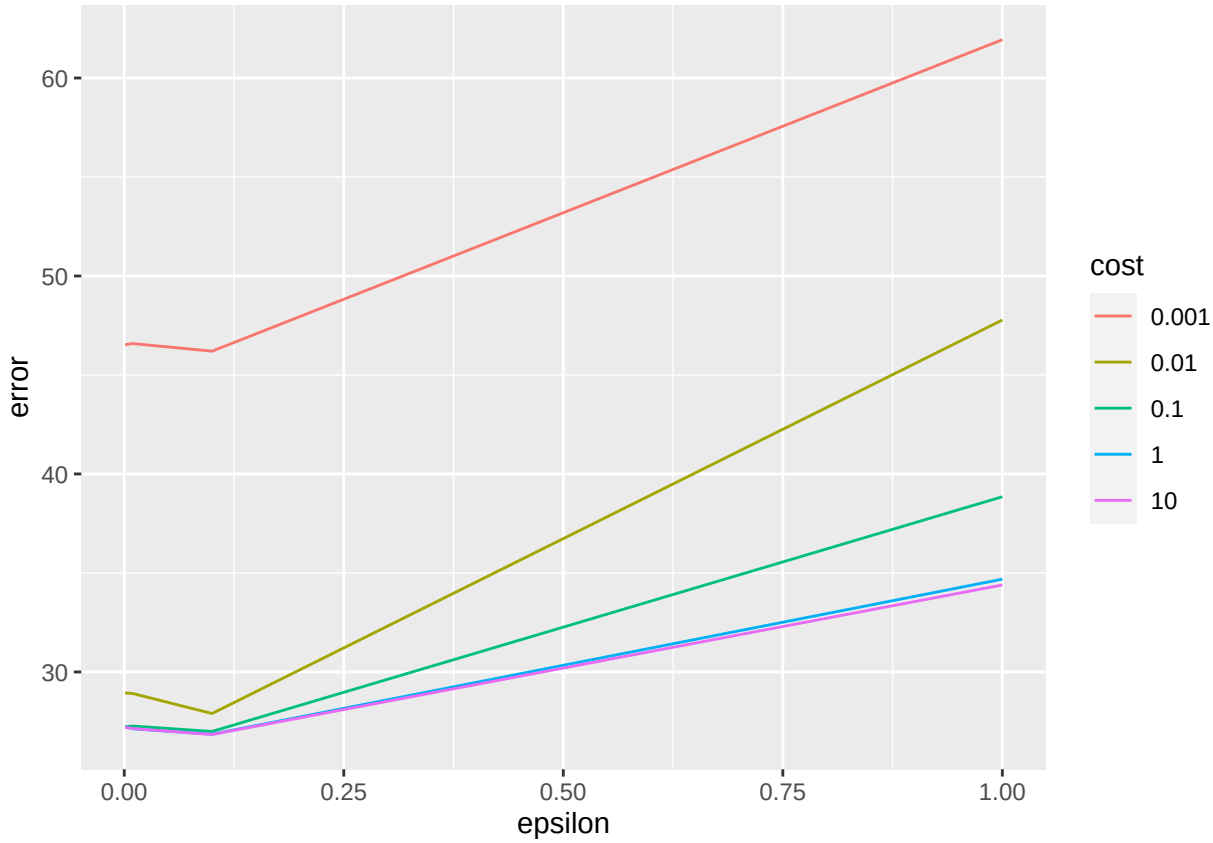
# stat_learning_assignment_11_JD

Joshua Damm

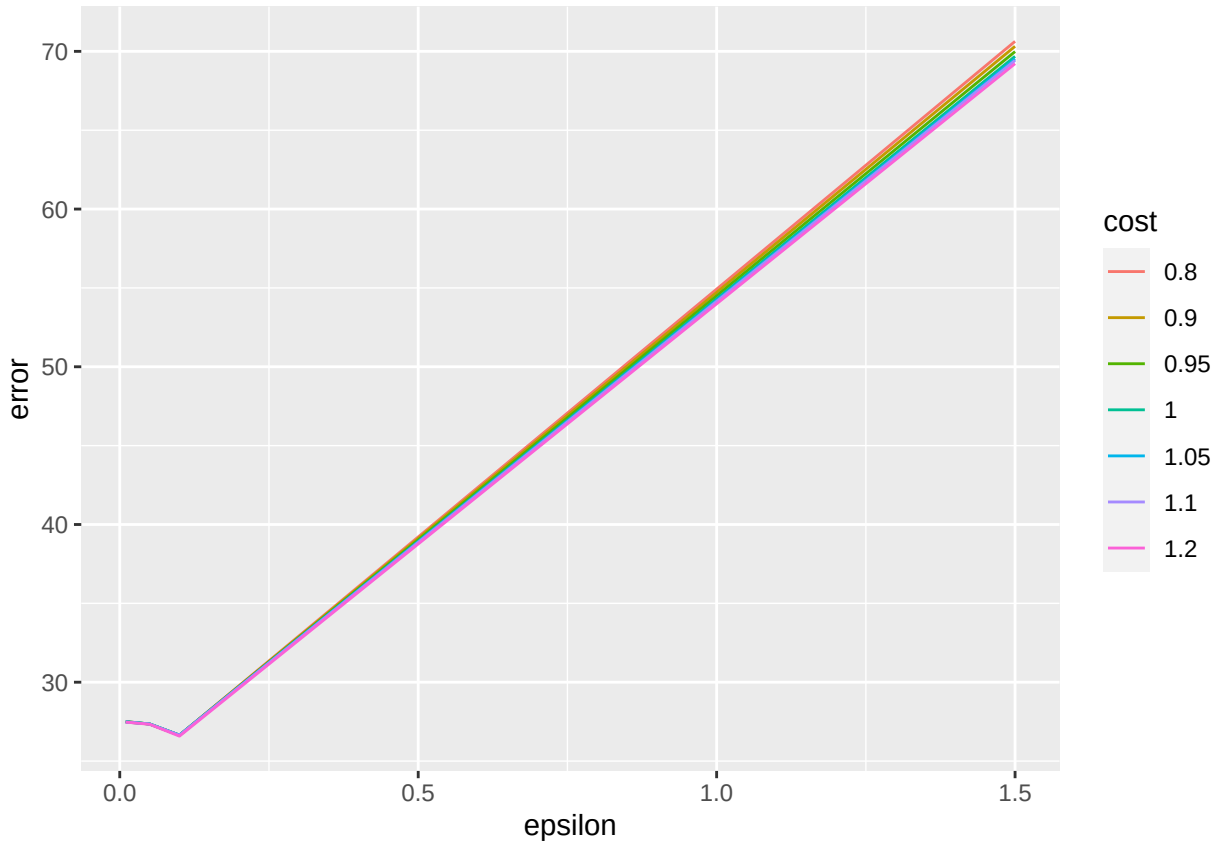
2024-05-06

## Question a)




```
##    cost epsilon
## 15   10     0.1
```

Best tuning values after grid search are cost = 1 and epsilon = 0.1.




```
##    cost epsilon
## 21  1.2     0.1
```

Also after trying finer values for epsilon and c (cost), cost = 1 and