CM146, Fall 2018
Problem Set 02: Jonathan Chu

November 10, 2018

1 (on CCLE)

2 (on CCLE)

3 Understanding Linear Separability

(a) **Solution:** If $\delta = 0$, we have:

$$y_i(\boldsymbol{w^T x_i} + \theta) \geq 1$$

Trivially, this inequality holds when $sgn(y_i) = sgn(\boldsymbol{w^T x_i} + \theta)$ and $|\boldsymbol{w^T x_i} + \theta| \geq 1$.

The matching signs of y_i and $\boldsymbol{w^T x_i}$ indicate these values of $\boldsymbol{w}$ and θ satisfy equation (1). Therefore D is linearly separable.

(b) **Solution:** For $0 < \delta < 1$, it still holds that $sgn(y_i) = sgn(\boldsymbol{w^T x_i} + \theta)$, meaning that D remains linearly separable, and $|\boldsymbol{w^T x_i} + \theta| < 1$ for some i, which is not of any concern.

However, if the minimum $\delta \geq 1$, we have

$$y_i(\boldsymbol{w^T x_i} + \theta) \geq c, \; c \leq 0$$

and D is not linearly separable.

(c) **Solution:** The minimum value we can achieve for δ is 0. If this is the case, we seek $\boldsymbol{w}$ and $\boldsymbol{x}$ that satisfy the following:

$$y_i(\boldsymbol{w^T x_i} + \theta) \geq 0$$

Trivially, $\boldsymbol{w} = \boldsymbol{0}$, $\theta = 0$ satisfy the above inequality yet clearly will not separate D.

(d) **Solution:** Trivially, we see that $\boldsymbol{w} = [1, 1, 1]$, $\theta = 0$, $\delta = 0$ is an optimal solution.

In fact, any solution with

$$\delta = 0$$

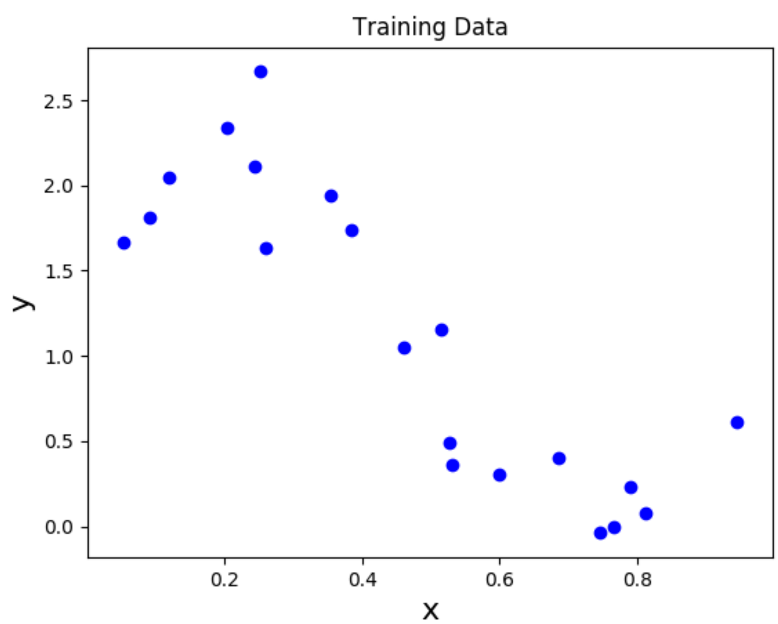
$$|\boldsymbol{w^T x_i} + \theta| = |w_1 x_1 + w_2 x_2 + w_3 x_3 + \theta| \geq 1$$

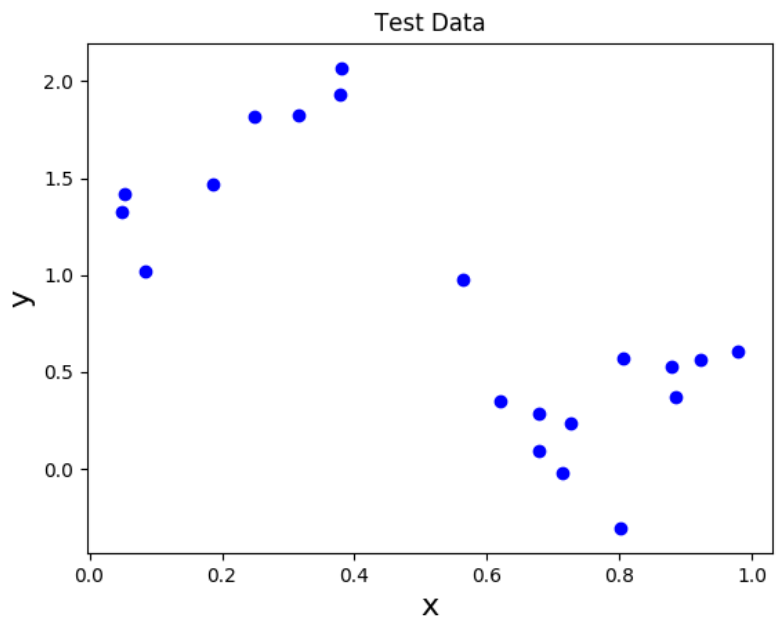
$$\implies |w_1 + w_2 + w_3 + \theta| \geq 1 \text{ and } |-w_1 + -w_2 + -w_3 + \theta| \geq 1$$

is optimal.

4 Implementation: Polynomial Regression

(a) By inspection, linear regression would not model the data very accurately. Although lines with negative slope on both the training and test data seem like they would perform moderately well, there would still be high training and test error, since the square distance from points to the line would be somewhat large for many of the data points.





(b)
```
Phi = np.concatenate((np.ones(np.shape(X)), X), 1)
```

(c)
```
y = np.dot(self.coef_, np.transpose(X))
```

(d) Investigating linear regression...

–The model cost with zero weights is 40.233847

η	# Iterations	Final Coefficients	Final Value of $J(\theta)$
0.00407	10000	-9.40470×10^{18}, -4.65229×10^{18}	1.35545×10^{38}
0.001	764		0.195629
0.0001	7020		0.195629
0.00001	10000		0.204320