HW 5

IDs:

- 206337768
- 311515381

(1) a. $(XY+1)^3 = (XY)^3 + 3(XY)^2 + 3XY + 1 = \left[(X_1Y_1+X_2Y_2)^3 + 3\cdot(X_1Y_1+X_2Y_2)^2 + 3(X_1Y_1+X_2Y_2) + 1\right] =$

$= (X_1Y_1)^3 + 3\cdot(X_1Y_1)^2\cdot(X_2Y_2) + 3\cdot(X_1\cdot Y_1)(X_2Y_2)^2 + (X_2Y_2)^3 + 3\left[(X_1Y_1)^2 + 2(X_1Y_1X_2Y_2) + (X_2Y_2)^2\right] + 3X_1Y_1 + 3X_2Y_2 + 1 =$

$= (X_1Y_1)^3 + 3(X_1Y_1)^2\cdot(X_2Y_2) + 3\cdot(X_1Y_1)\cdot(X_2\cdot Y_2)^2 + (X_2Y_2)^3 + 3(X_1Y_1)^2 + 6(X_1Y_1X_2Y_2) + 3(X_2Y_2)^2 + 3X_1Y_1 + 3X_2Y_2 + 1$

$\varphi(x) = \{X_1^3, X_2^3, \sqrt{3}X_1^2X_2, \sqrt{3}X_1X_2^2, \sqrt{3}X_1^2, \sqrt{3}X_2^2, \sqrt{6}X_1X_2, \sqrt{3}X_1, \sqrt{3}X_2, 1\}$

b. The full rational variety of order 3

c. There are 10 multiplications using $\varphi(x)\cdot\varphi(y)$ and 4 multiplications using $K(x,y)$, hence we save 6 multiplications.

(2) $f(x,y) = 2x-y$, find minimum and maximum points for f under the constraints

$g(x,y) = \frac{x^2}{4} + y^2 = 1 \Rightarrow L(x,y) = 2x - y + \lambda\left(\frac{x^2}{4} + y^2 - 1\right)$

(i) $\frac{\partial}{\partial x}L(x,y) = 2 + \frac{\lambda x}{2} = 0 \Rightarrow 2 = -\frac{\lambda x}{2} \Rightarrow \lambda x = -4 \Rightarrow \lambda = -\frac{4}{x}$

(ii) $\frac{\partial}{\partial y}L(x,y) = -1 + 2\lambda y = 0 \Rightarrow 2\lambda y = 1 \overset{(i)}{\Rightarrow} 2\cdot\frac{-4}{x}\cdot y = 1 \Rightarrow \frac{-8y}{x} = 1 \Rightarrow x = -8y \Rightarrow x^2 = 64y^2$

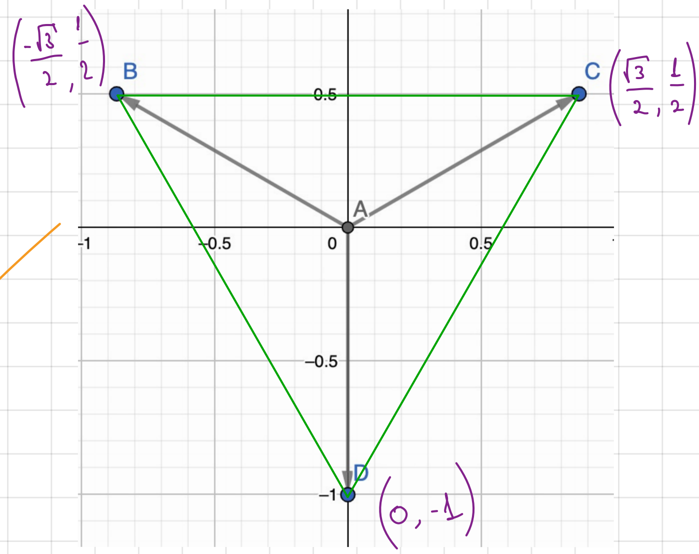
(iii) $\frac{\partial}{\partial \lambda}L(x,y) = \frac{x^2}{4} + y^2 = 1 \overset{(ii)}{\Rightarrow} \frac{64y^2}{4} + y^2 = 1 \Rightarrow 16y^2 + y^2 = 1 \Rightarrow 17y^2 = 1 \Rightarrow y = \pm\sqrt{\frac{1}{17}}$

