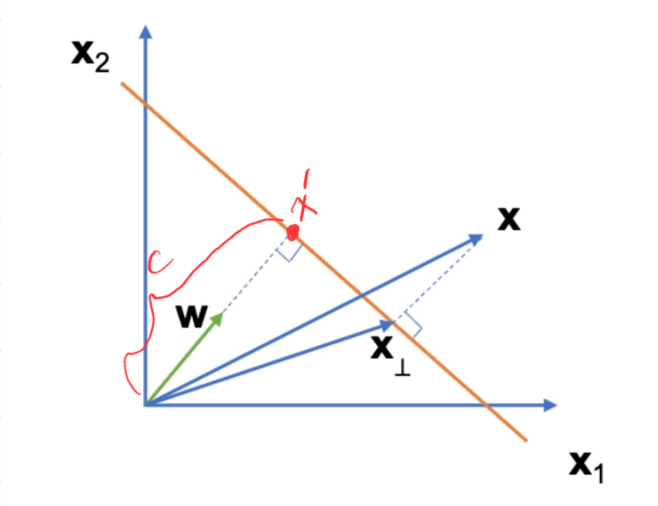
the distance between two arbitrary points on the decision boundary

= any vector on the db

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

$w \perp$ 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 $= cw$

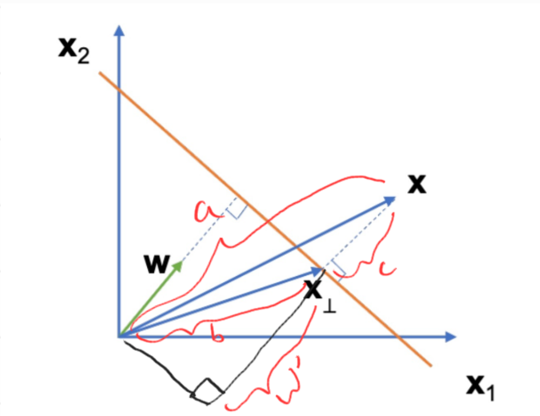
Given x' is on DB,
then $w^T x' + b = 0$

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

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

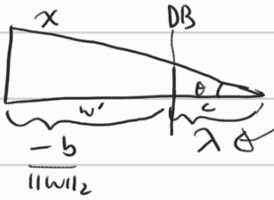
The absolute distance of $\vec{x}' = \frac{|b|}{||w||_2}$ //

c) Show distance from x to decision boundary is $\frac{|s(x)|}{||w||_2} = \frac{w^T x + b}{||w||_2}$, where x is arbitrary point.



From the diagram, we see
$$a = b + c$$
$$c = a - b = \lambda \cdot \frac{w}{||w||_2}$$

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



the length of c is λ as $\frac{w}{||w||_2}$ is a unit vector

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

$$||w' + c|| = \frac{-b}{||w||_2} + \lambda = \cos\theta \, ||x||$$
$$= \frac{w \cdot x}{\cancel{||x||}\,||w||} \cancel{||x||} = \frac{w^T x}{||w||_2} \qquad \therefore \lambda = \frac{w^T x}{||w||_2} + \frac{b}{||w||_2} \text{ //}$$

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 : $t_i(w^T x_i + b) = 1 \;=\; \sum_{i \in S} t_i (w^T x_i + b) = 1$, 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

Plug (12) into (8) $\sum_{i \in S} t_i \left(\sum_{j=1}^{N} a_j t_j x_j^T x_i + b \right) = 1$ $\boxed{\begin{aligned} x_j^T x_i &= x_i^T x_j \\ &= k(x_i, x_j) \end{aligned}}$

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

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

multiply both sides by t_i

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

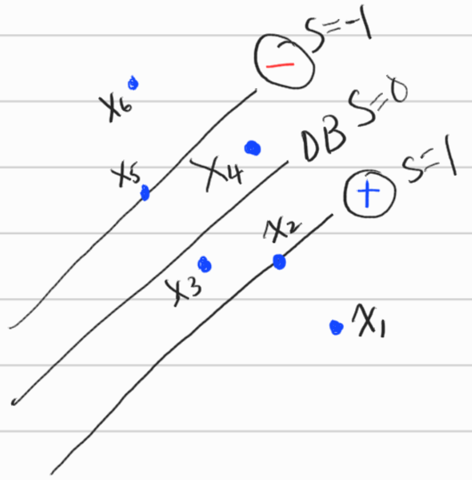
$t_i^2 = 1$

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

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

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 $\varepsilon_i \begin{cases} 0 & \text{if } x_i \text{ on the correct margin or correct side of the correct margin} \\ |t_i - S(x_i)| & \text{otherwise} \end{cases}$



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 ≥ RHS?
x_1	N	>1	>1	0	1	✓
x_2	Y	$=1$	$=1$	0	1	✓
x_3	Y	$0<S(x_3)<1$	$S(x_3)$	$1-S(x_3)$	$S(x_3)$	✓
x_4	Y	$-1<S(x_4)<0$	$S(x_4)$	$1-S(x_4)$	$S(x_4)$	✓
x_5	Y	-1	-1	2	-1	✓
x_6	Y	$S(x_6)<-1$	$S(x_6)$	$1-S(x_6)$	$S(x_6)$	✓

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 //