Introduction to Machine Learning

Homework 2

Names: Obaida Khateeb, 201278066

Adham Dahli, 318575834

Problem 4 – Nonlinear SVM

1. Find the feature vector mapping, i.e., the φ that maps each sample to its feature space.

   Giving that

$$K(x,y) = \frac{x^t y}{||x|| * ||y||}$$

   We aim to express $K(x,y)$ as $< \phi(x), \phi(y) >$:

$$K(x,y) = \frac{x^t y}{||x|| * ||y||} = \left(\frac{x^t}{||x||}\right) * \left(\frac{y}{||y||}\right) = \left(\frac{x}{||x||}\right)^t * \left(\frac{y}{||y||}\right)$$

$$= \left(\frac{x}{||x||}\right) \cdot \left(\frac{y}{||y||}\right) = < \phi(x), \phi(y) >$$

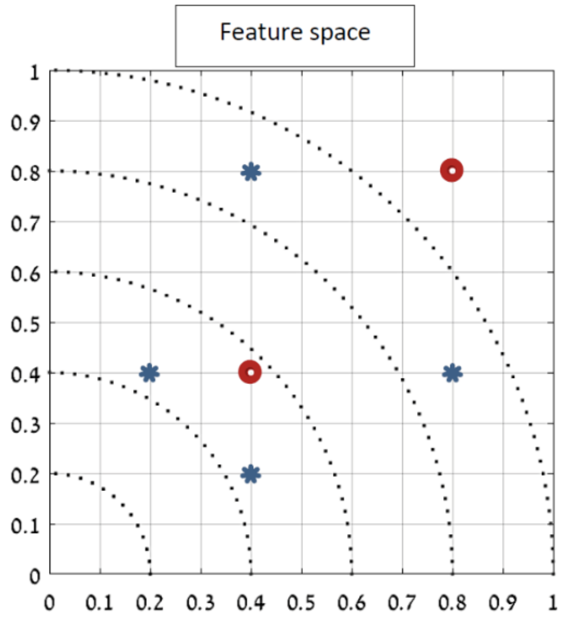
   Hence we conclude that:

$$\phi(x) = \frac{x}{||x||}, \phi(y) = \frac{y}{||y||}$$

   And particularly:

$$\boldsymbol{\phi}(\boldsymbol{x}) = \frac{\boldsymbol{x}}{||\boldsymbol{x}||}$$

2. Given the following data, map the points to their new feature representations using the figure as the feature space. Draw the 6 new points (with colors) on the figure.



The points by their coordinates are: (0.2, 0.4), (0.4, 0.2), (0.4, 0.4), (0.4, 0.8), (0.8, 0.4), (0.8, 0.8).

In order to map the points we're applying the feature vector $\phi(x)$ from the 1st section:

Firstly, lets <u>calculate the norm of each one</u> of the points:

$$||(0.2, 0.4)|| = \sqrt{0.2^2 + 0.4^2} = \sqrt{0.2} = 0.447$$

$$||(0.4, 0.2)|| = \sqrt{0.4^2 + 0.2^2} = \sqrt{0.2} = 0.447$$

$$||(0.4, 0.4)|| = \sqrt{0.4^2 + 0.4^2} = \sqrt{0.32} = 0.565$$

$$||(0.4, 0.8)|| = \sqrt{0.4^2 + 0.8^2} = \sqrt{0.8} = 0.894$$

$$||(0.8, 0.4)|| = \sqrt{0.8^2 + 0.4^2} = \sqrt{0.8} = 0.894$$

$$||(0.8, 0.8)|| = \sqrt{0.8^2 + 0.8^2} = \sqrt{1.28} = 1.131$$

Secondly, to apply $\phi(x) = \frac{x}{||x||}$ for each of the points:

$$\phi\big((0.2, 0.4)\big) = \frac{(0.2,0.4)}{0.447} = (0.447,0.894)$$

$$\phi((0.4, 0.2)) = \frac{(0.4,0.2)}{0.447} = (0.894,0.447)$$

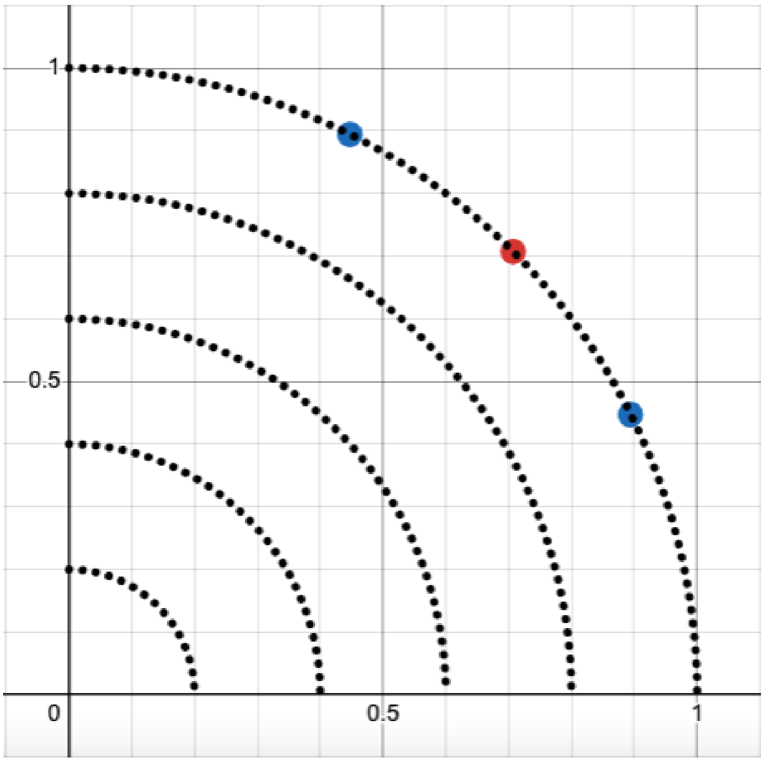
$$\phi\big((0.4, 0.4)\big) = \frac{(0.4,0.4)}{0.565} = (0.707,0.707)$$

$$\phi\big((0.4, 0.8)\big) = \frac{(0.4,0.8)}{0.894} = (0.447,0.894)$$

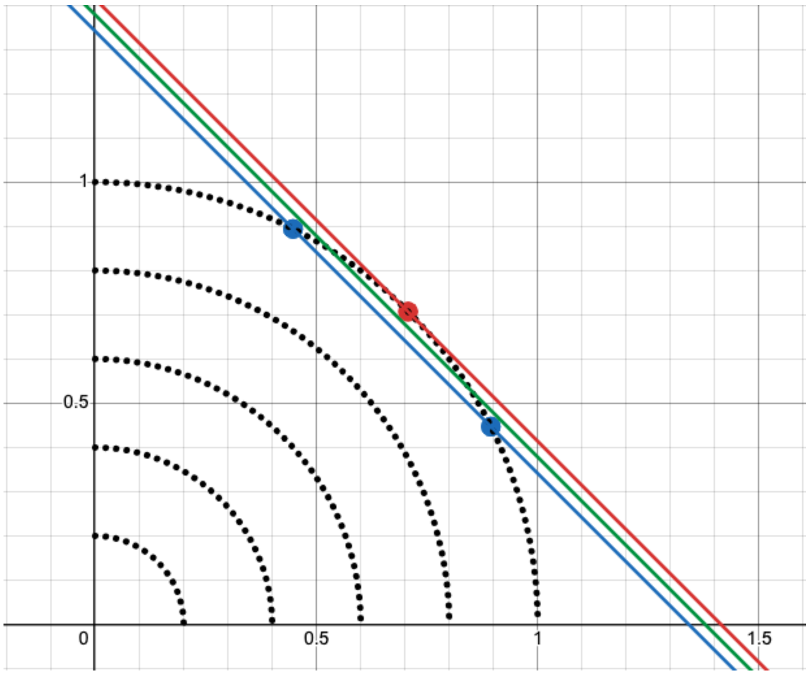
$$\phi\big((0.8, 0.4)\big) = \frac{(0.8,0.4)}{0.894} = (0.894,0.447)$$

$$\phi\big((0.8, 0.8)\big) = \frac{(0.8,0.8)}{1.131} = (0.707,0.707)$$

Lastly, to draw the points on the figure:



