

IEMS-469 Dynamic Programming HW1

Weijia Zhao ¹
Kellogg Finance

This version: October 21, 2021

¹Weijia.Zhao@kellogg.northwestern.edu

1 Single Class Problem

The order of events at time t is: dispatch of a shuttle \rightarrow new consumers come \rightarrow calculate waiting cost.²

I use pretty strict converging criteria in all the exercise, the maximal iteration possible is 10^7 (for value iteration, the outer loop and inner loop of policy iteration) and the error tolerance is 10^{-6} (again, for value iteration, the outer loop and inner loop of policy iteration).

I finished the homework assignment before this week's lecture and I use V for both the value function with only state as its argument (" V " in reinforcement learning) and the value function with both state and action as its argument (" Q " in reinforcement learning) following Econ/Fin tradition.

The initial value function or initial policy function in any of the three methods can be initialized randomly³. I initialize the value function to be all 0 and policy function to be all 0.

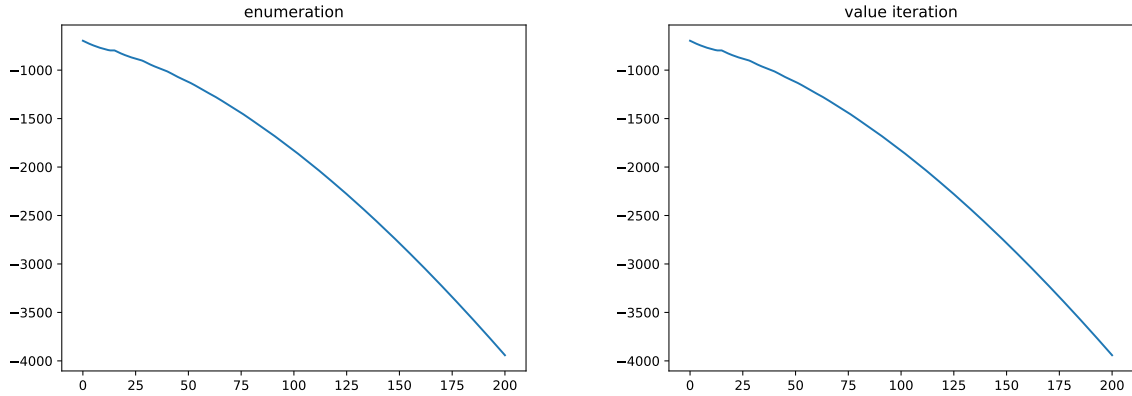
The usage of $V(s)$ and $Q(s,a)$ can be interchangeable as these two functions are closely related to each other. In enumeration and value iteration methods, $V(s) = \max_a Q(s,a)$. In policy iteration, actions are given by policy π . Then we only fill the only policy cells when construct $Q(s,a)$, leave all the off policy cells blank, and only focus on those non blank cells, we get back to $V(s)$.

All the graphs follow my expectation, we have a cut-off strategy, dispatch a shuttle if and only if the number of people waiting is above a certain threshold. The value function is a decreasing function of the number of people waiting.

²One can easily modify the code to reflect a different assumption under which new consumers can be picked by immediately by the shuttle. This is even more straightforward given that I use `s_remaining` and `s_remaining_new` to represent the number of consumers waiting in the line after the shuttle pick up and after the new consumers come. In the end the results do not change much for these two assumptions except that the value function under primary assumption equals to the value function under alternative assumption minus a constant, as all consumers need to wait at least one period under the primary assumption

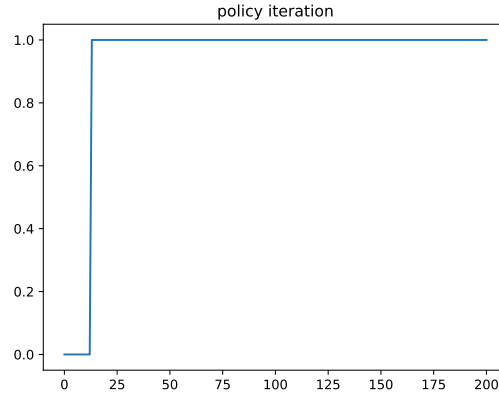
³Enumeration method requires V to be all 0 as the residual value after the last period should be 0 but in this example, T is so large that it will converge after 500 iterations so that doesn't really matter. So the enumeration method is actually identical to value iteration method except the counters/termination requirement of iteration are different

