**INF552 Assignment3**

**6.8.3.**

*(a) As we increase s from 0, the training RSS will:*

iv. Steadily decrease. As we increase s from 0, the coefficients will gradually increase to their least squares estimates, and so the model is becoming more and more flexible which provokes a steady decrease in the training RSS.

*(b) Repeat (a) for test RSS.*

ii. Decrease initially, and then eventually start increasing in a U shape. As we increase s from 0, the coefficients will increase to their least squares estimates. The flexibility will increase, so the test RSS will first decrease then increase.

*(c)Repeat (a) for variance.*

iii. Steadily increase. Same as above, when model becomes more flexible, the variance will increase correspondingly.

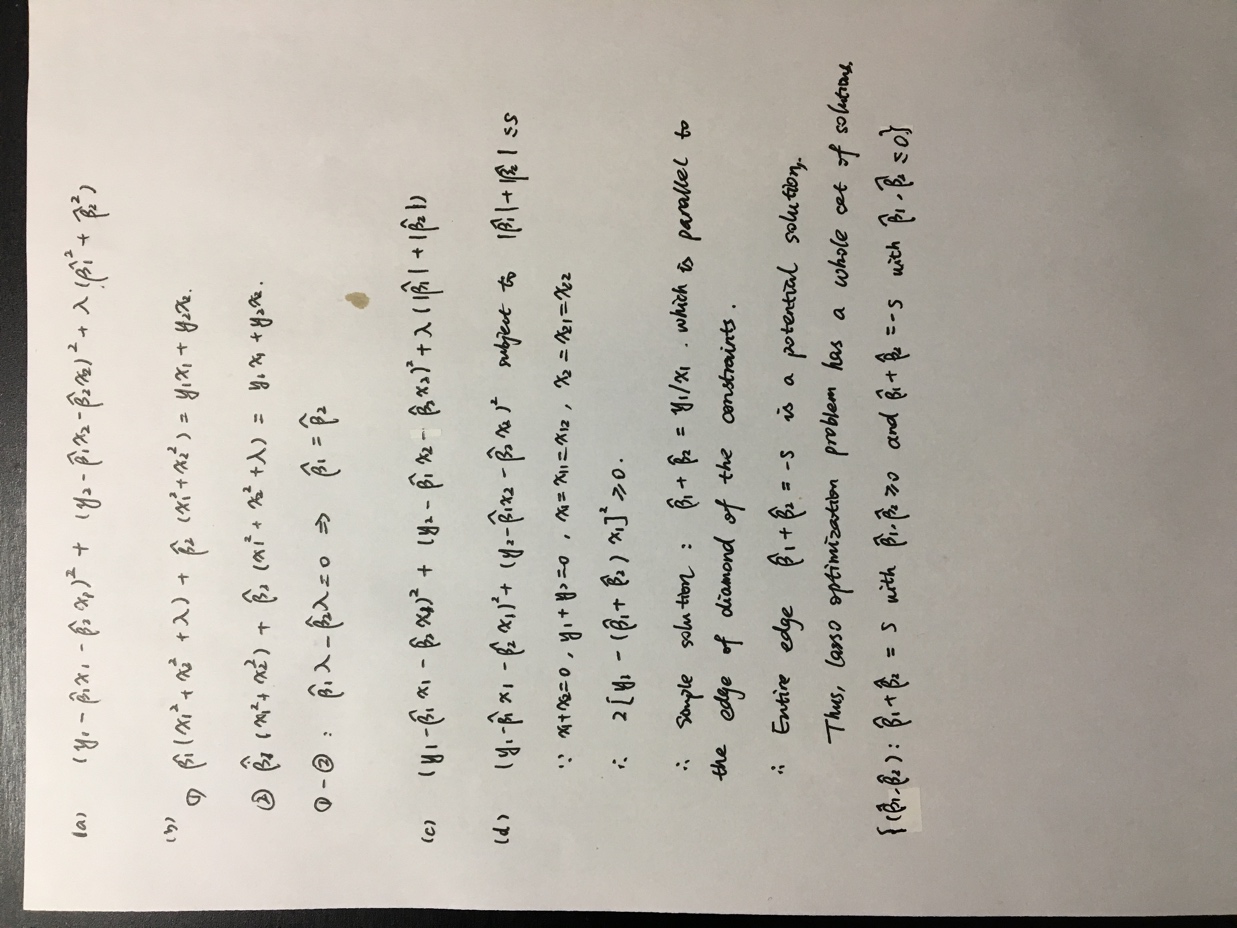
*(d) Repeat (a) for (squared) bias.*

