Solutions on attached pages.

In-class Quiz 1 (September 10, 2024)

YOUR NAME HERE: ____________________

1 Linear classifiers (25 + 5 bonus points)

Consider a 2×1 feature vector $\mathbf{x} = [0.3, 0.4]$ that has been labeled with $y = +1$. Answer each of the questions below and **show your work**.

1. Calculate the L2 norm of $\mathbf{x}$ (its length).

2. Calculate the dot product $\mathbf{w}^T\mathbf{x}$, where $\mathbf{w} = [-1, 1]$ is a 2×1 weight vector.

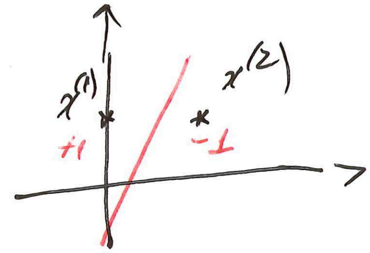
3. Find a weight vector $\mathbf{w} = [w_1, w_2]$ such that $\mathbf{w}^T\mathbf{x} - 2 \geq 0$.

4. Consider a linear model for binary classification with feature-wise parameters $\mathbf{w} = [w_1, w_2] = [-1, 1]$ and bias parameter $w_0 = -0.5$.

   (a) What is the label $\hat{y}$ predicted by this model for the example $\mathbf{x}$?

   (b) What is the equation for the decision boundary of this linear classifier?

   (c) [Bonus] What is the distance between example $\mathbf{x}$ and the decision boundary?



2 The Perceptron algorithm (25 points)

Consider a dataset with two training examples (bias feature not shown):

$\mathbf{x}^{(1)} = [0, 1]$ with label $y_1 = +1$.

$\mathbf{x}^{(2)} = [1, 1]$ with label $y_1 = -1$.

1. Will the Perceptron algorithm converge on this dataset? Explain your answer.

Yes, b.c. they are lin. sep.

Consider the Perceptron algorithm discussed in class, where the parameters are initialized as $\mathbf{w} = [w_0, w_1, w_2] = [-0.5, 0, 0]$. Run one epoch of the algorithm on the two training examples $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ above, in this order, and compute the following:

2. The weight vector after processing example $\mathbf{x}^{(1)}$. Show your work.

$W = [-0.5, 0, 0]$ $\quad W^T \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = -0.5 + 0 \times 0 + 0 \times 1 = -0.5 \Rightarrow \hat{y}_1 = -1$

$W \leftarrow W + [1, 0, 1] \Rightarrow W = [0.5, 0, 1]$

3. The weight vector after processing example $\mathbf{x}^{(2)}$. Show your work.

$W^T \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \cancel{0.5 \cdot 1 + 0 \cdot 1 + 1}$ $\quad 0.5 \times 1 + 0 \times 1 + 1 \times 1 = 1.5$

$\Rightarrow \hat{y}_2 = +1$

$W \leftarrow W + [1, 1, 1] \times -1 = [-0.5, -1, 0]$

Assume examples $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are stored in the rows of a 2D NumPy array named `X`, and the model parameters are stored in a 1D array named `w`. Write Numpy code that:

X.shape[0]

4. Adds a bias feature to each example in `X` and stores the results in an array `Xnew`.

$X = \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix}$ $\quad X_{New} = \begin{bmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \end{bmatrix}$ $\quad$ np.hstack((np.ones(2), X)).

column_stack.

5. Computes in `pred` the label predicted by `w` on the second example from `Xnew`.

pred = np.sign(w.dot(Xnew[1]) > 0)

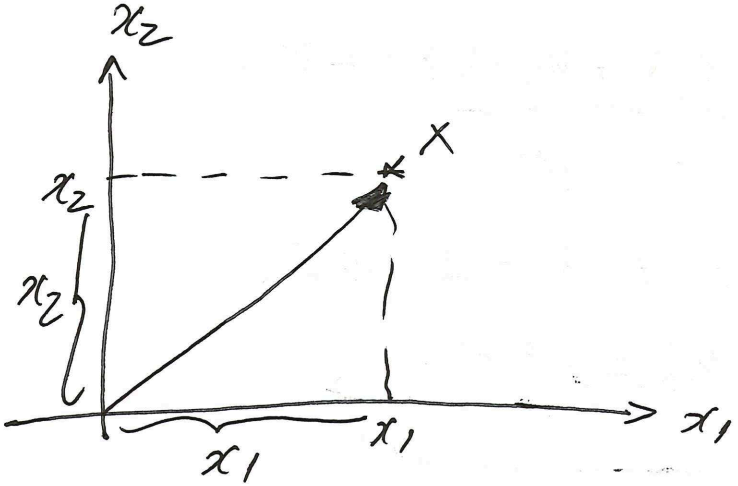
= +1 if ... else -1.

① $X = [x_1, x_2, \dots, x_N]$

L_2 norm ≡ Euclidian norm ≡ length

$$\sqrt{x_1^2+x_2^2+\dots+x_N^2} = \sqrt{\sum_{n=1}^{N} x_n^2} = \left(\sum_{n=1}^{N} x_n^2\right)^{1/2}$$

$$L_k \text{ norm} = \left(\sum_{n=1}^{N} |x_n|^k\right)^{1/k}, \quad L_0,\ L_\infty \dots$$



$X = [0.3, 0.4]$

$$\|X\| = \|X\|_2 = \sqrt{0.3^2+0.4^2} = \sqrt{0.09+0.16}$$

$$= \sqrt{0.25} = 0.5$$

(2) Dot-product of $x = \underset{2\times 1}{\begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix}}$ and $w = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is $w^T x = x^T w =$

$$= \underset{-1\cdot 0.3}{-0.3} + \underset{1\cdot 0.4}{0.4} = 0.1.$$

(3) $x = \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix}$. Find $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ s.t. $w^T x - 2 \geq 0$

$$\underbrace{w_1 \cdot 0.3 + w_2 \cdot 0.4}_{} - 2 \geq 0$$

$$10 \qquad 10$$

$$3 + 4 - 2 \geq 0.$$

(4) $x = \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix}$ $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$, bias $w_0 = -0.5$.

(4.a) $\underline{w^T x + w_0} \geq 0$? Yes $\Rightarrow \hat{y} = +1$; No $\Rightarrow \hat{y} = -1$.

$$0.1 - 0.5 = -0.4 < 0 \Rightarrow \hat{y} = -1$$

(4.b) The decision boundary is $\{ \underline{\underline{x}} \in \mathbb{R}^2 \mid w^T \underline{\underline{x}} + w_0 = 0 \}$

4.c Distance between x & d.b is. $\dfrac{|w^T x + w_0|}{\|w\|} = \dfrac{|-0.4|}{\sqrt{2}} = 0.2 \cdot \sqrt{2}$

$w = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \|w\| = \sqrt{2}$