Figure 1: Single Class Problem



(a) Enumeration: Value Function

(b) Value Iteration: Value Function



(c) Policy Iteration: Policy Function

2 Multi Class Problem

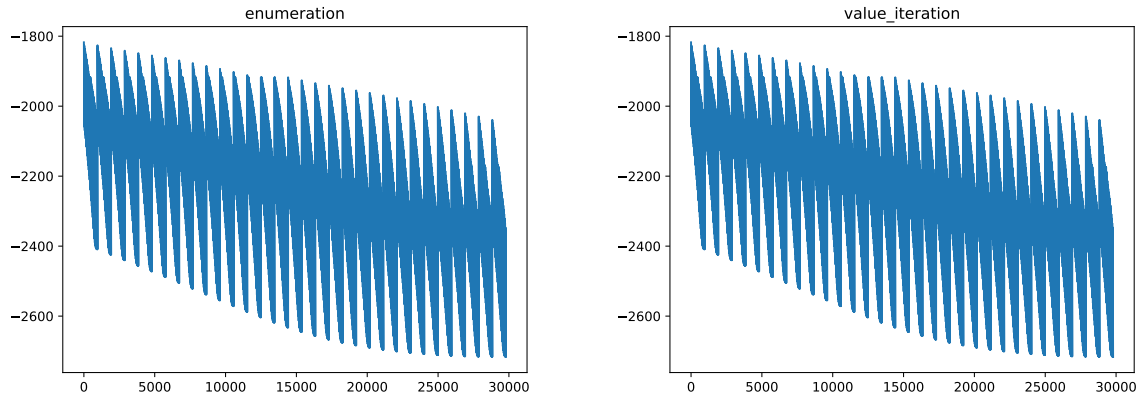
Apparently with the parameters suggested by the homework assignment, the state space is $100 * 100 * 100 * 100 * 100 = 100000000000 = 10^{10}$ and the transition matrix contains 10^{20} elements. This type of curse of dimensionality prevents us from getting any meaningful results in reasonable time period. Instead, I reduce the state space by looking at individuals with waiting time $\{2, 2.5, 3\}$ and the limit of each type is 30^4 . I get the following results, where states are indexed by $(num1, num2, num3)$ where num_i is the number of people waiting in the line that are of i th type.⁵

⁴Parameters suggested in the homework are listed in commented area

⁵The states are first sorted by $num1$ then $num2$ then $num3$. That is the reason why the density of policy function taking value 0 decrease because at the beginning, $num1$ is small and we allow $num2$ and $num3$ to be slightly larger but as $num1$ getting larger the threshold for $num2$ and $num3$ also decrease.

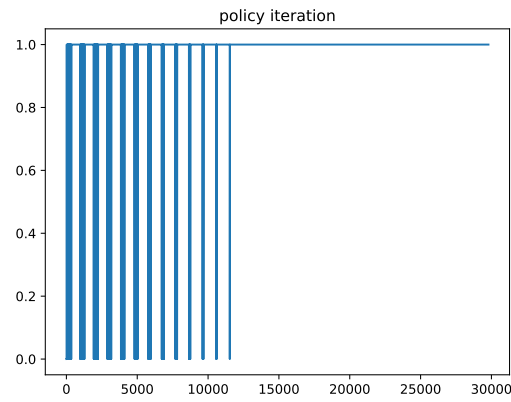
I also tried to use the parameters given in Q1 to verify the correctness of my code and it turns out they generate exactly the same results.

Figure 2: Option Characteristics



(a) GDP in year 2016

(b) Bond value matched



(c) GDP in year 2016