ISLR | Chapter 9 Exercises

Marshall McQuillen

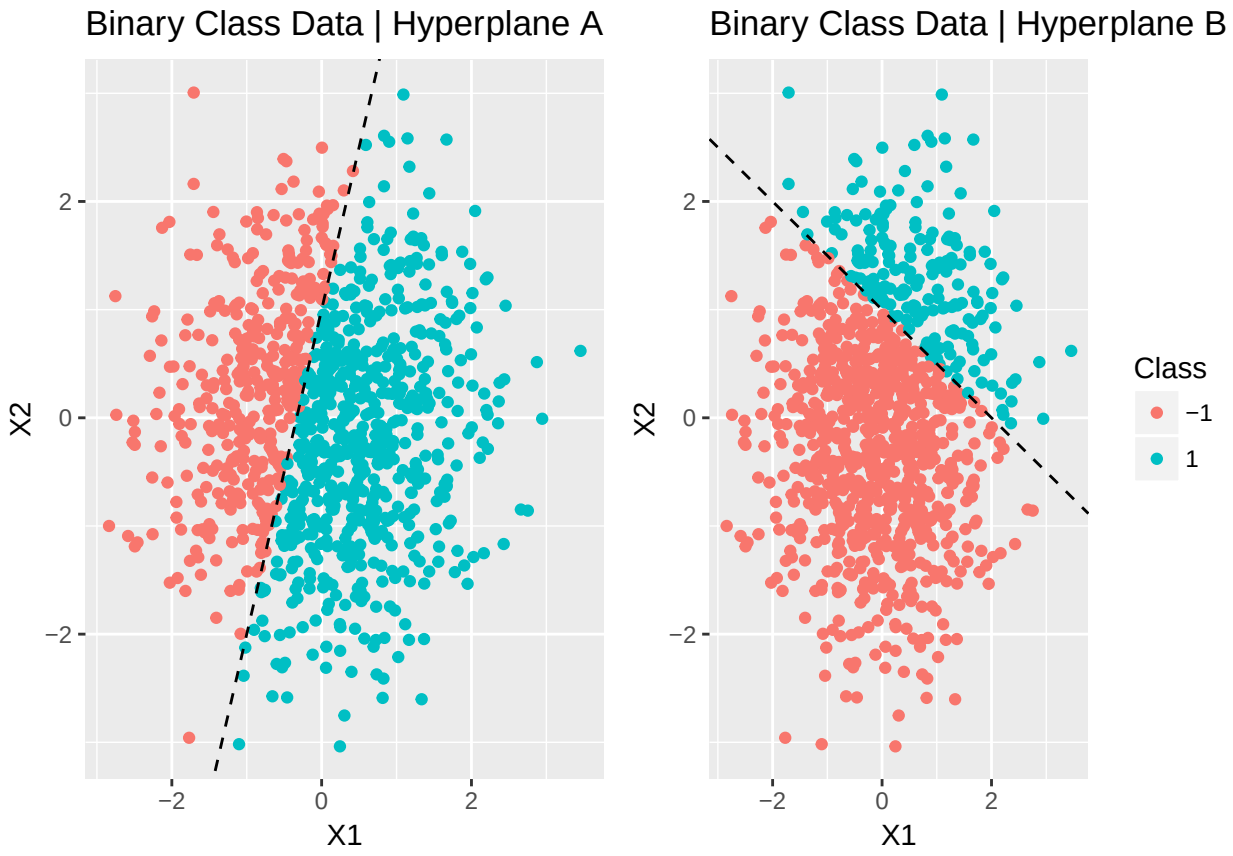
11/14/2018

Conceptual

1

- **A & B.**

```
suppressPackageStartupMessages(library(gridExtra))
suppressPackageStartupMessages(library(ggplot2))

x <- matrix(rnorm(1000*2), ncol = 2)

hyperplane1 <- 1 + 3*x[, 1] - x[, 2]
y <- ifelse(hyperplane1 > 0, 1, -1)

hyperplane2 <- -2 + x[, 1] + 2*x[, 2]
y2 <- ifelse(hyperplane2 > 0, 1, -1)

par(mfrow = c(1, 2))

plot1 <- ggplot(data.frame(x = x,
                  y = factor(y, levels = c(-1, 1))), aes(x.1, x.2, color=y)) +
             geom_point(show.legend = FALSE) +
             geom_abline(intercept = 1,
                         slope = 3,
                         linetype = 'dashed') +
             ggtitle("Binary Class Data | Hyperplane A") +
             xlab("X1") +
             ylab("X2") +
             labs(color = "Class")

plot2 <- ggplot(data.frame(x = x,
                  y = factor(y2, levels = c(-1, 1))), aes(x.1, x.2, color=y)) +
             geom_point() +
             geom_abline(intercept = 1,
                         slope = -0.5,
                         linetype = 'dashed') +
             ggtitle("Binary Class Data | Hyperplane B") +
             xlab("X1") +
             ylab("X2") +
             labs(color = "Class")

grid.arrange(plot1, plot2, ncol = 2)
```