epsilon = 0.1 yield the best results (lowest CV error in training data). One can see that we got a rather convex curve since the CV error reaches a minimum and then increases again. The original was fine enough, since we reached the same optimal values for a much finer grid.

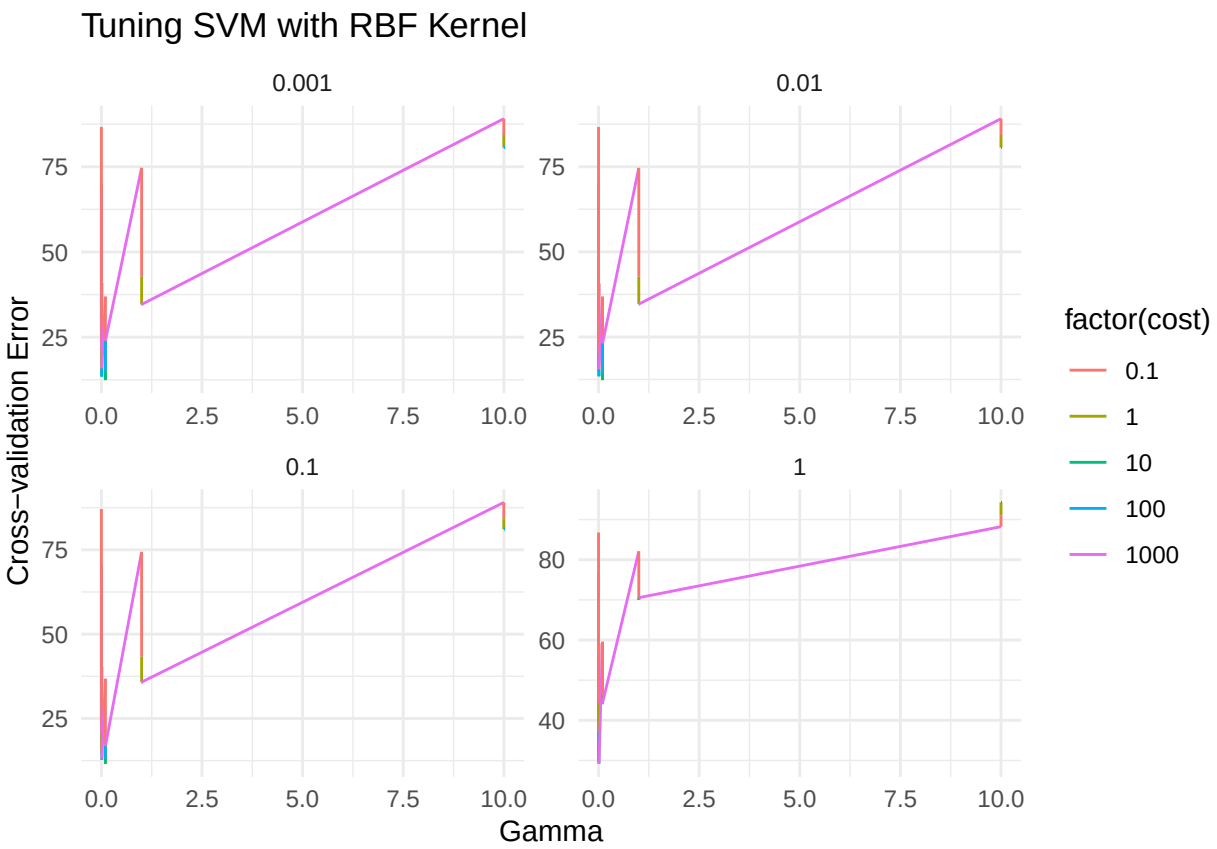
```
## [1] "MSE: 21.7482157379689"
```

```
## [1] "MAE: 2.95978491588339"
```

