

Part 1 Question 1- The convergence method that I picked was based off of the lecture slides which was to see how much the current iteration differed from the previous iteration. The way I measured this was I stored all the deltas (absolute value of current iteration - previous iteration) for each state in the matrix in an array, and then I checked to see if any of the deltas are greater than the threshold value I picked which was $1e-6$ which was also from the lecture slides. The algorithm was considered to be converged if the delta for every single state in the matrix was below the threshold.

Part 2 Question 2- I did the same convergence method that I described above for the policy iteration. However for this algorithm, because we are picking the max value and all the numbers in the matrix are integers, and because the threshold of $1e-6$ is so low, it is basically just going to keep going until the matrix from the last iteration is exactly the same as the matrix from the current iteration.

Part 2 Question 3- It took 493 iterations to converge for the policy iteration, and only 5 iterations for the value iteration. Someone may choose to use policy iteration for a multitude of reasons such as in an environment with infinite actions because this would be impossible for a value iteration as it would not be possible to calculate the value for every single possible action or in cases where you want to see the development of the policy for each iteration rather than just the optimal policy