Problem 1.3

Solution

a

$$y(t)\mathbf{w}^T(t)\mathbf{x}(t) < 0$$
$$y(t) \in \{-1, 1\}$$
$$y(t) \neq sign(\mathbf{w}^T(t)\mathbf{x}(t))$$

case 1

$$\mathbf{w}^T(t)\mathbf{x}(t) < 0 \rightarrow$$
$$sign(\mathbf{w}^T(t)\mathbf{x}(t)) = -1 \rightarrow$$
$$y(t) = 1$$

$$1 * \mathbf{w}^T(t)\mathbf{x}(t) < 0 \rightarrow$$
$$y(t) * \mathbf{w}^T(t)\mathbf{x}(t) < 0$$

case 2

$$\mathbf{w}^T(t)\mathbf{x}(t) > 0 \rightarrow$$
$$sign(\mathbf{w}^T(t)\mathbf{x}(t)) = +1 \rightarrow$$
$$y(t) = -1$$

$$-1 * \mathbf{w}^T(t)\mathbf{x}(t) < 0 \rightarrow$$
$$y(t) * \mathbf{w}^T(t)\mathbf{x}(t) < 0$$

b

$$\mathbf{w}(t+1) = \mathbf{w}(t) + y(t)\mathbf{x}(t)$$

,

$$\mathbf{x}(t) = [1, ...] \rightarrow\rightarrow \mathbf{x}(t) \neq \mathbf{0} \rightarrow \mathbf{x}^T(t)\mathbf{x}(t) > 0$$

,

$$\mathbf{y}(t) * \mathbf{y}(t) = 1$$

,

$$y(t)\mathbf{w}^T(t+1)\mathbf{x}(t) > y(t)\mathbf{w}^T(t)\mathbf{x}(t) \rightarrow$$

$$y(t)[\mathbf{w}(t) + y(t)\mathbf{x}(t)]^T\mathbf{x}(t) > y(t)\mathbf{w}^T(t)\mathbf{x}(t) \rightarrow$$

$$y(t)\mathbf{w}^T(t)\mathbf{x}(t) + y^2(t)\mathbf{x}^T(t)\mathbf{x}(t) > y(t)\mathbf{w}^T(t)\mathbf{x}(t) \rightarrow$$

$$\underbrace{y^2(t)}_{=1}\underbrace{\mathbf{x}^T(t)\mathbf{x}(t)}_{>0} > 0$$

c Assume that $\mathbf{x}(t)$ is misclassified by $\mathbf{w}(t) \rightarrow$

$$y(t)\mathbf{w}^T(t)x(t) < 0$$

$$y(t)\mathbf{w}^T(t+1)\mathbf{x}(t) > y(t)\mathbf{w}^T(t)\mathbf{x}(t)$$

Therefore $y(t)\mathbf{w}^T(t+1)\mathbf{x}(t)$ is closer to being positive and is therefore $\mathbf{w}(t)$ is closer to classifying $\mathbf{x}(t)$ correctly.

Problem 1.5

Solution

a Learning

b Design

c Learning

d Design

e Learning

Problem 1.6

Solution

a Supervised Learning

Input: Book with info i.e. Genre, length, year, topic, ...

Output: User liked the book $+1$, disliked -1

b Reinforcement Learning:

Input: Board State after our turn

Output: We eventually win from this state, We eventually lose from this state, We eventually draw from this state

Corresponding Scores: 1, -1, 0

c Unsupervised Learning - group the movies but do not actually label their type

Input: Movie video and audio

Supervised Learning:

Input: Movie video and audio

Output: Its known type

d Supervised Learning: Try to map a sequence of notes already played with the next note

Input: prefix of the notes played by a musician

Output: The actual note played by the musician

e Reinforcement Learning:

Input: customer details like debt, income, whether they take/have taken MLFD, etc. and how much this customer was loaned

Output: How much we gained or lost from loaning this much to the customer.

Score: Tangent function applied to ratio of customer's $p\&l$ to the total holdings of the bank at the beginning of the year, where net-loss would be a negative number.