iv. Steadily decrease. The bias will decrease when flexibility of the model increase.

*(e) Repeat (a) for irreducible error.*

v. Remain constant. The irreducible error is independent of the model, so the change of parameters in model won’t change irreducible error.

**6.8.5**



**8.4.5**

*There are two common ways to combine these results together into a single class prediction. One is the majority vote approach discussed in this chapter. The second approach is to classify based on the average probability. In this example, what is the final classification under each of these two approaches ?*

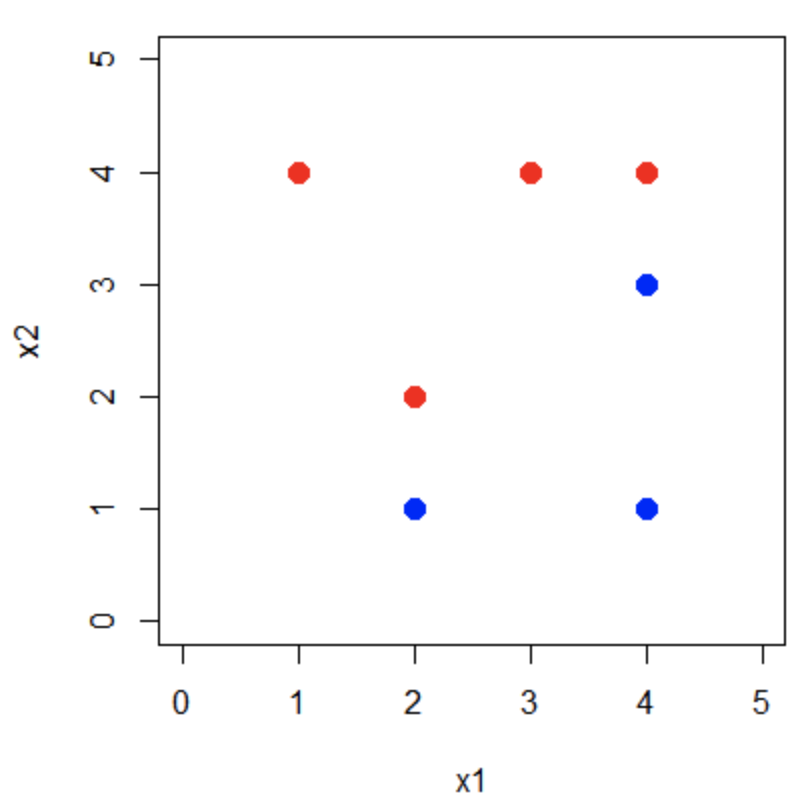
With the majority vote approach, we classify X as Red because 6 predictions (in 10 predictions) are classified as red.

With the average probability approach, we classify X as Green as the average of the 10 probabilities is 0.45*.*

