Homework 9

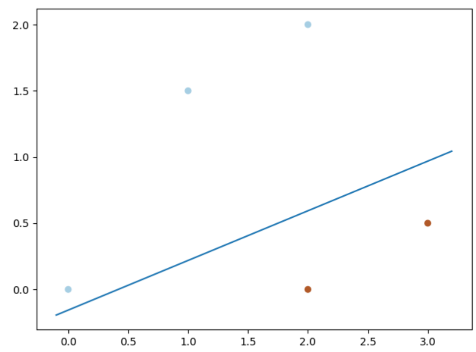
Name: Yuan Li
ID: N19728558 netID yl6606
Email: foxerlee1@gmail.com

Part I: Written Exercises

1. (a)

Answer:

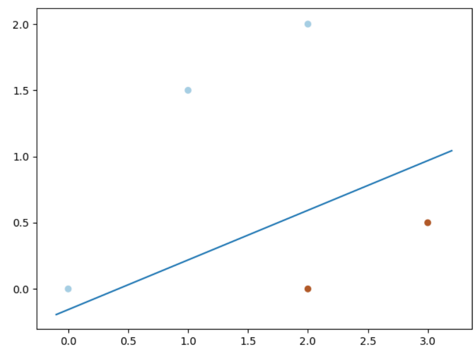
The training data and hyperplane should be as follow:



1. (b)

Answer:

The training data and hyperplane should be as follow, which will be the same as 1. (a):



1. (c)

Answer:

They are linearly separable.

1. (d)

Answer:

For $w^T = [6 \ -16]$ and $w_0 = -2.5$:

The functional margin for each training example is $[2.5, 22.5, 9.5, 20.5, 7.5]$. And the functional margin is:

$$
\begin{aligned}
\gamma &= \min_{i=1..n} \gamma^{(i)} \\
&= \min_{i=1..n} y^{(i)}(w^T x^{(i)} + w_0) \\
&= 2.5
\end{aligned}
$$

The geometric margin for each training example is $[0.14630143, 1.31671291, 0.55594545, 1.19967176, 0.4389043]$. And the geometric margin is:

$$
\begin{aligned}
\gamma &= \min_{i=1..n} \gamma^{(i)} \\
&= \min_{i=1..n} y^{(i)}(w^T x^{(i)} + w_0)/||w||_2 \\
&= 0.14630143
\end{aligned}
$$

Based on functional margin, the canonical weights should be:

$$
\begin{aligned}
w^{(')T} &= \frac{1}{2.5} w^T = [2.4 \quad -6.4] \\
w_0^{(')} &= \frac{1}{2.5} w_0 = -1
\end{aligned}
$$

For $100 * w^T = [100 * 12 \quad -100 * 32]$ and $w_0 = 100 * (-5)$:

The functional margin for each training example is $[500, 4500, 1900, 4100, 1500]$. And the functional margin is:

$$
\begin{aligned}
\gamma &= \min_{i=1..n} \gamma^{(i)} \\
&= \min_{i=1..n} y^{(i)}(w^T x^{(i)} + w_0) \\
&= 500
\end{aligned}
$$

The geometric margin for each training example is $[0.14630143, 1.31671291, 0.55594545, 1.19967176, 0.4389043]$. And the geometric margin is:

$$
\begin{aligned}
\gamma &= \min_{i=1..n} \gamma^{(i)} \\
&= \min_{i=1..n} y^{(i)}(w^T x^{(i)} + w_0)/||w||_2 \\
&= 0.14630143
\end{aligned}
$$

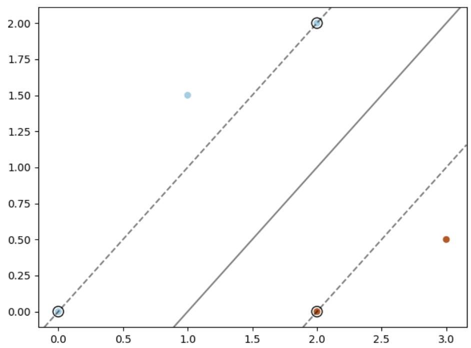
Based on functional margin, the canonical weights should be:

$$
\begin{aligned}
w^{(')T} &= \frac{1}{500} w^T = [2.4 \quad -6.4] \\
w_0^{(')} &= \frac{1}{500} w_0 = -1
\end{aligned}
$$

1. (e)

Answer:

After I use a SVM to train my examples, I can get result as follow:



Points with a circle are support vectors, which are: $(0,0)^T, (2,0)^T, (2,2)^T$

1. (f)

Answer:

If we add point $x = (1,3)^T$, The margin and hyperplane don't change. They are:

$$margin : y = x \text{ } and \text{ } y = x - 2$$
$$hyperplane : y = x - 1$$

1. (g)

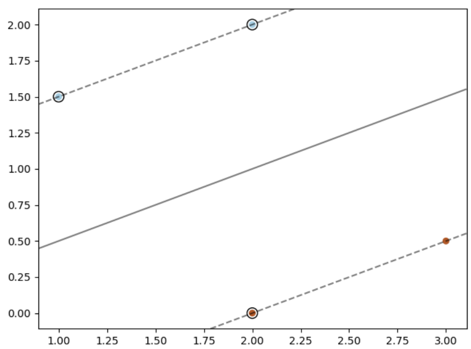
Answer:

If we remove point $x = (1, 1.5)^T$, The margin and hyperplane don't change.

1. (h)

Answer:

If we remove point $x = (0,0)^T$, The margin and hyperplane will change. They become:



1. (i)

Answer:

We can set as:

$$max \text{ } \gamma$$
$$subject \text{ } to \text{ } y^{(i)}(w_0 + W^T x^{(i)}) >= \gamma, \text{ } i = 1, .., N, \text{ } ||W||_2 = 1$$

2.

Answer:

The new features are:

$$\phi(x^{(1)}) = [1.0, 0.6065, 0.1353, 0.0111]$$
$$\phi(x^{(2)}) = [0.6065, 1.0, 0.6065, 0.1353]$$
$$\phi(x^{(3)}) = [0.1353, 0.6065, 1.0, 0.6065]$$
$$\phi(x^{(4)}) = [0.0111, 0.1353, 0.6065, 1.0]$$

The coefficients are: $w = [-0.656, 2.108, 2.108, -0.656], w0 = -1.900$

The predictions are: $y_{1.2} = -1$, $y_{2.5} = 1$, $y_{3.25} = 1$

3.

Answer:

They create the same decision boundary.

Because $||w||_2 = 1$, $max\ \gamma$ equals $max\ \frac{\gamma}{||w||_2}$. We also know that $max\ \frac{\gamma}{||w||_2}$ equals $max\ \frac{\gamma/\gamma}{||w/\gamma||_2}$. Then it becomes $max\ \frac{1}{||w||_2}$, which equals $min\ ||w||_2$