Q36

Rahul Atre

2023-11-20

Q36 [Classification]

We have seen that in p = 2 dimensions, a linear decision boundary takes the form $\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$. We now investigate a non-linear decision boundary.

1. Sketch the curve $(1 + X_1)^2 + (2 - X_2)^2 = 4$.

Let us expand and simplify this to the exact equation of a circle:

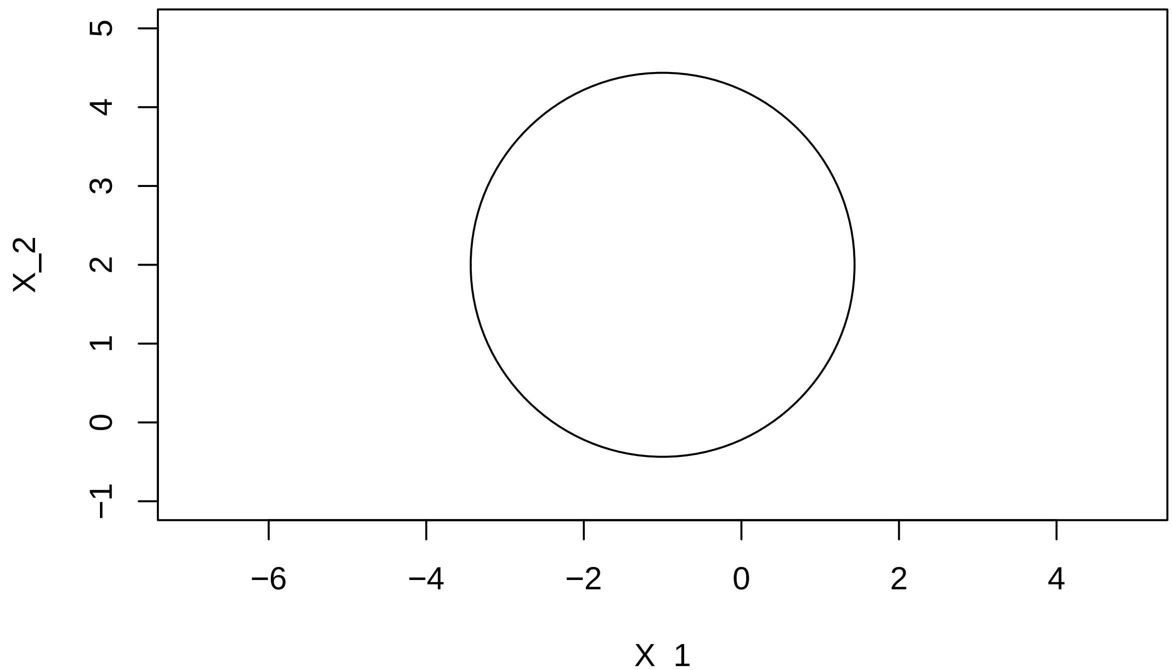
$(1 + X_1)^2 + (2 - X_2)^2 = 4$

$= (X_1 - (-1))^2 + ((-1)(X_2 - 2))^2 = 2^2$

$= (X_1 - (-1))^2 + (X_2 - 2)^2 = 2^2$

Thus, the center of the circle is (-1, 2), and radius r = 2.

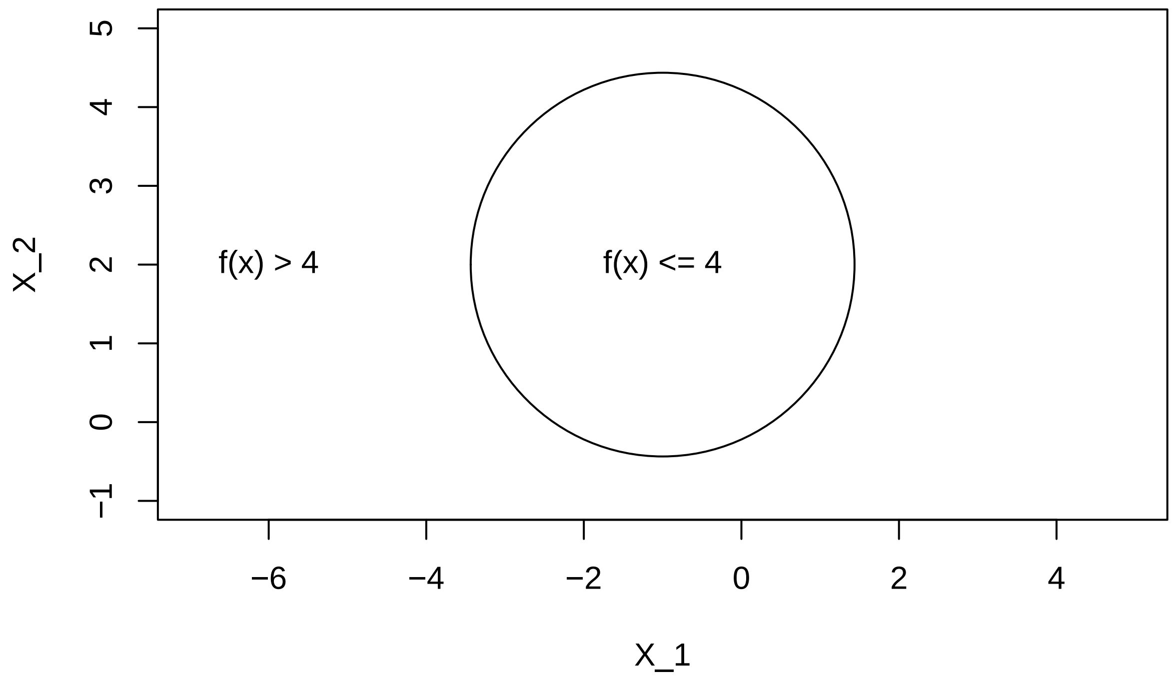
```
plot(NA, NA, xlim = c(-4, 2), ylim = c(-1, 5), asp = 1, xlab = "X_1", ylab = "X_2")
symbols(c(-1), c(2), circles = c(2), add = TRUE)
```



2. On your sketch, indicate the set of points for which $(1 + X_1)^2 + (2 - X_2)^2 > 4$, as well as the set of points for which $(1 + X_1)^2 + (2 - X_2)^2 \leq 4$.