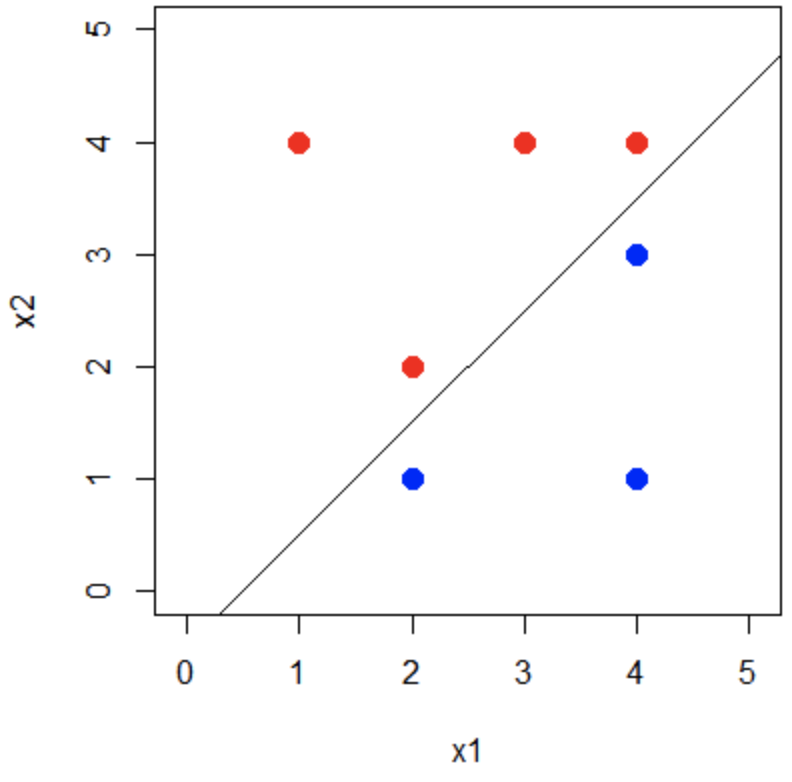
**9.7.3**

*Here we explore the maximal margin classiﬁer on a toy data set.*

*(a) We are given n = 7 observations in p = 2 dimensions. For each observation, there is an associated class label.*



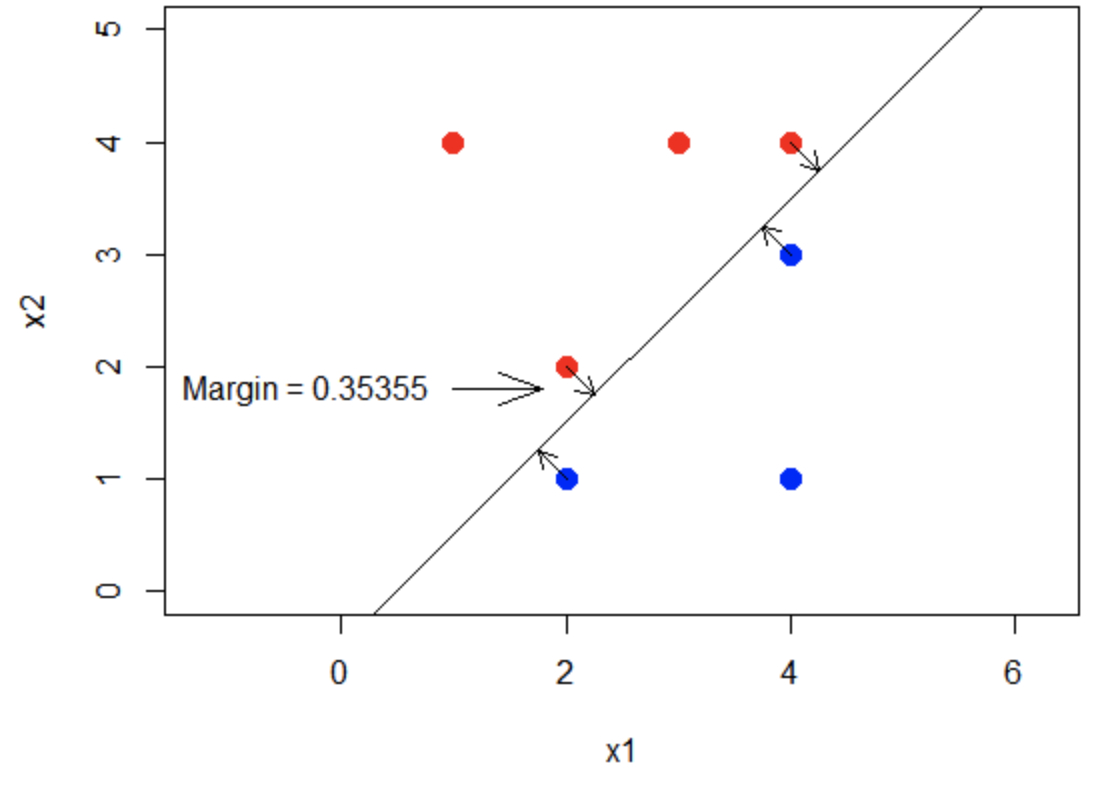
*(b) Sketch the optimal separating hyperplane, and provide the equation for this hyperplane (of the form (9.1)).*



