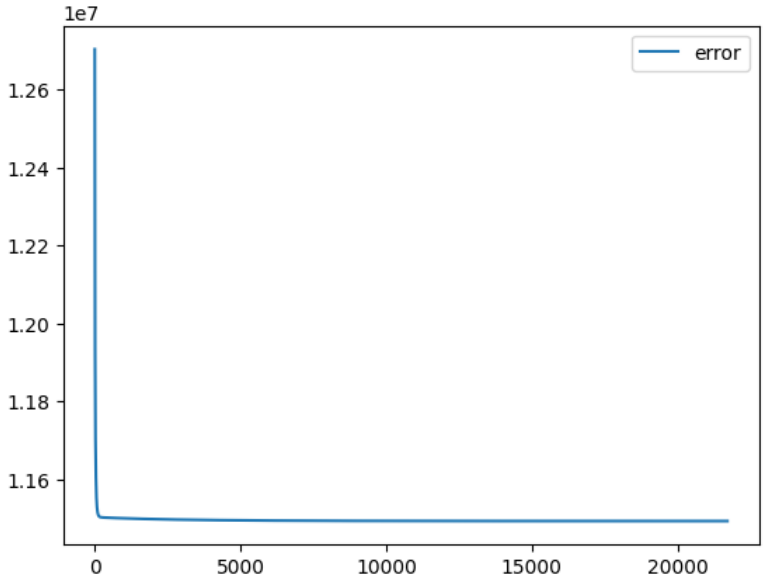
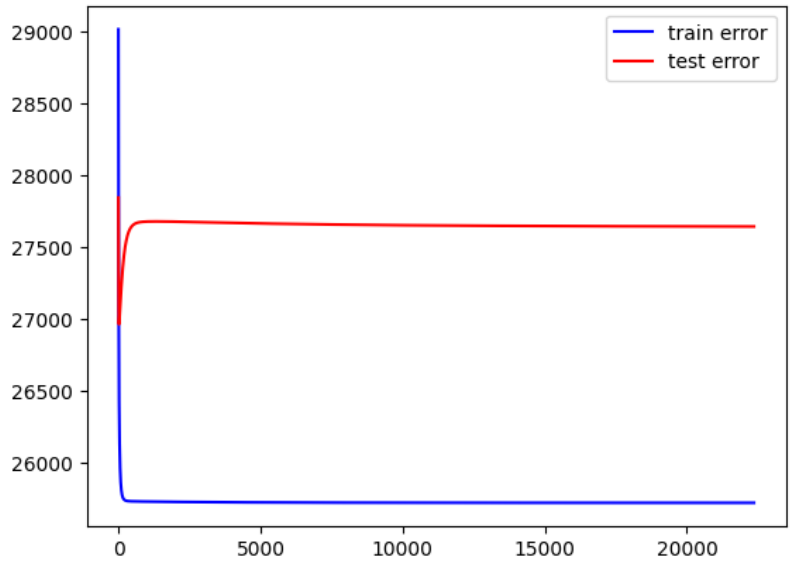
Ex1:

With step size $\epsilon = 1e - 2$ and stop condition $\delta = 1e - 1$:



Ex2:



The algorithm overfits, as you can see based on the graph after a few iterations the test error is around 27000 and afterwards begins to increase.

thus we can conclude that the model over performed on the training data and that made the model errors on the test data larger.

Ex3:

average train error :25836.823136084182, average test error :27362.638156956877

minimum train error :25270.332808434698, minimum test error :24295.857802405324

relative average : 0.059055829458448915

relative minimum : 0.038562017105849694

based on the errors we get from x* (the optimal x gradient descent returns). After calculating the relative error between train error and test error we can conclude that we cant see a pattern of overfitting in here.

That is because the relative average error and the relative minimum error are both small.

In the previous exercise we indeed saw a pattern of overfitting which we haven't seen now. Also ,We do get average train error around 26000 which, although the X's are vector of a big size which may cause this large error, perhaps we could use a different model which will fit a better predictions for the data. thus we may be in an underfitting case which we cant know for sure.