Problem 1.7

Solution

a $g(x) =$ ' • '
3: 1
2: 3
1: 3
0: 1

b $g(x) =$ ' • '
3: 1
2: 3
1: 3
0: 1

c $g(101) =$ ' ∘ '
$g(110) =$ ' ∘ '
$g(111) =$ ' • '
3: 1
2: 3
1: 3
0: 1

d $g(101) =$ ' • '
$g(110) =$ ' • '
$g(111) =$ ' ∘ '
3: 1
2: 3
1: 3
0: 1

Problem 1.1

Solution

$$P(B_1|B_2) = \frac{P(B_1 \wedge B_2)}{P(B_1)} = \frac{1/2}{3/4} = \frac{2}{3}$$

Problem 1.2

Solution

a

$$h(x) = \text{sign}(\mathbf{w}^T\mathbf{x})$$

$$h(x) = +1 \rightarrow w_0 + w_1x_1 + w_2x_2 > 0 \rightarrow$$

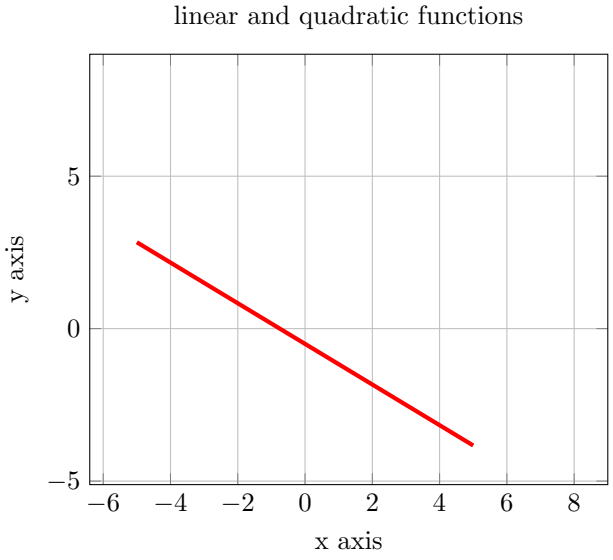
$$\left(x_2 > -w_0/w_2 - (w_1/w_2)x_1\right)$$

$$h(x) = -1 \rightarrow w_0 + w_1x_1 + w_2x_2 < 0 \rightarrow$$

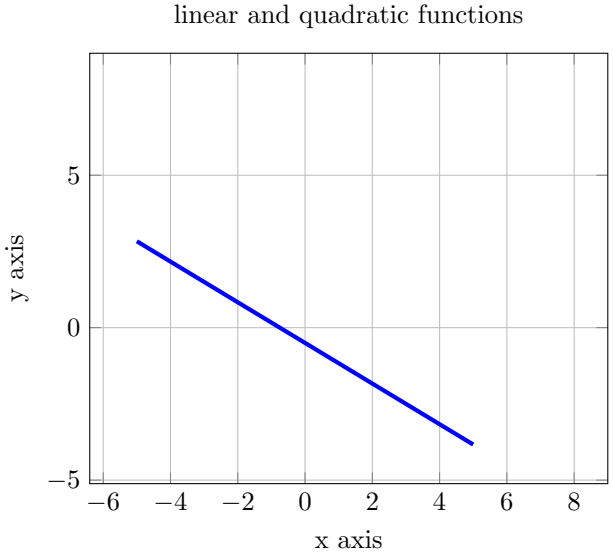
$$\left(x_2 < -w_0/w_2 - (w_1/w_2)x_1\right)$$

values picked for x_2 below $-w_o/w_2 - (w_1/w_2)x_1$ cause $h(x) = -1$ and above cause $h(x) = 1$ i.e. they are separated by a line.

$a = -w_1/w_2, b = -w_0/w_2$



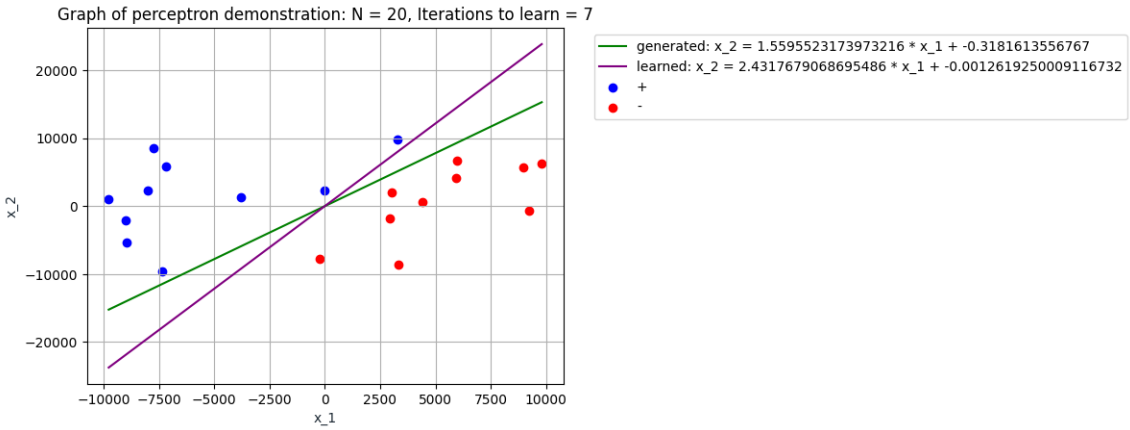
b



Problem 1.4 (a - e)

