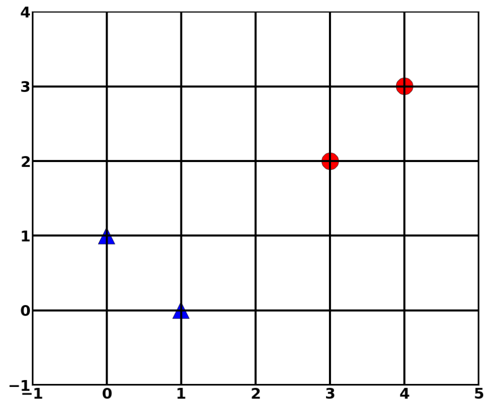
CSE 250B: Section 7 - Sharad Vikram

1. SVM Example

   You're presented with the following set of data (triangle = +1, circle = -1):



   Find the equation (by hand) of the hyperplane $w^T x + b = 0$ that the SVM algorithm would learn. Which points are support vectors?

   Solution: The equation of the hyperplane will pass through point $(2, 1)$, with a slope of -1, since it's the halfway point between the two closest points between the classes. The equation of this line is $x_1 + x_2 = 3$. We know that from this form, $w_1 = w_2$. We also know that the at the support vectors, $w^T x + b = \pm 1$. This gives us the equations:

$$1w_1 + 0w_2 + b = 1$$

$$3w_1 + 2w_2 + b = -1$$

   Solving this system of equations, we get $\vec{w} = [-\frac{1}{2}, -\frac{1}{2}]^T$ and $b = \frac{3}{2}$.

   The support vectors are $(1, 0), (0, 1)$, and $(3, 2)$.

2. RBF Kernel

   Show that the simple RBF kernel for 1-dimensional data

$$k(x_1, x_2) = \exp\left\{-(x_1 - x_2)^2\right\}$$

   is equivalent to an infinite dimensional feature mapping. Hint: the Taylor expansion of $e^{f(x)}$ is $\sum_{i=0}^{\infty} \frac{f(x)^i}{i!}$.

Solution:

$$
\begin{aligned}
k(x_1, x_2) &= \exp\left\{-(x_1 - x_2)^2\right\} \\
&= \exp\left\{-x_1^2\right\} \exp\left\{-x_2^2\right\} \exp\left\{2x_1x_2\right\} \\
&= \exp\left\{-x_1^2\right\} \exp\left\{-x_2^2\right\} \sum_{i=0}^{\infty} \frac{(2x_1x_2)^i}{i!} \\
&= \exp\left\{-x_1^2\right\} \exp\left\{-x_2^2\right\} \sum_{i=0}^{\infty} \frac{2^i x_1^i x_2^i}{i!} \\
&= \exp\left\{-x_1^2\right\} \exp\left\{-x_2^2\right\} \sum_{i=0}^{\infty} \left(\left(\sqrt{\frac{2^i}{i!}} x_1^i\right)\left(\sqrt{\frac{2^i}{i!}} x_2^i\right)\right) \\
&= \sum_{i=0}^{\infty} \left(\left(\exp\left\{-x_1^2\right\} \sqrt{\frac{2^i}{i!}} x_1^i\right)\left(\exp\left\{-x_2^2\right\} \sqrt{\frac{2^i}{i!}} x_2^i\right)\right) \\
&= \phi(x_1)^T \phi(x_2)
\end{aligned}
$$

where

$$
\begin{aligned}
\phi(x) &= [\phi_1(x), \phi_2(x), \ldots] \\
\phi_i(x) &= \exp\left\{-x^2\right\} \sqrt{\frac{2^i}{i!}} x^i
\end{aligned}
$$

3. Kernel Math

To show that a kernel is valid, you need to show its corresponding feature mapping exists i.e. $k(x_1, x_2) = \phi(x_1)^T \phi(x_2)$. Let $k_1(x_1, x_2)$ and $k_2(x_1, x_2)$ be valid kernel functions. Show that the following kernel functions are valid:

(a) $k'(x_1, x_2) = ck_1(x_1, x_2)$ where $c > 0$.

Solution:

$$
\begin{aligned}
k'(x_1, x_2) &= c\phi(x_1)^T \phi(x_2) \\
&= \phi'(x_1)^T \phi'(x_2)
\end{aligned}
$$

where

$$
\phi'(x) = \sqrt{c}\phi(x)
$$

(b) $k'(x_1, x_2) = f(x_1)k_1(x_1, x_2)f(x_2)$ where f is any function.

Solution:

$$
\phi'(x) = f(x)\phi(x)
$$

(c) $k'(x_1, x_2) = k_1(x_1, x_2)k_2(x_1, x_2)$

Solution:

$$\begin{aligned}
k'(x_1, x_2) &= \left(\phi_1(x_1)^T \phi_1(x_2)\right) \left(\phi_2(x_1)^T \phi_2(x_2)\right) \\
&= \left(\sum_i \phi_1(x_1)_i \phi_1(x_2)_i\right) \left(\sum_j \phi_2(x_1)_j \phi_2(x_2)_j\right) \\
&= \sum_{ij} \left(\phi_1(x_1)_i \phi_2(x_1)_j\right) \left(\phi_1(x_2)_i \phi_2(x_2)_j\right) \\
&= \sum_{ij} \phi'(x_1)_{ij} \phi'(x_2)_{ij}
\end{aligned}$$

where

$$\phi'(x)_{ij} = \phi_1(x)_i \phi_2(x)_j$$