HOMEWORK # 3 | Max Sokolich

$\boxed{\text{Example 12.2}}$

12.29: $$\min_{\beta_0, \alpha} \sum_{i=1}^{N} (1 - y_i f(x_i))_+ + \frac{\lambda}{2} \alpha^T \vec{K} \alpha$$

12.25: $$\min_{\beta_0, \beta} \sum_{i=1}^{N} [1 - y_i f(x_i)]_+ + \frac{\lambda}{2} \|\beta\|^2$$

→ solution is the same for a particular kernel

define $\boxed{K(x,y) = h(x)^T h(y)}$ and let $\beta = \sum_{i=1}^{N} \alpha_i h(x_i)$

$$f(x) = \beta_0 + \sum_{i=1}^{N} \alpha_i h(x)^T h(x_i)$$

$$= \beta_0 + h(x)^T \sum_{i=1}^{N} \alpha_i h(x_i)$$

$$= \beta_0 + h(x)^T \beta$$

also $$\|\beta\|^2 = \beta^T \beta = \left(\sum \alpha_i h(x_i)\right)^T \left(\sum \alpha_i h(x_i)\right)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j K(x_i, x_j)$$

$$= \alpha^T K \alpha$$

Problem # 2:

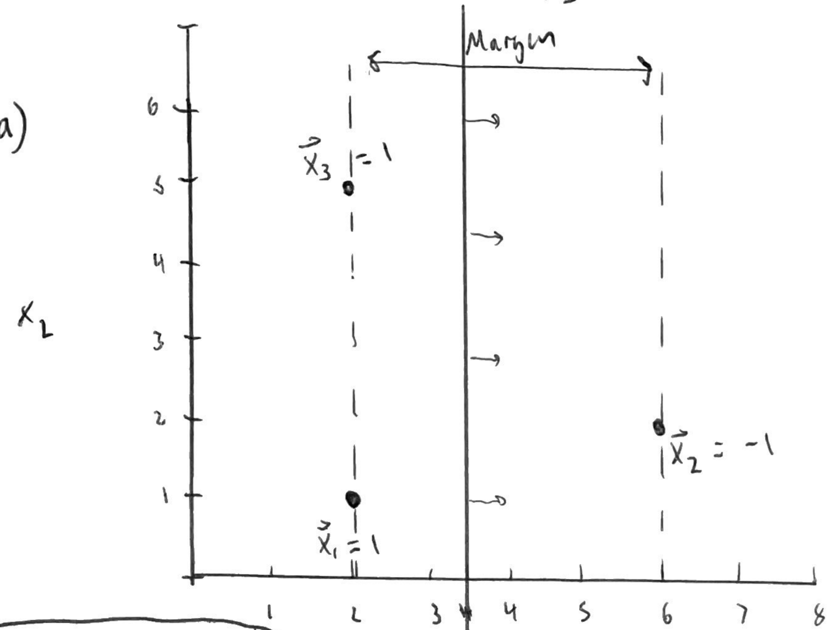
consider: $\vec{x}_1 = (2,1)$ $y_1 = 1$

$\vec{x}_2 = (6,2)$ $y_2 = -1$

$\vec{x}_3 = (2,5)$ $y_3 = 1$

NEED TO SHOW THERE ARNT ANY OTHER ~~SOLUTIONS~~ LINE that seperates the points w/ a bigger margin

a)



whats the line $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0$?

let each of thes points be support vectors on the margin

$$\boxed{X_1 = 4}$$

$$\boxed{\begin{array}{l} \text{if } X_1 < 4 \Rightarrow 1 \\ \text{if } x_1 > 4 \Rightarrow -1 \end{array}}$$

x_1

I mean the line to be at $X = 4$

b) prove $x = 4$ is the unique solution to $\min_{\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^2} \frac{1}{2} \|\beta\|_2^2$

subject to $y_i(x_i^T \beta + \beta_0) \geq 1$

So $\beta_1 = 1$

$\beta_0 = -4$

$\beta_2 = 0$

point 1: $1\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} [\beta_1 \ \beta_2] + \beta_0\right) > 1$

