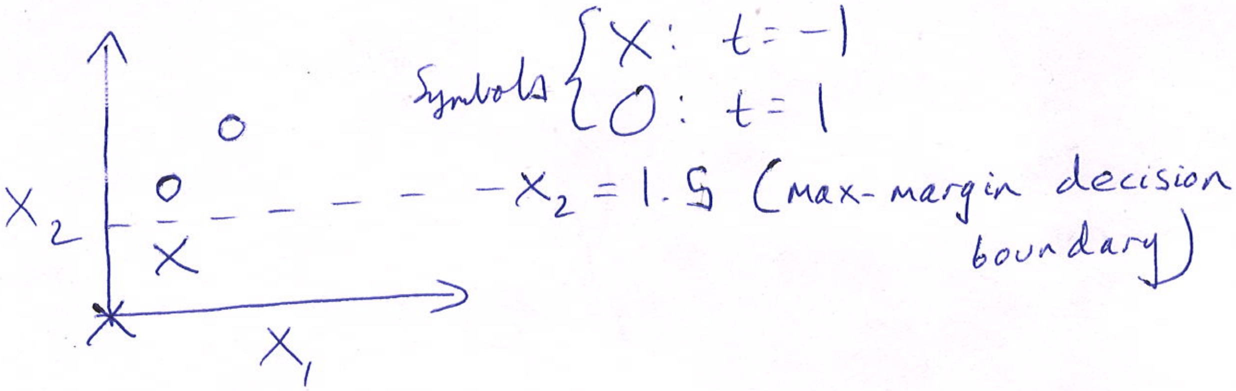
MINOR II

1. (a)



(b) $$\mathcal{L}(\underline{w}, w_0, \underline{\mu}) = \frac{\|\underline{w}\|^2}{2} + \mu_1\big(1-(w_1+w_2+w_0)(-1)\big)$$

$$+ \mu_2\big(1-(w_0)(-1)\big) + \mu_3\big(1-(2w_1+3w_2+w_0)\big)$$

$$+ \mu_4\big(1-(w_1+2w_2+w_0)\big)$$

(c) $$\frac{\partial \mathcal{L}}{\partial w_0} = \mu_1+\mu_2-\mu_3-\mu_4 \qquad -(1)$$

$$\frac{\partial \mathcal{L}}{\partial w_1} = w_1+\mu_1-2\mu_3-\mu_4 \qquad -(2)$$

$$\frac{\partial \mathcal{L}}{\partial w_2} = w_2+\mu_1-3\mu_3-2\mu_4 \qquad -(3)$$

Setting these to 0:

$$\hat{w}_1 = -(\mu_1-2\mu_3-\mu_4) \qquad \text{from } (2)$$

$$\hat{w}_2 = -(\mu_1-3\mu_3-2\mu_4) \qquad \text{from } (3)$$

Rewriting $\mathcal{L}$:

$$\mathcal{L} = \frac{1}{2}(w_1^2+w_2^2) + \mu_1+\mu_2+\mu_3+\mu_4$$

$$+ \omega_1 (\mu_1 - 2\mu_3 - \mu_4) + \omega_2 (\mu_1 - 3\mu_3 - 2\mu_4)$$

$$+ \omega_0 \cancel{(\mu_1 + \mu_2 - \mu_3 - \mu_4)} \to 0 \text{ (from (1))}$$

$$\Rightarrow \mathcal{L} = \frac{1}{2}\omega_1^2 + \frac{1}{2}\omega_2^2 + \sum_{i=1}^{4} \mu_i + \omega_1(\mu_1 - 2\mu_3 - \mu_4)$$

$$+ \omega_2(\mu_1 - 3\mu_3 - 2\mu_4)$$

Now plug in $\hat{\omega}_1, \hat{\omega}_2$:

$$\Rightarrow \underset{\text{(dual)}}{\tilde{\mathcal{L}}(\underline{\mu})} = \sum_{i=1}^{4} \mu_i - \frac{1}{2}(\mu_1 - 2\mu_3 - \mu_4)^2$$

$$- \frac{1}{2}(\mu_1 - 3\mu_3 - 2\mu_4)^2$$

(d) Clearly the points $(0,0)$ and $(2,3)$ are non-SVs, so we must have $\mu_2 = \mu_3 = 0$. This gives:

$$\tilde{\mathcal{L}}(\underline{\mu}) = \mu_1 + \mu_4 - \frac{1}{2}(\mu_1 - \mu_4)^2 - \frac{1}{2}(\mu_1 - 2\mu_4)^2$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \mu_1} = 1 - (\mu_1 - \mu_4) - (\mu_1 - 2\mu_4)$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \mu_4} = 1 + (\mu_1 - \mu_4) + 2(\mu_1 - 2\mu_4)$$

Setting these to 0:

$$1 - 2\mu_1 + 3\mu_4 = 0$$
$$1 + 3\mu_1 - 5\mu_4 = 0$$

Solving, we get $\hat{\mu}_1 = 8$, $\hat{\mu}_4 = 5$

(e) Plugging back $\hat{\mu}_1, \hat{\mu}_4$:

$$\hat{w}_1 = -(8-5) = -3$$

$$\hat{w}_2 = -(8 - 2\times 5) = 2$$

But for drawn decision boundary:

$$x_2 = 1.5 \quad \text{or} \quad x_2 - 1.5 = 0$$

So $w_1 = 0$, $w_2 = 1$, $w_0 = -1.5$

It doesn't match, because when maximising the dual we did not enforce any constraints.

Actually, the dual needs to be maximised with the constraints

$$-\sum_{i=1}^{N} \mu_i t_i = 0 \quad (\text{Eqn. } ①)$$

to get the correct SVM solution.

2. (a) i) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

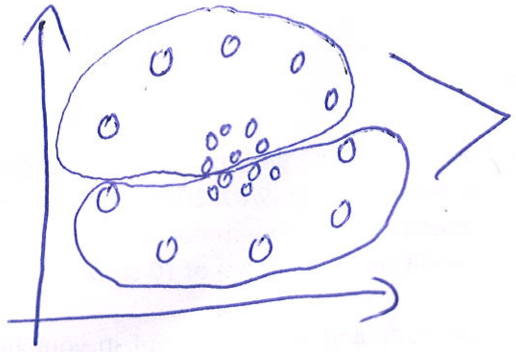
ii) $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

(b) Case ii) variance in 1^{st} component is more

Case i), same variance in both directions

(c) Same in both cases; data is 2-D, so 2 components always capture 100% of variance.

3. (a)



K-means (K=2) will give something like this; but natural clusters are inner and outer groups

(b) Not a convergence problem; the objective fn. itself fails here. The natural clusters do not accord with the idea of 'tightness' or closeness to a cluster centre which K-means relies on. A clustering approach instead based on progressively grouping nearby points (without assuming a proximal cluster centre) would work better here: e.g., single-linkage clustering.