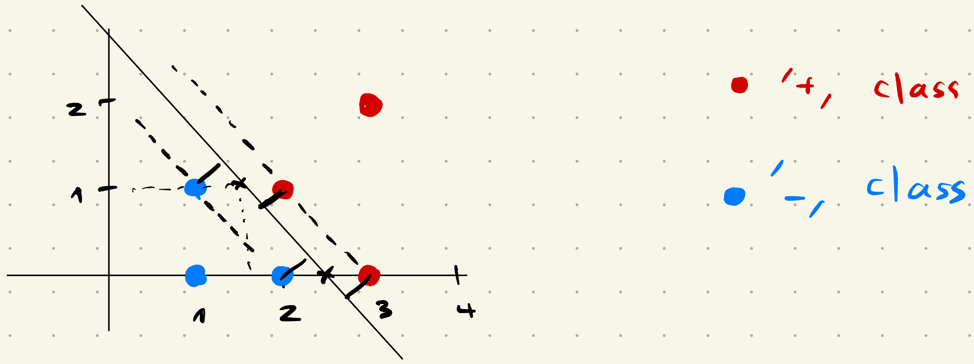
2




b) Inspection: The line above, passing from (2.5,0) and (1.5,1) satisfies $x_1 + x_2 = 2.5$. Thus, parametrically our separating hyperplane is:

$$w^T z + b = 0, \text{ where}$$

$w_0 = w_1 = 1$, $b = -2.5$, $|w| = \sqrt{2}$

$$\delta^{((2,1))} = \delta^{((2,0))} = \delta^{((1,1))} = \delta^{((3,0))} = \frac{0.5}{\sqrt{2}}$$

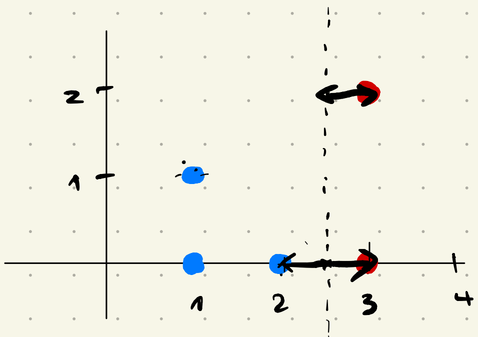
$$\delta^{((1,0))} = \frac{1.5}{\sqrt{2}}$$

$$\delta^{((3,2))} = \frac{2.5}{\sqrt{2}}$$

Support vectors: (2,0), (1,1), (2,1), (3,0). Margin $= \frac{0.5}{\sqrt{2}}$

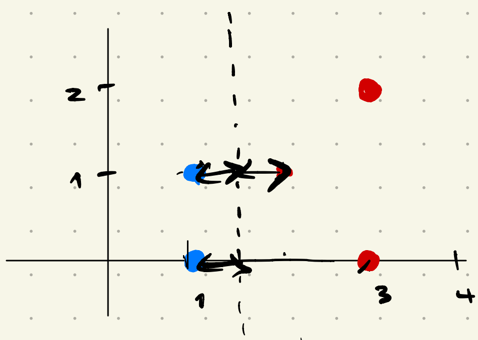
c) Remove (2,0), (2,1) margin increases

" (1,1), (3,0) " stays the same



without (2,1)

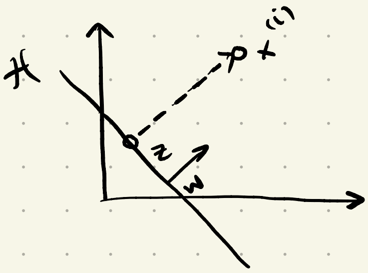
margin $= 1 > \frac{0.5}{\sqrt{2}}$



without (2,0)

margin $= 1 > \frac{0.5}{\sqrt{2}}$

3




Let $\mathcal{H} = \{x \mid w^T x + b = 0\}$ denote the hyperplane defined through $w$ and $b$.