$2\beta_1 + \beta_2 + \beta_0 > 1$ ——— 1

point 2: $-1\left(\begin{bmatrix} 6 \\ 2 \end{bmatrix} [\beta_1 \ \beta_2] + \beta_0\right) > 1$

$-6\beta_1 - 2\beta_2 + \beta_0 > 1$ ——— 2

point 3 $1\left(\begin{bmatrix} 2 \\ 5 \end{bmatrix} [\beta_1 \ \beta_2) + \beta_0\right) > 1$

$2\beta_1 + 5\beta_2 + \beta_1 > 1$ ——— 3

Forget this neg 4

from hint b

$$\left.\begin{aligned} 2\beta_1 + \beta_2 + \beta_0 &> 1 \\ -6\beta_1 - 2\beta_2 + \beta_0 &> 1 \\ 2\beta_1 + 5\beta_2 + \beta_0 &> 1 \end{aligned}\right\} \begin{bmatrix} 2 & 1 & 1 \\ -6 & -2 & 1 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

no solution to this

$$\frac{1}{2}\|\beta\|^2 = \frac{1}{2}\left(\sqrt{\beta_1^2 + \beta_2^2}\right)^2 = \frac{1}{2}\beta_1^2 + \frac{1}{2}\beta_2^2$$

$$\min\left(\frac{1}{2}\beta_1^2 + \frac{1}{2}\beta_2^2\right), \quad \text{also } \|\beta\| = \frac{1}{M}$$

when observation $M = 2$

generate standard support vector classifier

$$\left.\begin{aligned} \min\|\beta\| \quad \text{subject to} \quad & y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad \sum_{i=1}^{n} \xi_i \leq C \end{aligned}\right\} \quad \xi_i = 2 \text{ for all } i$$

$i=1$: $2\beta_1 + \beta_2 + \beta_0 \cancel{\geq 1 - 2} = 1$

$i=2$: $-6\beta_1 - 2\beta_2 + \beta_0 \cancel{\geq 1 - 2} = 1$

$i=3$: $2\beta_1 + 5\beta_2 + \beta_0 \cancel{\geq 1 - 2} = 1$

$$\boxed{\begin{aligned} \beta_0 &= -\frac{1}{2} \\ \beta_2 &= 2 \\ \beta_1 &= 0 \end{aligned}}$$

$$-\frac{1}{2} + 0 + 2x_2 = 0$$

$$\boxed{x_1 - 4 = 0}$$

First look at lagrangian primal function

$$\cancel{L_p = \left(\frac{1}{2}\beta_1^2 + \frac{1}{2}\beta_2^2\right) - \Big[\alpha_1(2\beta_1 + \beta_2 + \beta_0 - 1) + \alpha_2(-6\beta_1 - 2\beta_2 + \beta_0 - 1) + \alpha_3(2\beta_1 + 5\beta_2 + \beta_1 - 1)\Big]}$$

$$2x_2 = \frac{1}{2}$$

$$x_2 = \frac{1}{4}$$

$$\frac{\partial L_P}{\partial \beta} = \beta - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$$

$$(\beta_1, \beta_2) = (\alpha_1)(1)(2, 1) + (\alpha_2)(-1)(6, 2) + (\alpha_3)(1)(2, 5)$$

$$(\beta_1, \beta_2) = (2\alpha_1, \alpha_1) + (-6\alpha_2, -2\alpha_2) + (2\alpha_3, 5\alpha_3)$$

$$\boxed{\begin{aligned} \beta_1 &= 2\alpha_1 + 6\alpha_2 + 2\alpha_3 \\ \beta_2 &= \alpha_1 - 2\alpha_2 + 5\alpha_3 \end{aligned}}$$