Therefore we get $\longrightarrow$ $y = -\frac{1}{\sqrt{17}}, x = \frac{8}{\sqrt{17}}$ $\cup$ $y = \frac{1}{\sqrt{17}}, x = -\frac{8}{\sqrt{17}}$, now check for the min & max:

- $f\left(\frac{8}{\sqrt{17}}, \frac{-1}{\sqrt{17}}\right) = 2\cdot\frac{8}{\sqrt{17}} + \frac{1}{\sqrt{17}} = \frac{17}{\sqrt{17}} = \sqrt{17}$
- $f\left(\frac{-8}{\sqrt{17}}, \frac{1}{\sqrt{17}}\right) = 2\cdot\frac{-8}{\sqrt{17}} - \frac{1}{\sqrt{17}} = \frac{-17}{\sqrt{17}} = -\sqrt{17}$

Therefore minimum point is $\left(\frac{-8}{\sqrt{17}}, \frac{1}{\sqrt{17}}\right)$ and maximum point is $\left(\frac{8}{\sqrt{17}}, \frac{-1}{\sqrt{17}}\right)$

(3) Example of one origin-centered upright equilateral triangle:



upright

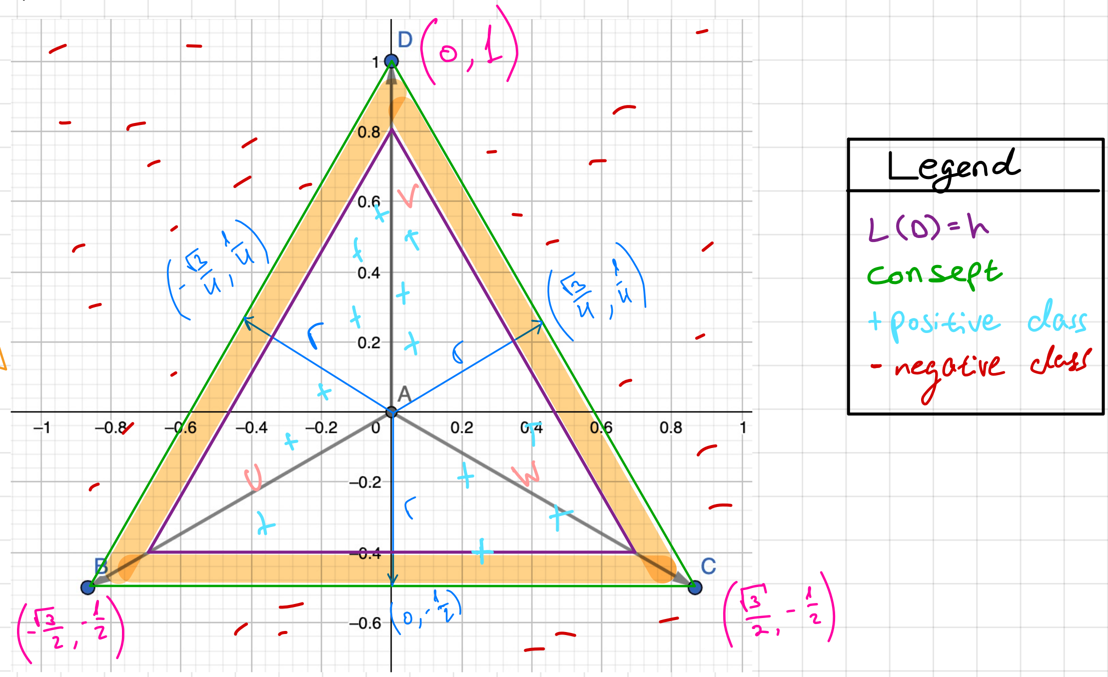
$$\vec{DC} = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) - (0, 1) = \left(\frac{\sqrt{3}}{2}, -\frac{3}{2}\right)$$

$$\vec{r} \perp \vec{DC} \Rightarrow (r_1, r_2) \cdot \left(\frac{\sqrt{3}}{2}, -\frac{3}{2}\right) = 0 \Rightarrow \frac{\sqrt{3}}{2} r_1 - \frac{3}{2} r_2 = 0 \Rightarrow \frac{\sqrt{3}}{2} r_1 = \frac{3}{2} r_2 \Rightarrow r_1 = \sqrt{3}\, r_2$$