Let $x^{(i)}$ a training example. Let $z$ be the projection of $z$ onto $\mathcal{H}$. Since $x^{(i)} - z \perp \mathcal{H}$ and $w \perp \mathcal{H}$ $\exists \alpha \in \mathbb{R}$:

$x^{(i)} - z = \alpha w$. Note that

$$|\alpha| = \frac{|x^{(i)} - z|}{|w|} \iff \alpha = \pm \frac{|x^{(i)} - z|}{|w|} \iff \alpha = \pm \frac{\delta^{(i)}}{|w|}$$

$$\frac{\pm(x^{(i)} - z)}{\delta^{(i)}} = \frac{w}{|w|}.$$ Note that $\alpha = \frac{\delta^{(i)}}{|w|}$ when $y^{(i)} = +1$

and $\alpha = -\frac{\delta^{(i)}}{|w|}$ when $y^{(i)} = -1$.

$$\frac{y^{(i)}(x^{(i)} - z)}{\delta^{(i)}} = \frac{w}{|w|} \overset{\cdot w^T}{\iff} y^{(i)} \frac{(w^T x^{(i)} - w^T z)}{\delta^{(i)}} = \frac{w^T w}{|w|}$$

$$\frac{y^{(i)}(w^T x^{(i)} + b)}{\delta^{(i)}} = \frac{|w|^{\cancel{2}}}{\cancel{|w|}},$$ where we used $z \in \mathcal{H}$.

Thus, $$\delta^{(i)} = \frac{y^{(i)}(w^T x^{(i)} + b)}{|w|}$$ ∎

④ Hard margin: $\min_{w,b} \; w^T w$ (*)

s.t. $\forall i \;\; y_i(w^T x_i + b) \geq 1$

a) If $w^*, b^*$ optimal for (*).

Claim: $\exists \; i$ s.t. $y_i(w^{*T} x_i + b^*) = 1$

Proof: Suppose that $y_i(w^{*T} x_i + b^*) > 1$, $\forall i$. Then, $\exists \varepsilon > 0$:

$y_i(w^{*T} x_i + b^*) \geq 1 + \varepsilon, \; \forall i$

$y_i\left[\dfrac{w^{*T}}{1+\varepsilon} x_i + \dfrac{b^*}{1+\varepsilon}\right] \geq 1, \; \forall i$. Denote $\tilde{w} = \dfrac{w^*}{1+\varepsilon}$, $\tilde{b} = \dfrac{b^*}{1+\varepsilon}$

$y_i(\tilde{w}^T x_i + \tilde{b}) \geq 1, \forall i$ But since $(w^*, b^*)$ optimal for (*), then $w^{*T} w^* = |w^*|^2 \leq \tilde{w}^T \tilde{w} = |\tilde{w}|^2$, or

$$|w^*| \leq |\tilde{w}| = \frac{|w^*|}{1+\varepsilon},$$

which entails $1+\varepsilon \leq 1$ which cannot hold as $\varepsilon > 0$.

By contradiction the claim holds. So, $\exists \, i_0$ s.t.

$y_{i_0}(w^{*T} x_{i_0} + b^*) = 1 \;\Rightarrow\; \dfrac{y_{i_0}(w^{*T} x_{i_0} + b^*)}{|w^*|} = \dfrac{1}{|w^*|} \Rightarrow$

$\delta^{(i_0)} = \dfrac{1}{|w^*|}$, where $\delta^{(i_0)}$ is the dist of $x_{i_0}$ from the optimal hyperplane $\{z \mid w^{*T} z + b^* = 0\}$.

4) b) $(z, d)$ separating hyperplane with margin $M$

$z' = \frac{z}{|z| \cdot M}$, $d' = \frac{d}{|z| \cdot M}$

$|z'| = \frac{1}{M}$.