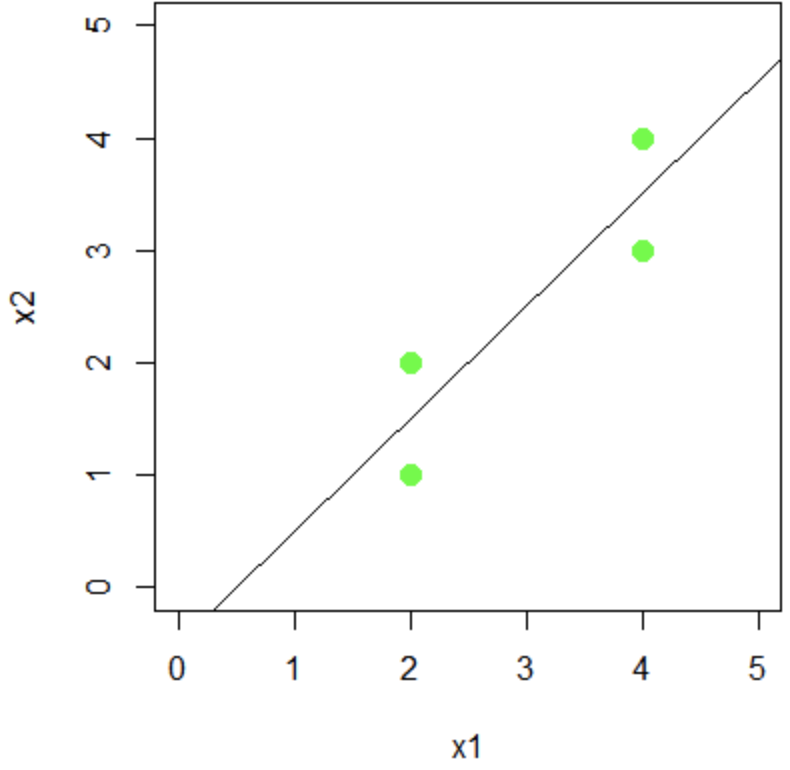
*(c) Describe the classiﬁcation rule for the maximal margin classiﬁer. It should be something along the lines of “Classify to Red if β 0 + β 1 X 1 + β 2 X 2 > 0, and classify to Blue otherwise.” Provide the values for β 0 , β 1 , and β 2 .*

The classification rule is “Classify to Red if X1−X2−0.5<0, and classify to Blue otherwise.”

*(d) On your sketch, indicate the margin for the maximal margin hyperplane.*



*(e) Indicate the support vectors for the maximal margin classiﬁer.*

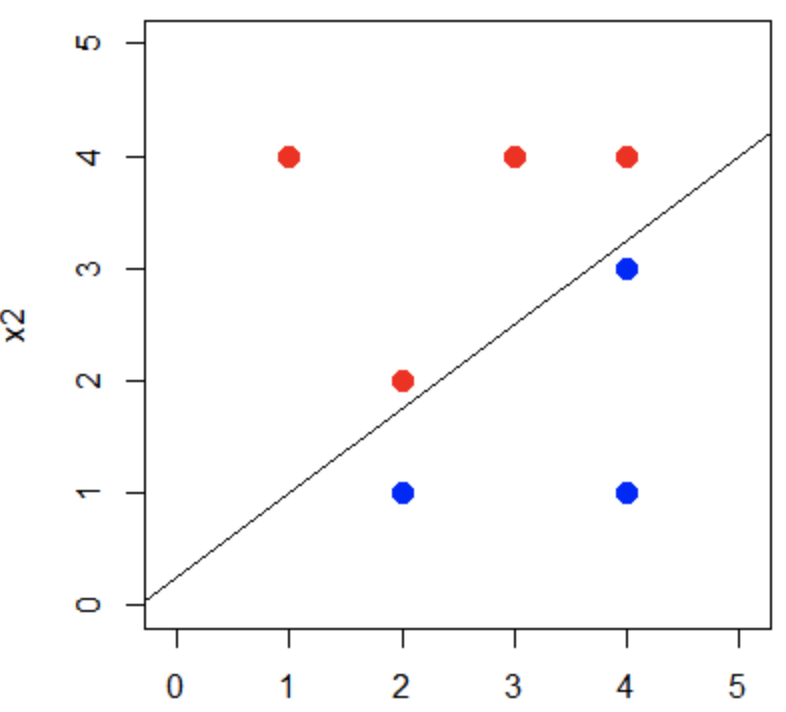


Support vectors in green (2,1), (2,2), (4.3), (4,4)

*(f) Argue that a slight movement of the seventh observation would not aﬀect the maximal margin hyperplane.*

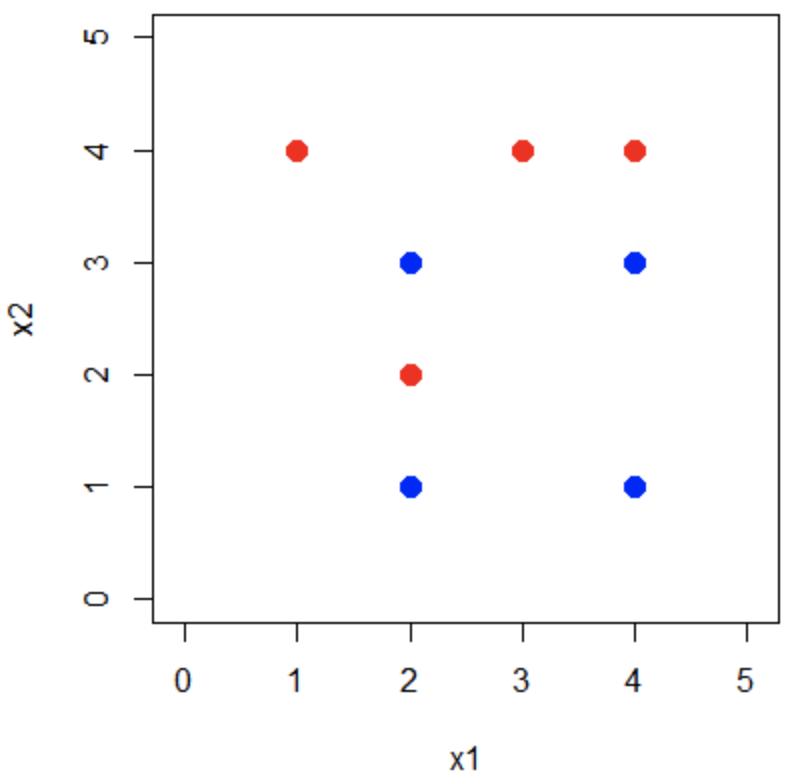
Since the seventh observation (4,1) is much further away from the hyperplane than the support vectors, and since the maximal margin hyperplane is only determined by the support vectors, a slight movement of the point would have no affect on the hyperplane.

*(g) Sketch a hyperplane that is not the optimal separating hyperplane, and provide the equation for this hyperplane.*



0.25+0.75X1−X2=0

*(h) Draw an additional observation on the plot so that the two classes are no longer separable by a hyperplane.*

**