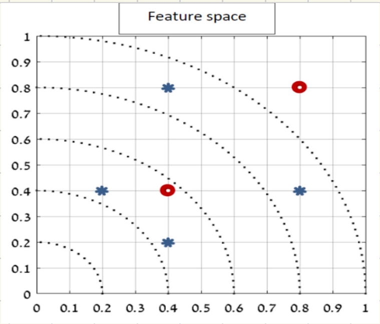
Problem 4: Nonlinear SVM

We introduce the normalized linear kernel:

$$K(X,y) = \frac{X^t y}{\|X\| \cdot \|y\|}$$



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

To find the feature vector mapping $\varphi(X)$ that maps each sample to its feature space using the Kernel $K(X,y) = \frac{X^T y}{\|X\| \cdot \|y\|}$ we normalize each vector using the function $\boxed{\varphi(X) = \frac{X}{\|X\|}}$

---

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 on the figure.

First let's calculate the new points according to the function $\varphi(X) = \frac{X}{\|X\|}$ that we found before.

Blue points mapped to Green:

point 1: $(0.2, 0.4)$ $\varphi(x) = \left(\frac{0.2}{0.447}, \frac{0.4}{0.447}\right) = (0.447, 0.894)$

point 2: $(0.4, 0.2)$ $\varphi(x) = \left(\frac{0.4}{0.447}, \frac{0.2}{0.447}\right) = (0.894, 0.447)$

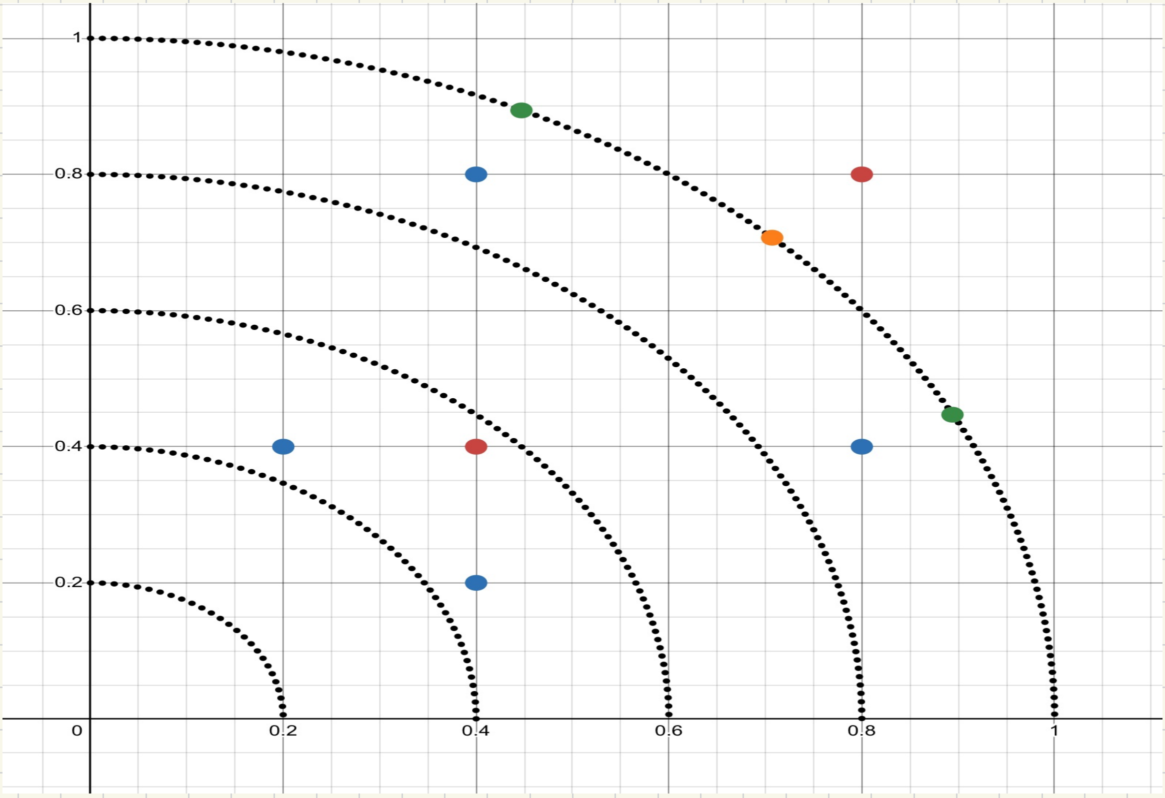
point 3: $(0.2, 0.4)$ $\varphi(x) = \left(\frac{0.4}{0.894}, \frac{0.8}{0.894}\right) = (0.447, 0.894)$

point 4: $(0.2, 0.4)$ $\varphi(x) = \left(\frac{0.8}{0.894}, \frac{0.4}{0.894}\right) = (0.894, 0.447)$