$\forall i \quad \frac{y_i(z^T x_i + d)}{|z|} \geq M \Leftrightarrow \forall i \quad \frac{y_i\,(z^T x_i + d)}{|z| \cdot M} \geq 1$

$\forall i \quad y_i\left(\frac{z^T}{|z| \cdot M} \cdot x_i + \frac{d}{|z| \cdot M}\right) \geq 1 \Leftrightarrow \forall i \quad y_i\,((z')^T x_i + d') \geq 1$,

so $(z', d')$ provides a feasible solution.

c) Since $(w^*, b^*)$ is optimal for $(*)$.

$|w^*|^2 \leq |z'|^2 \Leftrightarrow |w^*| \leq |z'| \Leftrightarrow \frac{1}{|w^*|} \geq \frac{1}{|z'|} = M$

$$\frac{1}{|w^*|} \geq M$$

$\frac{1}{|w^*|}$ was the margin for $(w^*, b^*)$ from a) and

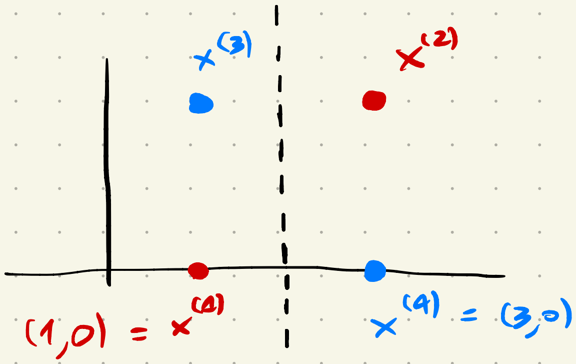
$M$ is the margin for $(z, d)$, so we have proven

the statement.

5 a) Suppose there is a solution $(w_0, w_1, b)$. Then, it must satisfy:

$$\left.\begin{array}{ll} (1) & w_0 + b \geq 1 \\ (2) & 3w_0 + 2w_1 + b \geq 1 \\ (3) & -w_0 - 2w_1 - b \geq 1 \\ (4) & -3w_0 - b \geq 1 \end{array}\right] \begin{array}{l} (1)+(4): \quad -2w_0 \geq 2 \Rightarrow w_0 \leq -1 \\ (2)+(3): \quad 2w_0 \geq 2 \Rightarrow w_0 \geq 1 \end{array}$$

b) $$\min_{w,b,\xi} \; w^T w + C\sum_{i=1}^{4} \xi^{(i)} \quad \text{s.t.} \quad \begin{array}{l} \xi_i \geq 0, \forall i \\ y_i(w^T x_i + b) \geq 1 - \xi_i, \forall i \end{array} \qquad (**)$$




This is a possible approach (not optimal). Let

$$\left.\begin{array}{l} w_0 = 1 \\ w_1 = 0 \\ b = -2 \end{array}\right\} \quad \text{(see the dotted line above)}$$

Constraints:

$$\left.\begin{array}{l} w_0 + b \geq 1 - \xi^{(1)} \\ 3w_0 + 2w_1 + b \geq 1 - \xi^{(2)} \\ -w_0 - 2w_1 - b \geq 1 - \xi^{(3)} \\ -3w_0 - b \geq 1 - \xi^{(4)} \end{array}\right\} \begin{array}{ll} -1 \geq 1 - \xi^{(1)} & \Rightarrow \xi^{(1)} \geq 2 \\ 1 \geq 1 - \xi^{(2)} & \Rightarrow \xi^{(2)} \geq 0 \\ 1 \geq 1 - \xi^{(3)} & \Rightarrow \xi^{(3)} \geq 0 \\ -1 \geq 1 - \xi^{(4)} & \Rightarrow \xi^{(4)} \geq 2 \end{array}$$

Thus, for $C = 1$ in $(**)$, we have that the min occurs at:

$$(w, b, \xi) = (1, 0, -2, 2, 0, 0, 2)$$