β_0

how to find $\alpha_1, \alpha_2, \alpha_3$ by hand

---

maybe I need to show uniqueness given $\beta_0 = -4$, $\beta_1 = 0$, $\beta_2 = 1$

how ~~the~~ min $\frac{1}{2}\|\beta\|^2$ is subject to

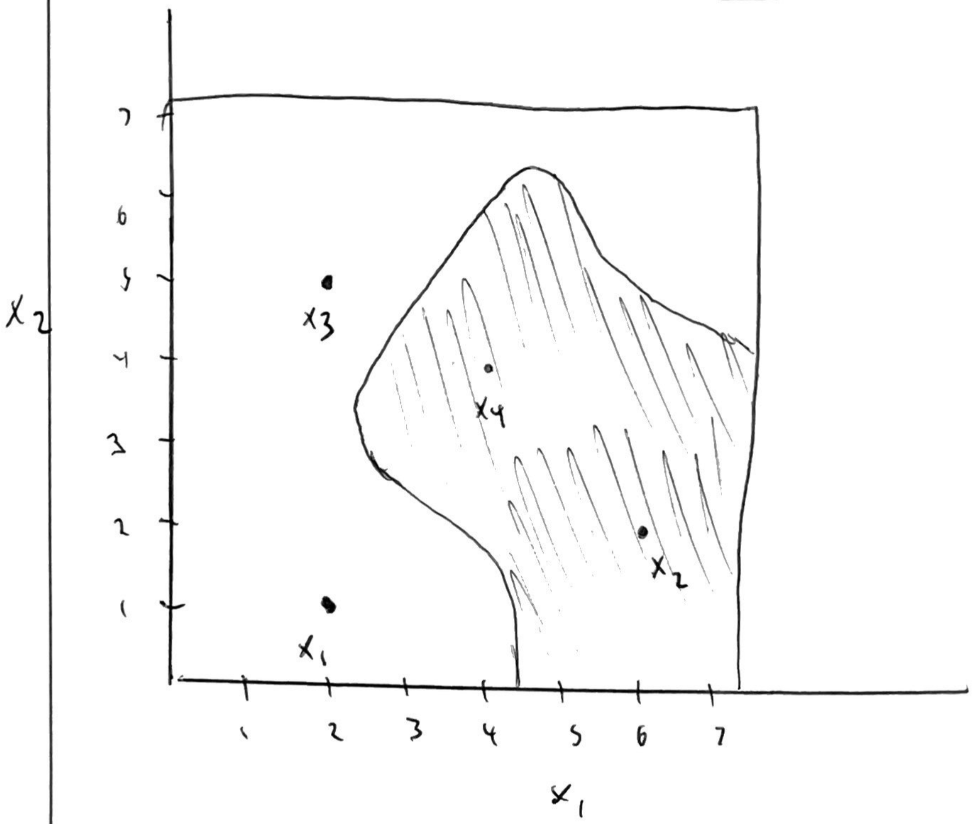
$$\left.\begin{aligned} y_1(x_1 - 4) &\geq 1 \\ y_2(x_2 - 4) &\geq 1 \\ y_3(x_3 - 4) &\geq 1 \end{aligned}\right\} \quad L(b, \alpha) = \frac{1}{2}\|\beta\|^2 - \alpha_1(y_1(x_1 - 4) - 1) - \alpha_2(y_2(x_2 - 4) - 1) - \alpha_3(y_3(x_3 - 4) - 1)$$

deriv gives $\rightarrow \beta = \sum \alpha_i y_i x_i$

wrt $\alpha \longrightarrow y_i x_i \beta - 1 = 0$ plugging β and solving for α gives

$\alpha_1 = \alpha_3 = \frac{1}{4}$, $\alpha_2 = 0 \longrightarrow \beta = (0.5, 1.25)$

c) from python, let $\boxed{x_4 = (4,4)}$ $y_4 = -1$



$= -1$

$= 1$

the kernel function I used was radial basis function

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$$

which in python can be written as

$$\|x - x'\|^2 = \|x\|^2 + \|x'\|^2 - 2x^T \cdot x'$$

RBF seemed to be the most common kernel