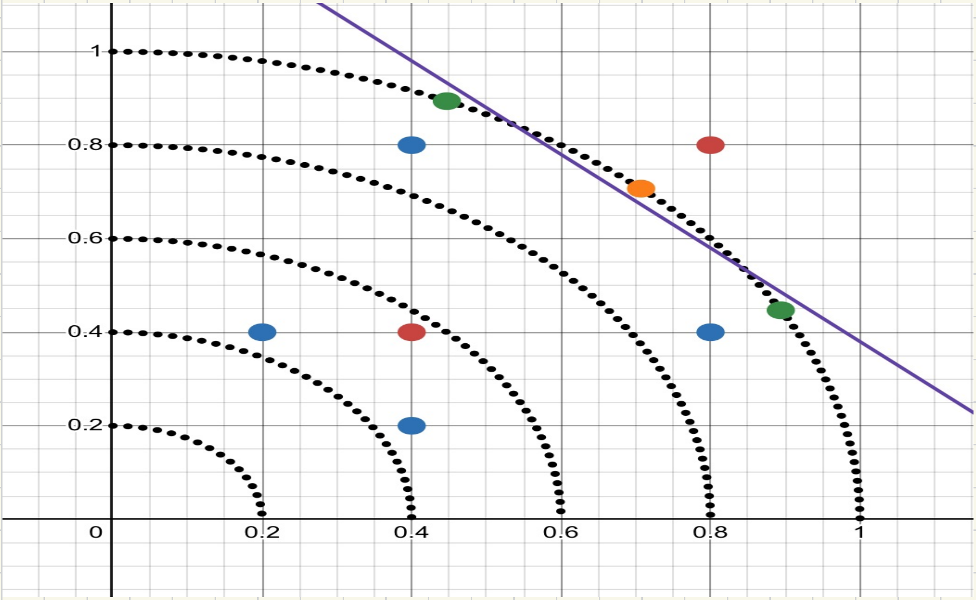
Red points mapped to Yellow:

point 1: $(0.4, 0.4)$ $\varphi(x) = \left(\frac{0.4}{0.566}, \frac{0.4}{0.566}\right) = (0.707, 0.707)$

point 2: $(0.8, 0.8)$ $\varphi(x) = \left(\frac{0.8}{1.131}, \frac{0.8}{1.131}\right) = (0.707, 0.707)$



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.

Given the points we found before and $w = \sum_{i=1}^{3} y_i \alpha_i \varphi_i$

we know that $w = (-1, -1)$ thus:

$$-1 = y_1\alpha_1 \cdot 0.447 + y_2\alpha_2 \cdot 0.894 + y_3\alpha_3 \cdot 0.707$$

$$-1 = y_1\alpha_1 \cdot 0.894 + y_2\alpha_2 \cdot 0.447 + y_3\alpha_3 \cdot 0.707$$

assuming that the points have the labels $\{1, 1, -1\}$

meaning the the equations are:

$-1 = 1 \cdot \alpha_1 \cdot 0.447 + 1 \cdot \alpha_2 \cdot 0.894 + (-1)\,\alpha_3 \cdot 0.707$

$-1 = 1 \cdot \alpha_1 \cdot 0.894 + 1 \cdot \alpha_2 \cdot 0.447 + (-1)\,\alpha_3 \cdot 0.707$

$\sum_{i=1}^{3} y_i \alpha_i = 1 \cdot \alpha_1 + 1 \cdot \alpha_2 - 1 \cdot \alpha_3 = 0$

now we need to solve the equations:

$-2 = 1.341 \cdot \alpha_1 + 1.341 \cdot \alpha_2 - 1.414 \cdot \alpha_3$. sum of the first 2 equations.

$\alpha_1 + \alpha_2 - \alpha_3 = 0 \rightarrow \alpha_3 = \alpha_1 + \alpha_2$ now substitute.