Let f(x) rep. $(1 + X_1)^2 + (2 - X_2)^2$

```
plot(NA, NA, xlim = c(-4, 2), ylim = c(-1, 5), asp = 1, xlab = "X_1", ylab = "X_2")
symbols(c(-1), c(2), circles = c(2), add = TRUE)
text(c(-1), c(2), "f(x) <= 4")
text(c(-6), c(2), "f(x) > 4")
```



3. Suppose that a classifier assigns an observation to the blue class if $(1 + X_1)^2 + (2 - X_2)^2 > 4$, and to the red class otherwise. To what class is the observation (0,0) classified? (-1,1)? (2,2)? (3,8)?

Let us check the x_1 and x_2 values for all the following observations, to see which class it falls under.

(0,0): $(1 + 0)^2 + (2 - 0)^2 > 4 \rightarrow 5 > 4$ ==> Blue Class

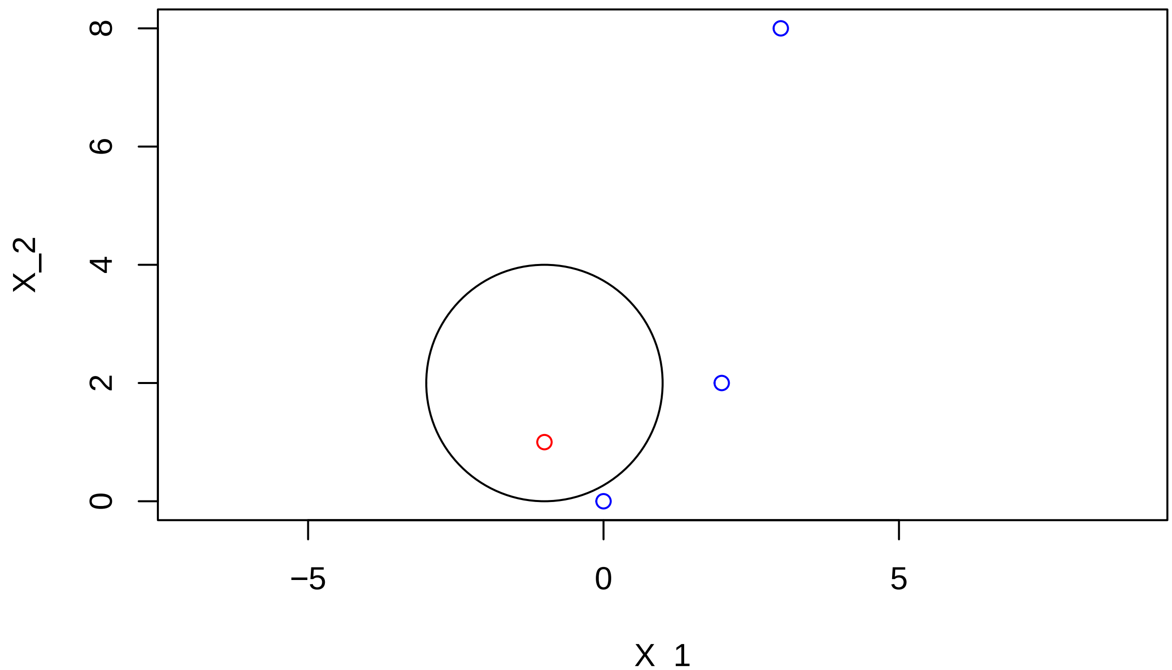
(-1,1): $(1 - 1)^2 + (2 - 1)^2 > 4 \rightarrow 1 > 4$ ==> Red Class

(2,2): $(1 + 2)^2 + (2 - 2)^2 > 4 \rightarrow 9 > 4$ ==> Blue Class

(3,8): $(1 + 3)^2 + (2 - 8)^2 > 4 \rightarrow 52 > 4$ ==> Blue Class

Plotting these results in our graph:

```r
plot(c(0, -1, 2, 3), c(0, 1, 2, 8), col = c("blue", "red", "blue", "blue"), asp = 1, xlab = "X_1", ylab
symbols(c(-1), c(2), circles = c(2), add = TRUE, inches = FALSE)
```



4. Argue that while the decision boundary in 3 is not linear in terms of X_1 and X_2, it is linear in terms of X_1, X_1^2, X_2, X_2^2.

Let us expand the equation of the decision boundary in 3 and see what result we get:

Let f(x) $= (1 + X_1)^2 + (2 - X_2)^2 - 4$.

f(x) $= X_1^2 + 2X_1 + 1 + X_2^2 - 4X_2 + 4 - 4$

f(x) $= X_1^2 + 2X_1 + X_2^2 - 4X_2 + 1$

For the above equation, we can rewrite the general form as the linear model $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_4 X_2^2$, where:

$\beta_0 = 1$ $\beta_1 = 2$ $\beta_2 = -4$ $\beta_3 = 1$ $\beta_4 = 1$

Therefore, the decision boundary is linear in terms of X_1, X_1^2, X_2, X_2^2.