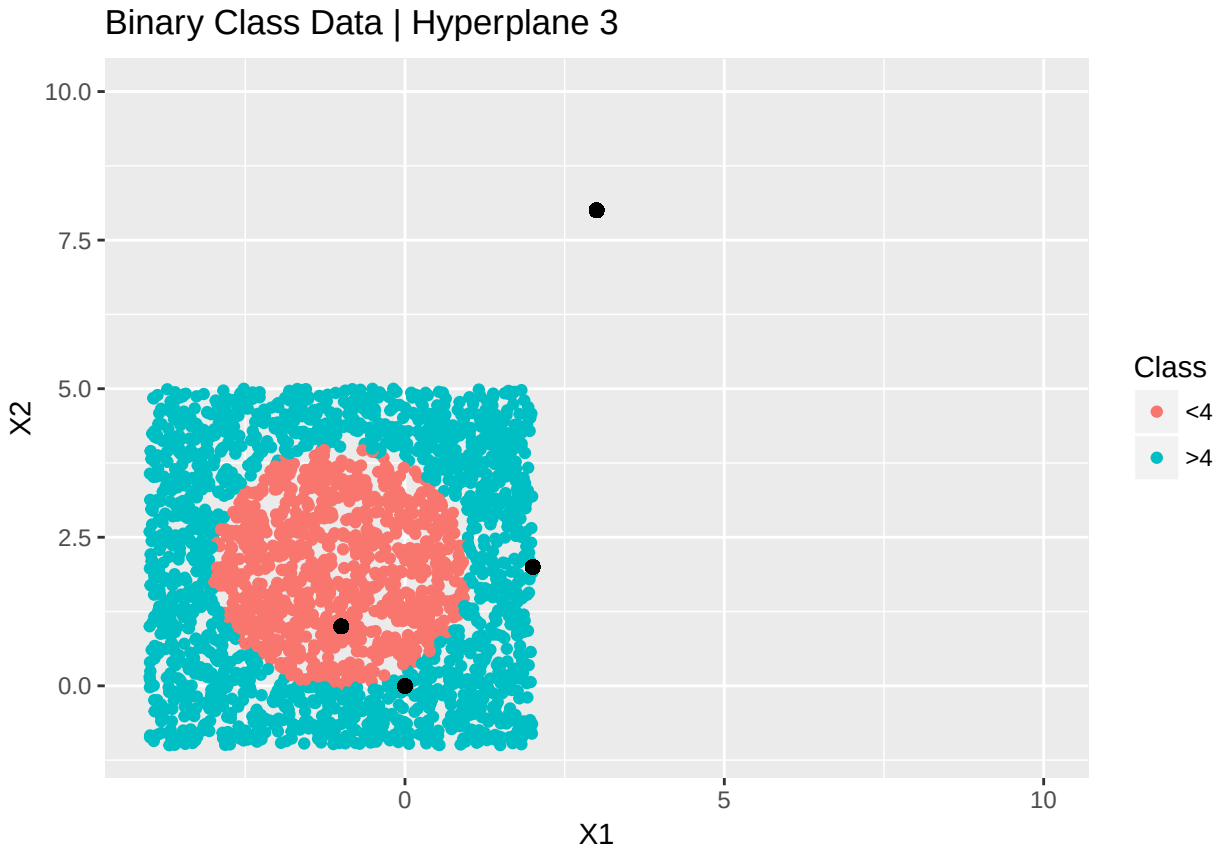


2

- **A, B & C**. The hyperplane is the circle encompassing the pink/orange data points below, where those data points are the ones who's value, when plugged into the equation $f(X_1, X_2) = (1 + X_1)^2 + (2 - X_2)^2 - 4$, will be negative. The blue data points output value of the above equation would be positive. The 4 data points plotted in black are those requested in part **C**, and it is clear which class they would fall into.

```
x_1 <- runif(2500, min = -4, max = 2)
x_2 <- runif(2500, min = -1, max = 5)

hyperplane3 <- (1 + x_1)^2 + (2 - x_2)^2 - 4
y3 <- factor(ifelse(hyperplane3 > 0, ">4", "<4"))

df <- data.frame(x1 = x_1,
                 x2 = x_2,
                 y = y3)

ggplot(df, aes(x1, x2, color=y)) +
    geom_point() +
    geom_point(x = 0,
               y = 0,
               col = 'black',
               cex = 2) +
    geom_point(x = -1,
```

```
            y = 1,
            col = 'black',
            cex = 2) +
  geom_point(x = 2,
            y = 2,
            col = 'black',
            cex = 2) +
  geom_point(x = 3,
            y = 8,
            col = 'black',
            cex = 2) +
  ggtitle("Binary Class Data | Hyperplane 3") +
  scale_y_continuous(limits = c(-1, 10)) +
  scale_x_continuous(limits = c(-4, 10)) +
  xlab("X1") +
  ylab("X2") +
  labs(color = "Class")
```



- **D**. One can see that the equation given in the text is non-linear with regard to X_1 and X_2. However, when expanded and refactored, it is clear that the hyperplane is linear with regard to to X_1, X_2, X_1^2 and X_2^2.

$$(1 + X_1)^2 + (2 - X_2)^2 = 4 \tag{1}$$
$$(1 + X_1)^2 + (2 - X_2)^2 - 4 = 0 \tag{2}$$
$$(1 + 2X_1 + X_1^2) + (4 - 4X_2 + X_2^2) - 4 = 0 \tag{3}$$
$$1 + 2X_1 + X_1^2 - 4X_2 + X_2^2 = 0 \tag{4}$$
$$\tag{5}$$