Lecture 3: Linear classifiers and perceptron

2. Review of last example

- Gaussian mixture models in $\mathbb{R}^d$ and $\{0,1\}$
- Training set $S_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$
- Classifier $h : \mathbb{R}^d \to \{0,1\}$ with $h(x) = \begin{cases} 1 & \text{if } \eta(x) \geq 1/2 \\ 0 & \text{otherwise} \end{cases}$
- Training error $\hat{L}(h)$
- Test error $L(h)$
- Let η_n denotes η at n

Review: training set

The Intent

- The set of linear classifiers $\mathcal{H}$
- Linear separable
- Perceptron algorithm $\hat{h} = \mathcal{A}(S_n, h_0)$

1

Concept Review Problem: Linear Classifier

Consider the following training data in 2-dimensional space:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Concept Review Problem: Linear Classifier

Consider the following training data in 2-dimensional space:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Concept Review Problem: Classifier and Training Error 1

Consider the following training data:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 5 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

0.25

Concept Review Problem: Classifier and Training Error 2

Consider the following training data:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 6 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

0.25

Concept Review Problem: Classifier and Training Error 3

Consider the following training data:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 7 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

0.25

3. Linear classifier mathematically defined

Linear classifier hard cases

$$\{h(x, y) = \eta(x) \neq \eta(x)\}$$

Decision boundary

$$\{(x, y) = (0, 0)\}$$

1. Linear classifier

$$\{h(x, y) = \eta(x) \neq \eta(x)\}$$

Decision boundary

$$\{(x, y) = (0, 0)\}$$

Linear classifier

Linear classifier

Consider the following training data in 2-dimensional space:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

- 0.25
- 0.5
- 0.75
- 1.0

Correct

Linear classifier

Linear separation: ex

Perceptron:

Training records $S_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$ are linearly separable, then there exists a perfect vector $\hat{\theta}$ such that perceptron $\hat{\theta}_n$ will find $\{(\hat{\theta} \cdot X_i) = Y_i\}$ for all $i \in 1, \dots, n$

Quiz 1

Consider the following training data in 2-dimensional space:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

- 0.25
- 0.5
- 0.75
- 1.0

Correct

Quiz 2

Consider the following training data:

x_1	x_2	y
1	1	1
2	1	1
1	2	1
2	2	1
3	2	1
4	2	1
1	3	1
2	3	1
3	3	1
4	3	1
1	4	1
2	4	1
3	4	1
4	4	1
1	1	0
2	1	0
3	1	0
4	1	0
1	2	0
2	2	0
3	2	0
4	2	0
1	3	0
2	3	0
3	3	0
4	3	0
1	4	0
2	4	0
3	4	0
4	4	0

Consider the following classifier:

$$h(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 + x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the training error?

- 0.25
- 0.5
- 0.75
- 1.0

Correct

Linear Separation 1

Consider the following training data:

Linear Separation 2

Consider the following training data:

Linear Separation 3

Consider the following training data:

2) The perceptron algorithm

- Training error for a linear classifier (through origin)

$$\hat{L}_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(\eta(X) \neq Y)$$

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(\eta(X) \neq Y)$$

- not through origin

$$\hat{L}_n(\theta, b) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(\eta(X^T \theta + b) \neq Y)$$

Question: Perceptron $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$

$\theta = 0$ (zero)

for $i = 1, \dots, n$ do

for $j = 1, \dots, n$ do

if $Y_j (\theta \cdot X_j) \leq 0$ then

$\theta = \theta + Y_j X_j$

return θ

Example

Perceptron with bias

Same as except

$\theta = (\theta_1, \dots, \theta_d)$

$\theta_0 = b$

return θ, θ_0

Perceptron Concept Questions 1

Consider the Perceptron Algorithm without bias. Which of the following are true?

Statement 1: If $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$, then $\theta^{(k+1)} = \theta^{(k)} + x^{(i)} y^{(i)}$

Statement 2: $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$ implies $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

- Statement 1 only
- Statement 2 only
- Both Statement 1 and Statement 2
- Neither Statement 1 nor Statement 2

Correct

Consider the Perceptron Algorithm with bias. Which of the following are true?

Statement 1: $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) + b^{(k)} \leq 0$, then $\theta^{(k+1)} = \theta^{(k)} + x^{(i)} y^{(i)}$, $b^{(k+1)} = b^{(k)} + y^{(i)}$

Statement 2: $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) + b^{(k)} \leq 0$ implies $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) + b^{(k+1)} \leq 0$

- Statement 1 only
- Statement 2 only
- Both Statement 1 and Statement 2
- Neither Statement 1 nor Statement 2

Correct

Perceptron Concept Questions 2

Consider the Perceptron Algorithm without bias.

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct

Perceptron Concept Questions 3

Consider the Perceptron Algorithm without bias.

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct

Perceptron Concept Questions 4

Consider the Perceptron Algorithm without bias.

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct

Suppose $\theta^{(k)} = (1, 1)$, $x^{(i)} = (1, 1)$, $y^{(i)} = 1$. Which of the following are true?

- $\theta^{(k)} \cdot (x^{(i)} y^{(i)}) \leq 0$
- $\theta^{(k+1)} = (2, 2)$
- $\theta^{(k+1)} \cdot (x^{(i)} y^{(i)}) \leq 0$

Correct