3. Draw the resulting margin decision boundary in the feature space. Also draw the separating line.



4. Given that the separating hyperplane is defined by ℓ:$-x-y+1.378=0$, so w=(-1,-1), b=1.378, find the alphas for each sample.

$$\alpha_1 = \frac{1000}{73} = 13.698$$
$$\alpha_2 = \frac{1000}{73} = 13.698$$
$$\alpha_3 = \frac{2000}{73} = 27.397$$

5. Draw inside the desmos link the decision boundary in the original input space, resulting from the kernel. Recall that the nonlinear hyperplane given by:

$$\sum_{i=1}^{3} y_i \alpha_i K(x, x_i) + b = 0$$

Where b is given to you, you found the alphas earlier.

$$\sum_{i=1}^{3} y_i * \alpha_i * K(x, x_i) + b = 0$$

$$= y_1 * \alpha_1 * K(x, x_1) + y_2 * \alpha_2 * K(x, x_2) + \; y_3 * \alpha_3 * K(x, x_3) + \; 1.378 = 0$$

$$= 1 * \; 13.698 * K\big((x, y), (0.447, 0.894)\big) + \; 1 * \; 13.698 * K\big((x, y), (0.894, 0.447)\big) - 1 * \; 27.397 * K\big((x, y), (0.707, 0.707)\big) + \; 1.378 = 0$$

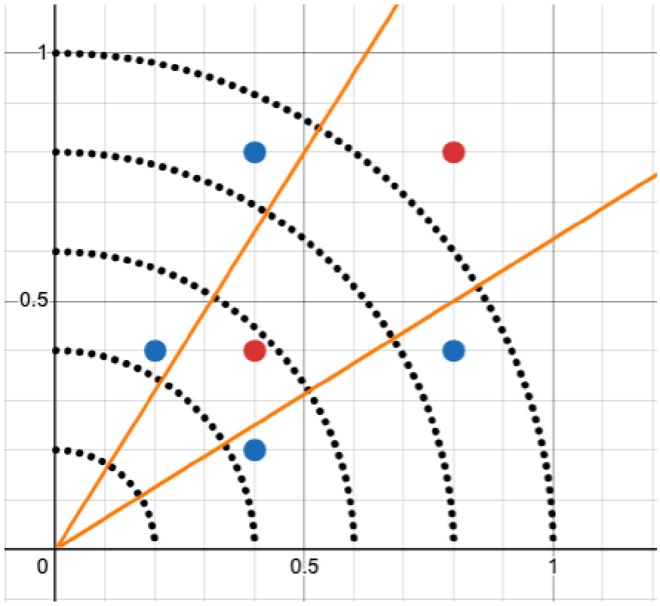
$$= 13.698 * \frac{(x, y)^t * (0.447, 0.894)}{\sqrt{x^2 + y^2} * \sqrt{0.447^2 + 0.894^2}} + 13.698 * \frac{(x, y)^t * (0.894, 0.447)}{\sqrt{x^2 + y^2} * \sqrt{0.894^2 + 0.447^2}} - 27.397 * \frac{(x, y)^t * (0.707, 0.707)}{\sqrt{x^2 + y^2} * \sqrt{0.707^2 + 0.707^2}} + 1.378 = 0$$

$$= 13.698 * \frac{0.447x + 0.894y}{\sqrt{x^2 + y^2}} + 13.698 * \frac{0.894x + 0.447y}{\sqrt{x^2 + y^2}} - 27.397 * \frac{0.707x + 0.707y}{\sqrt{x^2 + y^2}} + 1.378 = 0$$

$$= 13.698 * \frac{1.341 + 1.341y}{\sqrt{x^2 + y^2}} - 27.397 * \frac{0.707x + 0.707y}{\sqrt{x^2 + y^2}} + 1.378 = 0$$

$$= \frac{-x - y}{\sqrt{x^2 + y^2}} + 1.378 = 0$$

After placing the resulted equation in desmos, we got the following decision boundary:



6. Consider two distinct points $x_1, x_2 \in R^d$ with labels $y_1 = 1, y_2 = -1$. Compute the hyperplane that Hard SVM will return on this data, i.e., give explicit expressions for w and b as functions of x1, x2.

