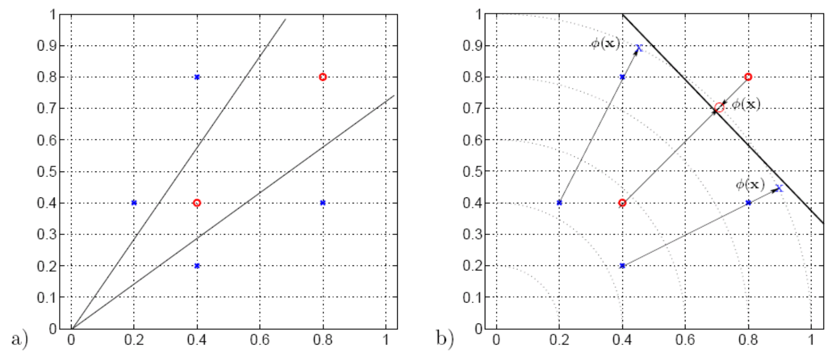
5 Kernel Method (20%)

Suppose we have six training points from two classes as in Figure (a). Note that we have four points from class 1: $(0.2, 0.4), (0.4, 0.8), (0.4, 0.2), (0.8, 0.4)$ and two points from class 2: $(0.4, 0.4), (0.8, 0.8)$. Unfortunately, the points in Figure (a) cannot be separated by a linear classifier. The kernel trick is to find a mapping of $\mathbf{x}$ to some feature vector $\phi(\mathbf{x})$ such that there is a function K called kernel which satisfies $K(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T\phi(\mathbf{x}')$. And we expect the points of $\phi(\mathbf{x})$ to be linearly separable in the feature space. Here, we consider the following normalized kernel:

$$K(\mathbf{x}, \mathbf{x}') = \frac{\mathbf{x}^T\mathbf{x}'}{||\mathbf{x}^T||||\mathbf{x}'||}$$

$$= \frac{x^T}{||x^T||} \cdot \frac{x'}{||x'||} = \phi(x) \cdot \phi(x')$$



1. (5%) What is the feature vector $\phi(\mathbf{x})$ corresponding to this kernel? Draw $\phi(\mathbf{x})$ for each training point $\mathbf{x}$ in Figure (b), and specify from which point it is mapped.

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

2. (5%) You now see that the feature vectors are linearly separable in the feature space. The maximum-margin decision boundary in the feature space will be a line in $\mathbb{R}^2$, which can be written as $w_1x + w_2y + c = 0$. What are the values of the coefficients w_1 and w_2? (Hint: you don't need to compute them.)

   (sol.)

$$(w_1, w_2) = (1, 1)$$

3. (3%) Circle the points corresponding to support vectors in Figure (b).

4. (7%) Draw the decision boundary in the original input space resulting from the normalized linear kernel in Figure (a). Briefly explain your answer.

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

1 $\phi((0.2, 0.4))$

$$= \frac{(0.2, 0.4)}{\sqrt{0.2^2 + 0.4^2}} \rightarrow x$$

class 1

$$= \frac{(0.2, 0.4)}{\sqrt{0.20}} = \frac{(0.2, 0.4)}{\sqrt{4 \times .05}}$$

$$= \frac{(0.2, 0.4)}{2\sqrt{0.05}} = \left(\frac{0.1}{\sqrt{0.05}}, \frac{0.2}{\sqrt{0.05}}\right)$$

$$= (0.45, 0.90)$$

2 $\phi((0.4, 0.8)) = \left(\frac{0.1}{\sqrt{0.05}}, \frac{0.2}{\sqrt{0.05}}\right)$

for 3 $(0.4, 0.2)$ & $(0.8, 0.4)$ 4 $\Longrightarrow (0.90, 0.45)$

$$\phi((0.4, 0.4)) = \frac{0.4}{4\sqrt{0.02}}, \frac{0.4}{4\sqrt{0.02}}$$

class 2

$$= \left(\frac{0.1}{\sqrt{0.02}}, \frac{0.1}{\sqrt{0.02}}\right) = (0.7, 0.7)$$

$$= \phi((0.8, 0.8))$$

So from 6 points, we have 3 distinct points

class 1 → (0.45, 0.90)
→ (0.90, 0.45)

class 2 → (0.7, 0.7)

& separable.