1. d) Report

i) Your most general code

Look above for general code

ii) Why you are sure your code works. That is, what test cases did you use and why are they general?

- The code works because the error vs iterations has reached a saturation point.
- Doing more updates will reduce error very minimally or not at all.
- The test cases have been generated the same way the train error was generated, thus general.
- In the CSV case the data was shuffled and split into 70/30 ratio, thus ensuring generality

iii) What were the convergence criterion and learning rate you used? Did you experiment with the learning rate? Do you have any comments on this?

Convergence was determined by looking at error vs iterations. A better way would be to stop when update size or error reduction reaches a small epsilon. Since this is toy data, plots were sufficient. Learning rate was a big issue, it was picked by testing values from both extremes so the lower values didn't reduce error much and higher ones shooter up the error by bouncing. Thus a value in the middle was chosen iteratively. Again plots were helpful.