   The primal problem is to find:

$$\min \frac{1}{2} * ||w||^2$$

   Subject to:

$$y_i(w * x_i + b) \geq 1, \forall i$$

   On our case, $i = 1{,}2$, then:

$$1 * (w * x_1 + b) \geq 1 \rightarrow w * x_1 + b \geq 1$$

$$(-1) * (w * x_1 + b) \geq 1 \rightarrow w * x_2 + b \leq 1$$

   From the primal form the problem we can drive the dual form of it by introducing the Lagrange multipliers α_1, α_2 for the constraints:

$$L_p = L(w, b, \alpha) = \frac{1}{2}\left||w|\right|^2 - \sum_{i=1}^{2} \alpha_i (y_i(x_i * w + b) - 1)$$

$$= \frac{1}{2}\left||w|\right|^2 - \sum_{i=1}^{2} \alpha_i y_i x_i w + b\alpha_i y_i - \alpha_i$$

Then, to find the partial derivatives of L_p with respect to w and b while equaling them to 0:

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{2} \alpha_i y_i \ = -(\alpha_1 * 1) - \left(\alpha_2 * (-1)\right) = \ -\alpha_1 + \alpha_2 = 0$$

$\rightarrow \ \boldsymbol{\alpha_1} = \ \boldsymbol{\alpha_2}$

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{2} \alpha_i y_i x_i \ = w - (\alpha_1 * 1 * x_1) - (\alpha_2 * (-1) * x_2)$$

$$= \ w - \alpha_1 x_1 + \alpha_2 x_2 = 0$$

$\rightarrow w = \alpha_1 x_1 - \alpha_2 x_2$

Since $\alpha_1 = \ \alpha_2$:

$\boldsymbol{w} = \boldsymbol{\alpha x_1} - \boldsymbol{\alpha x_2} = \ \boldsymbol{\alpha}(\boldsymbol{x_1} - \boldsymbol{x_2})$

Placing w in the margin constraints will get us the following:

$\boldsymbol{\alpha}(\boldsymbol{x_1} - \boldsymbol{x_2}) * x_1 + b = 1 \rightarrow \ \alpha x_1^2 - \alpha x_1 x_2 + b = 1$

$\boldsymbol{\alpha}(\boldsymbol{x_1} - \boldsymbol{x_2}) * x_2 + b = -1 \rightarrow \ \alpha x_1 x_2 - \alpha x_2^2 + b = -1$

$\rightarrow (\alpha x_1^2 - \alpha x_1 x_2) - (\alpha x_1 x_2 - \alpha x_2^2) = 2$

$\rightarrow \alpha x_1^2 - 2\alpha x_1 x_2 + \alpha x_2^2 = 2$

$\rightarrow \alpha(x_1 - x_2)^2 = 2$

$\rightarrow \ \alpha = \dfrac{2}{(x_1 - x_2)^2}$

We're placing the α in one of the margin constrains to find the b:

$$\frac{2}{(x_1 - x_2)^2} * (x_1 - x_2) * x_1 + b = 1 \rightarrow \frac{2x_1}{x_1 - x_2} + b = 1$$

$$\boldsymbol{b = 1 - \frac{2x_1}{x_1 - x_2} = \frac{-x_1 - x_2}{x_1 - x_2}}$$

So w is:

$$\boldsymbol{w = \frac{2}{(x_1 - x_2)^2}(x_1 - x_2) = \frac{2}{x_1 - x_2}}$$