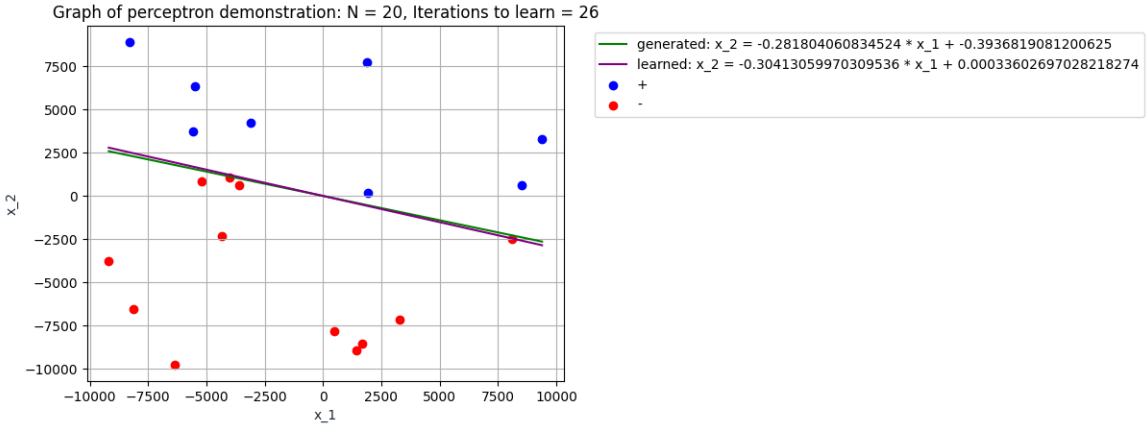
Solution

a, b



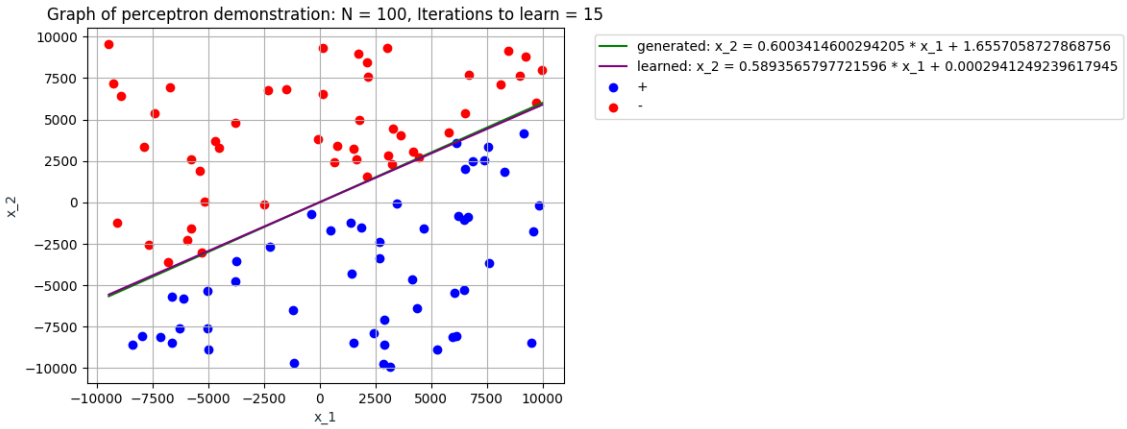
f is not very close to g. There are multiple possible g that output correct results for the data set.

c



f is much closer to g because the space of linear equations that correctly define the output of the data set is much smaller in this occurrence. It has also taken more iterations to finish the learning process in this test, compared to part b.

d f is much closer to g here than in part b. It has also taken more iterations to complete the learning process here than in part b.



e f is much closer to g here than in part b. It has also taken significantly more iterations to complete the learning process here than in part b.