$-2 = 1.341 \cdot \alpha_1 + 1.341 \cdot \alpha_2 - 1.414 \cdot (\alpha_1 + \alpha_2)$

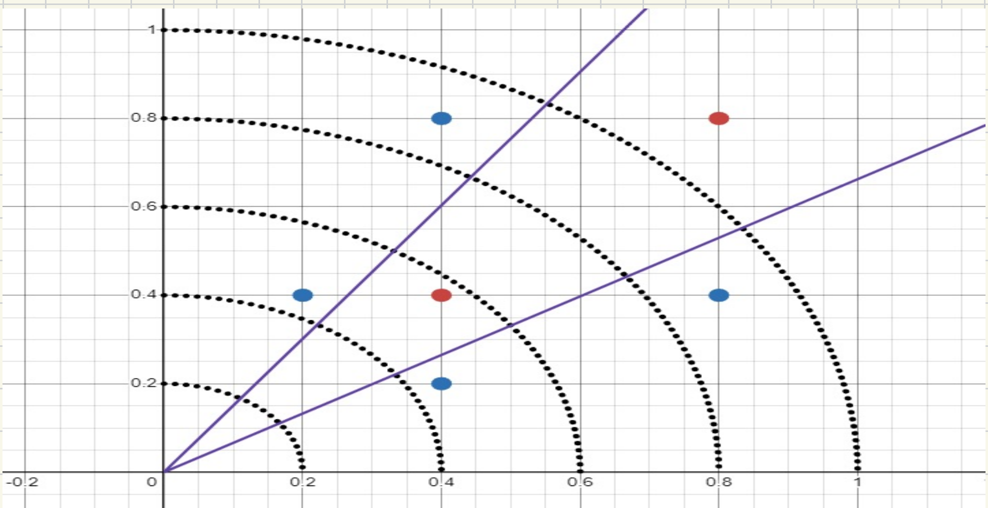
$-2 = -0.073 \cdot \alpha_1 - 0.073 \cdot \alpha_2 \Rightarrow \alpha_1 + \alpha_2 = \frac{2}{0.073} = 27.397$

since $\alpha_3 = \alpha_1 + \alpha_2$ then $\boxed{\alpha_3 = 27.397}$

and assuming that $\alpha_1 = \alpha_2$ we can infer $\boxed{\alpha_1 = \alpha_1 = 13.698}$

---

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.



6) Consider two distinct points $x_1, x_2 \in \mathbb{R}^d$ 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 x_1, x_2.

we start from the primal problem $\min \frac{1}{2}\|w\|^2$ and we subject it to the constraints $y_i(wx_i + b) \geq 1 \ \forall i$

now the dual problem as discussed in class was:

$L(w,b,a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{2} \alpha_i(y_i(wx_i + b) - 1)$ where $\alpha_i \geq 0$

we get this equation by deriving using lagrange multipliers. Now to solve it we need to substitute the $w = \sum_i \alpha_i y_i x_i$ and we need to maximize the equation:

$\max \sum_{i=1}^{2} \alpha_i - \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_i \alpha_j y_i y_j x_i x_j$ now subject to the constraint $\sum_{i=1}^{2} \alpha_i y_i = 0$ and since $y_1 = 1$, $y_2 = -1$ then we get $\alpha_1 = \alpha_2$ and then we can do a reduction to one variable $\alpha_1 = \alpha_2 = \alpha \Rightarrow \max 2\alpha - \alpha^2(x_1 \cdot x_1 + x_2 \cdot x_2 - 2x_1 \cdot x_2)$

$\Rightarrow \max 2\alpha - \alpha^2 \|x_1 - x_2\|^2$ now we solve for α, we derive:

$2 - 2\alpha\|x_1 - x_2\|^2 = 0 \Rightarrow \alpha = \frac{1}{\|x_1 - x_2\|^2}$ then we use the value of α to find w and b. $w = \alpha_1 y_1 x_1 + \alpha_2 y_2 x_2 = \frac{1}{\|x_1 - x_2\|^2}(x_1 - x_2)$ and to find b we use the support vectors: $w \cdot x_1 + b = 1$, $w \cdot x_2 + b = -1$

we add the two equations to get: $wx_1 + wx_2 + 2b = 0$

$\Rightarrow b = -\frac{1}{2}(w(x_1+x_2))$ now substitute w:

$b = -\frac{1}{2}\left(\frac{1}{\|x_1-x_2\|^2}(x_1-x_2)\cdot(x_1+x_2)\right) \Rightarrow b = -\frac{1}{2}\left(\frac{(x_1-x_2)(x_1+x_2)}{\|x_1-x_2\|^2}\right)$

$\Rightarrow b = -\frac{1}{2}\left(\frac{\|x_1\|^2-\|x_2\|^2}{\|x_1-x_2\|^2}\right)$ in conclusion we get the following expressions

$$\boxed{b = -\frac{1}{2}\left(\frac{\|x_1\|^2-\|x_2\|^2}{\|x_1-x_2\|^2}\right), \quad w = \frac{x_1-x_2}{\|x_1-x_2\|^2}}$$