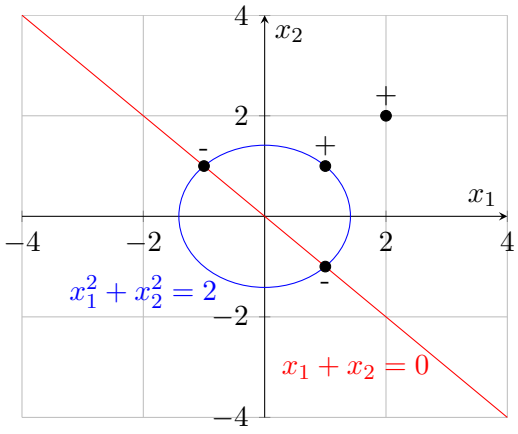
SUTD 2021 50.007 Homework 1

James Raphael Tiovalen 1004555

1. Classification

We can plot a visual, graphical representation of the given data points:



(a) Such a classifier does not exist. As demonstrated by the blue-colored circle on the graph above, a minimum radius of $\sqrt{2}$ would be needed to classify the points. However, since the classifier's boundary would intersect two positive points and one negative point (i.e., they lie on the boundary), such a classifier would be insufficient to correctly classify all said data points without any ambiguity. Any classifier of an origin-centered circle with $r < \sqrt{2}$ would classify all the data points with the same label (which is not correct for some of the points), while any classifier of an origin-centered circle with $r > \sqrt{2}$ would classify the two negative data points at $(-1, 1)$ and $(1, -1)$ along with at least the positive data point at $(1, 1)$ under the same label (which is also not correct for some of the points). Hence, no such classifier exists.

(b) One and only one possible linear classifier that goes through the origin would be one with a normal vector $\theta = (\theta_1, \theta_2) = (1, 1)$ (or $[\theta_1, \theta_2]^T = [1, 1]^T$). More formally, such a classifier can be defined as

$$h(x, \theta) = \text{sign}(x_1 + x_2) = \text{sign}(\theta \cdot x) = \begin{cases} +1, & \text{if } \theta \cdot x > 0 \\ -1, & \text{if } \theta \cdot x \leq 0 \end{cases}$$