1.Linear Classification

1.1

$$C = Sign(\langle \binom{V_h}{-c} | \binom{x}{1} \rangle)$$

$$C_1 = Sign(\langle \binom{V_h}{-c} | \binom{x_1}{1} \rangle) = Sign\left(\frac{-3}{\sqrt{2}} + \frac{-1}{2\sqrt{2}}\right) = -1$$

$$C_2 = Sign(\langle \binom{V_h}{-c} | \binom{x_2}{1} \rangle) = Sign\left(\frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} - \frac{1}{2\sqrt{2}}\right) = 1$$

1.2

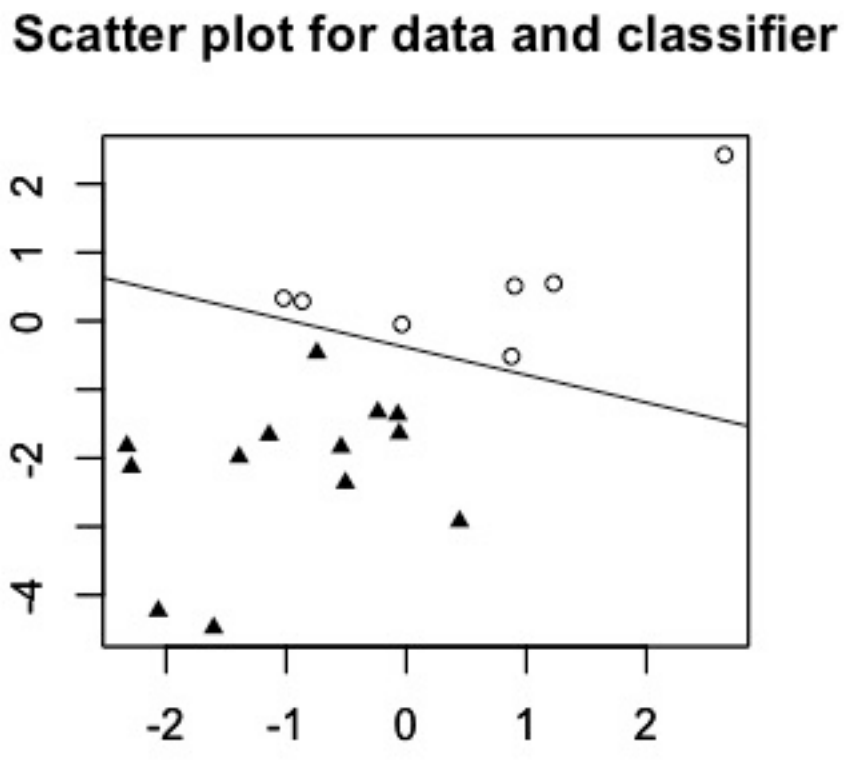
Result will not change because the classification result only depends on formula $C = Sign(\langle \binom{V_h}{-c} | \binom{x}{1} \rangle)$, though perceptron and SVM use different methods to get value of V_h and c.

1.3

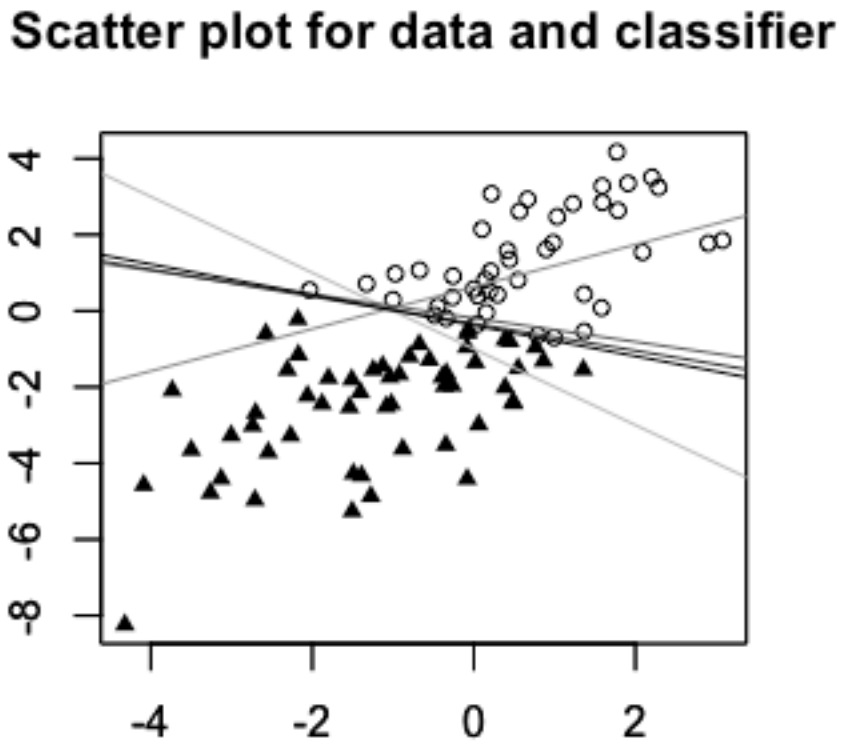
Perceptron cost function approximate to empirical risk function. Since empirical risk function is piecewise constant, and we want to checking the gradient function of cost function so we make it piecewise linear.

2. Perceptron

Test data and classifier hyper plane



Training data and classifier hyper plane trajactory

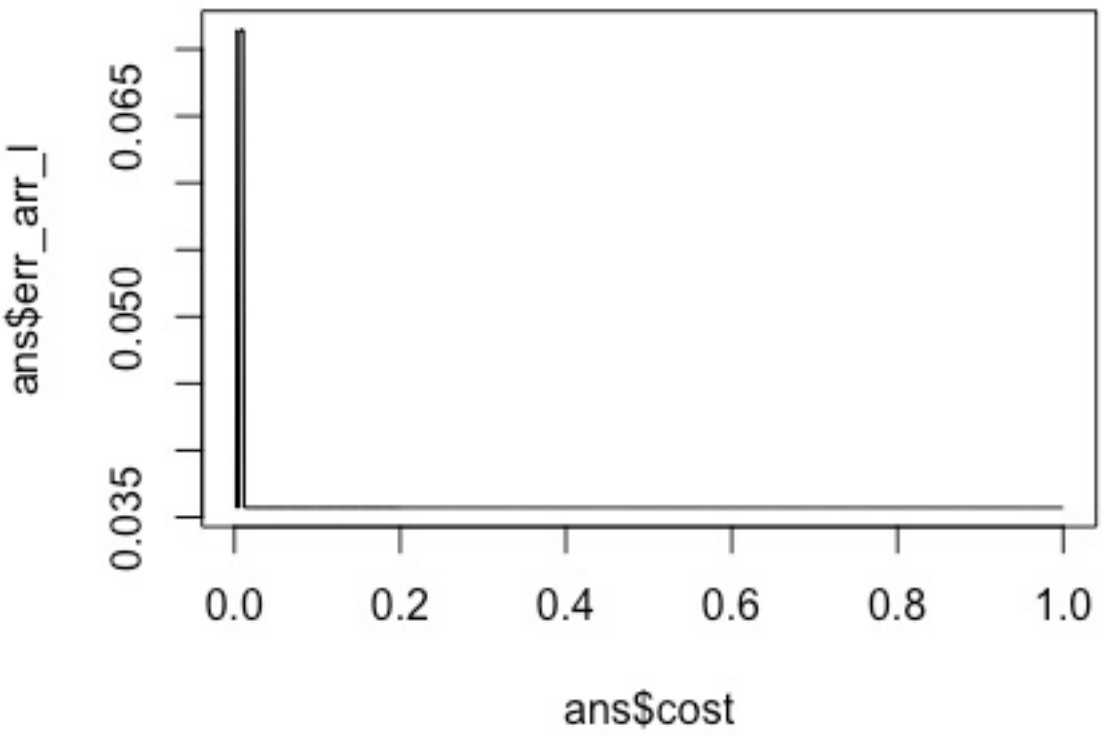


In this plot, lighter gray line is where I start, and black line is where I final arrived at.

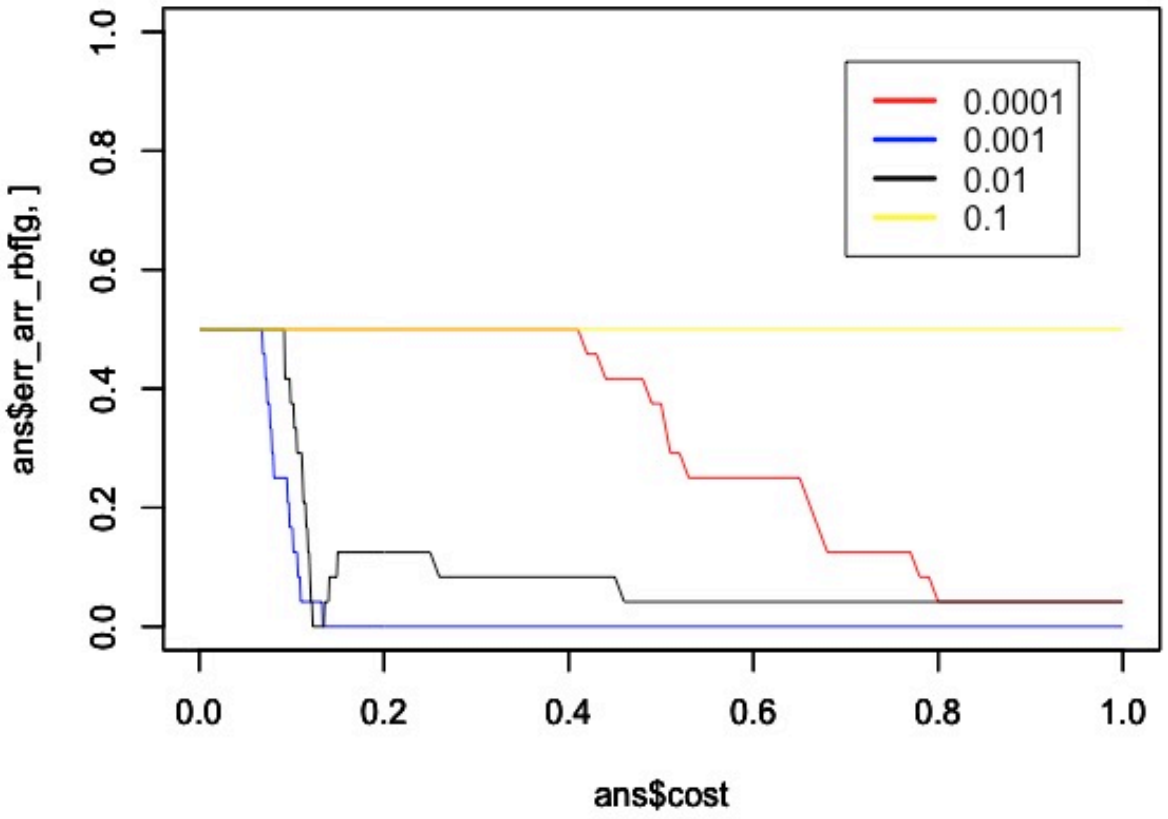
3. SVM

3.1

Cost vs Misclassication rate – Linear SVM



Cost vs Misclassication rate – Radial SVM



3.2

Since the data is random selected and I do whole process for 50 times, and I selected 3 out of 50. The results and parameters selected reported as followed:
[1] "The error rate of linear_svm on test data is 0.046512, with linear_cost 0.001000"
[1] "The error rate of radial_svm on test data is 0.116279, with rbf_cost 0.116000 and gamma 0.010000"

[1] "The error rate of linear_svm on test data is 0.000000, with linear_cost 0.093000"
[1] "The error rate of radial_svm on test data is 0.024390, with rbf_cost 0.091000 and gamma 0.001000"

[1] "The error rate of linear_svm on test data is 0.051282, with linear_cost 0.001000"
[1] "The error rate of radial_svm on test data is 0.051282, with rbf_cost 0.183000 and gamma 0.001000"

The error rates of both linear and non-linear is below 0.1 which is acceptable, and for three times experiment there is no sign that one totally outperforms another.
In addition, non-linear classifier cost more time to train and validate. So in all I will choose linear one because it is cheaper and as good as non-linear one.