Based on the MSE and MAE I would say that the support vector regression does a solid job of predicting the outcome.

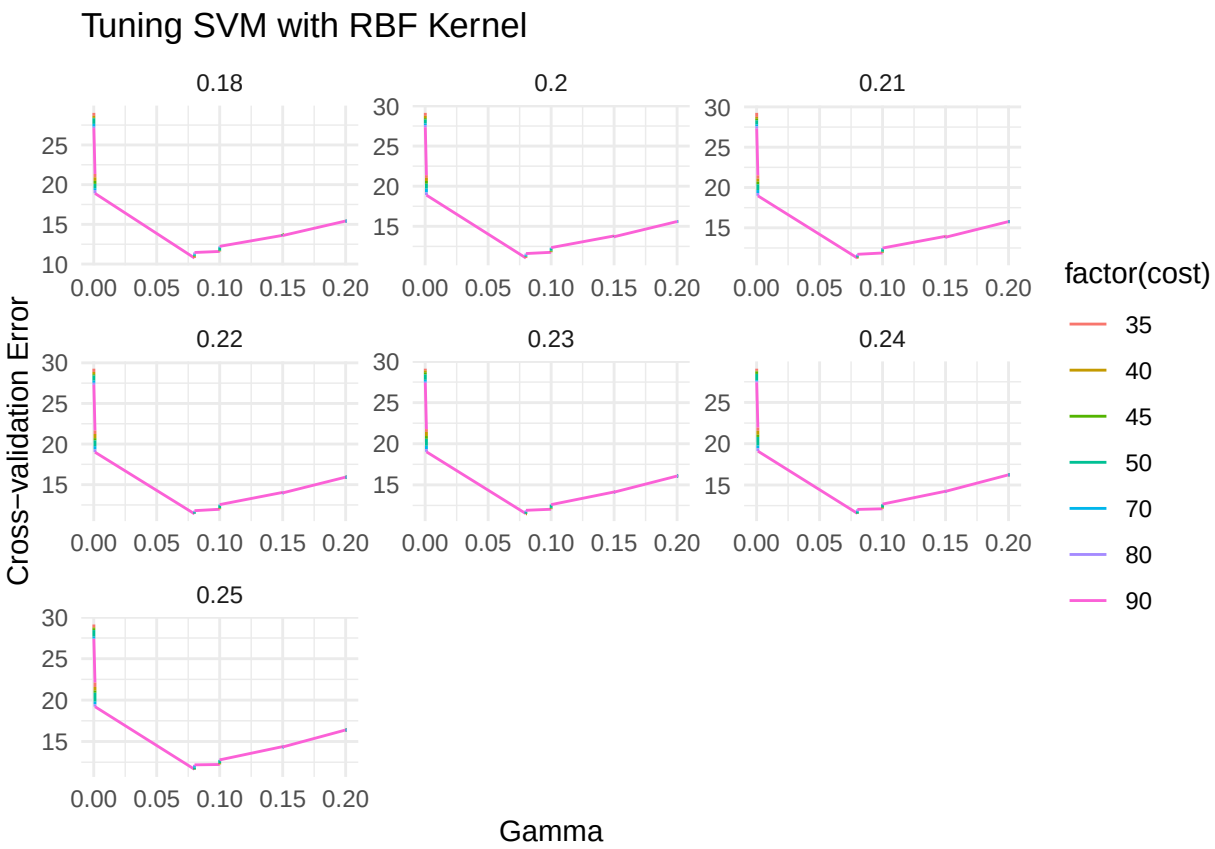
# Question b

```
##    gamma cost epsilon
## 76   0.1   10     0.1
```




In the intial grid, a cost parameter of 10, a gamma of 0.1 and epsilon of 0.1 were computed for the lowest CV error. Let's adjust the grid and see if we can increase the model performance.

```
##   gamma cost epsilon
## 3  0.08   35    0.18
```




After adjusting the grid a few times I got an optimal gamma of 0.08, cost of 35, and epsilon of 0.18 for lowest CV error on the training data. One can also see that we get a nicely convex curve. Let us refit the model and see how it fits the test data.

```
## [1] "MSE: 8.47466594386089"
```

```
## [1] "MAE: 1.93831266518836"
```

The MSE and MAE suggest a good model performance and a substantially lower prediction error than the SVM model with the linear kernel.