$\vec{r} = (\sqrt{3}\, r_2, r_2) \Rightarrow y = \frac{1}{\sqrt{3}} x \Rightarrow$ DC is perpendicular to r, therefore the slope is $-\sqrt{3}$ and the line equation DC is $y = -\sqrt{3}\, x + 1$

the intersection of $\vec{r}$ and DC is $\frac{1}{\sqrt{3}} x = -\sqrt{3}\, x + 1 \Rightarrow x = \frac{\sqrt{3}}{4}, y = \frac{1}{4}$

So $\vec{r} = \left(\frac{\sqrt{3}}{4}, \frac{1}{4}\right)$ and therefore $\|r\| = \sqrt{\frac{3}{16} + \frac{1}{16}} = \frac{1}{2}$



So the consistent learner L should be defined as follows:

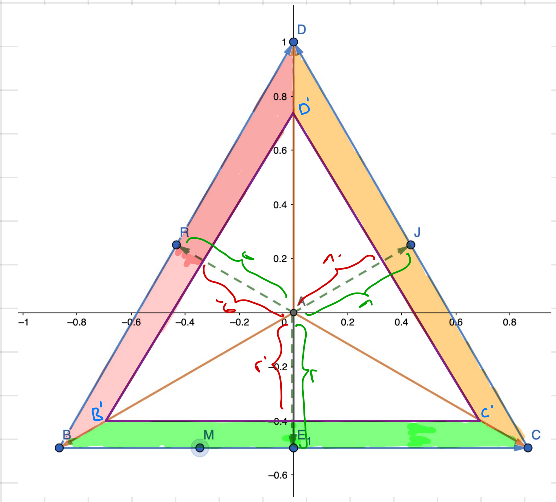
$$h = L(D) = \max \|X_{1i}, X_{2i}\| \left\{ \begin{matrix} (X_1, X_2)\cdot U \le r \\ (X_1, X_2)\cdot V \le r \\ (X_1, X_2)\cdot W \le r \end{matrix} \;\middle|\; i \in m \right\} \text{ such that } \forall\, x \in X \;\; h(x) = 1 \Rightarrow c(x) = 1$$

once we get the positive point with the maximal distance from $(0,0)$, we will draw an equilateral triangle such that this point determines the value of r^h.

therefore we get a boundary for all positive data points in the training data (the purple triangle in the example above).

<u>Time complexity</u> $\longrightarrow$ for each sample calculate 3 points, therefore $O(3m) = O(m)$

Now, we will divide the space between the concept and the hypothesis to 3 parts.



$$B_1 = B_2 = B_3 = \frac{\left(\|\vec{DC}\| + \|\vec{D'C'}\|\right)\cdot\left(\|r\| - \|r'\|\right)}{2}$$

<u>Sample complexity:</u>

Consider training data $D \in X^m$, the probability of the data D to be in either $B_1 \cup B_2 \cup B_3$ is $P_{B_1} = P_{B_2} = P_{B_3} = \frac{\varepsilon}{3}$. Assuming D visits each of the 3 B_is, we can evaluate the error $\varepsilon(h,c)$ as follows:

$$Err(L(D), concept) = Err(h, concept) = \varepsilon$$

for given ε, δ the required number of samples is:

$$P(\{D \in X^m : Err(h = L(D), concept) > \varepsilon\}) \le \delta \Rightarrow \sum_{i=1}^{3} [P(x - B_i)]^m \le 3\cdot\left[1 - \frac{\varepsilon}{3}\right]^m \le 3\cdot e^{-\frac{m\varepsilon}{3}} \le \delta \Rightarrow$$

$$\Rightarrow \ln(3) - \frac{m\varepsilon}{3} \le \ln(\delta) \Rightarrow \ln\left(\frac{3}{\delta}\right) \le \frac{m\varepsilon}{3} \Rightarrow \frac{3}{\varepsilon}\ln\left(\frac{3}{\delta}\right) \le m$$

$\Rightarrow$ Therefore the sample complexity is polynomial.

(4) With 95% confidence, the true error they can expect is up to 22.52%

(5) SVM: