Exercise 4.5 ***(programming)*** Here, as a check of my program, I will first replicate the results of the Example 4.2: Jack's Car Rental. We start from the policy where no cares are moved, see Fig. 1. The policy iteration requires only four iterations until the optimal policy is found, see Fig. 2. The final state-value function is depicted in Fig. 3. There is a slight difference between the policies found by my program and those from the textbook. Reading the source-code of this Example provided on GitHub, one can find a difference in the treatment of this problem. They assumed an upper bound (of the value 11) for the number of returned or requested cars, which means that the probabilities of getting greater value than the upper bound are truncated to zero. However, my treatment uses the cumulative probabilities of getting values of returned or requested cars which would lead to the same outcome (using the fact that there cannot be rented more cars than available and there cannot be stored more cars than 20 at each location). Arbitrarily cutting the probabilities is not a rigorous treatment of the problem.

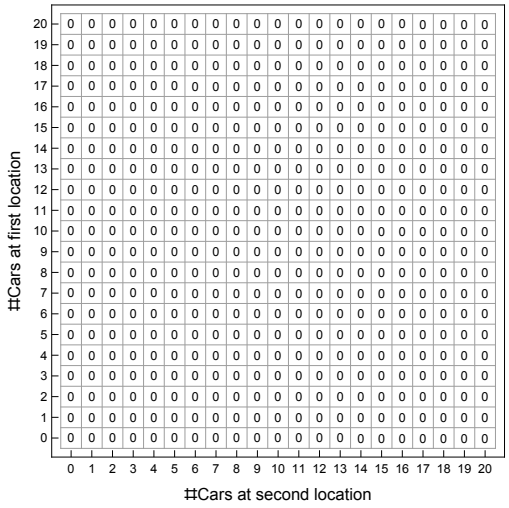

Figure 1: *Example 4.2*: The initial policy π_0.

Now, let us present the results for the Exercise 4.5, which is a modified version of the original example. Again, we start from the policy π_0 where no cars are moved. Fig. 4 shows the sequence of policies found by the policy iteration. The final state-value function is showed in Fig. 5.

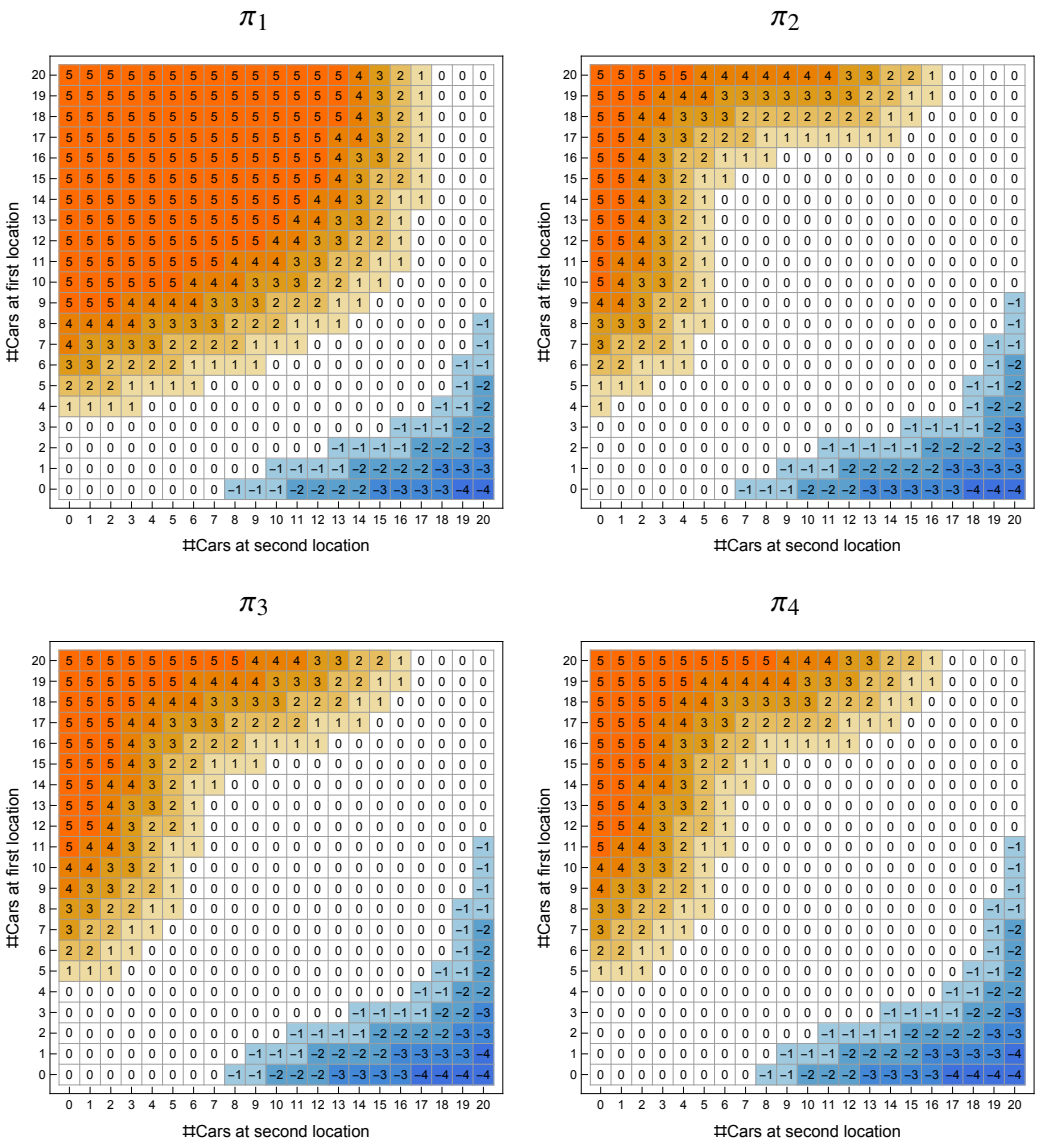

Figure 2: *Example 4.2*: The sequence of policies found by policy iteration on Jack's car rental problem. The policy iteration algorithm was initialized by the policy that never moves any cars π_0, depicted in Fig. 1.

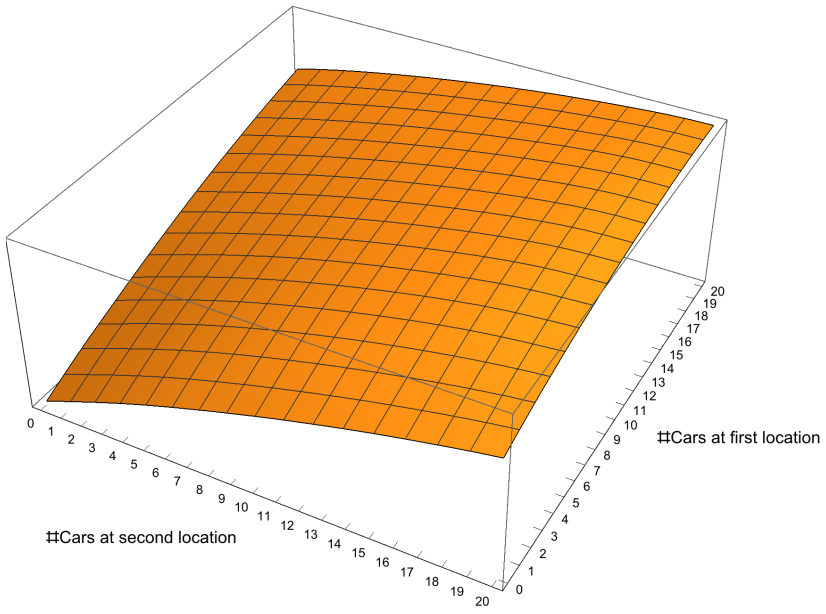

Figure 3: *Example 4.2*: Final state-value function for the Example 4.2.

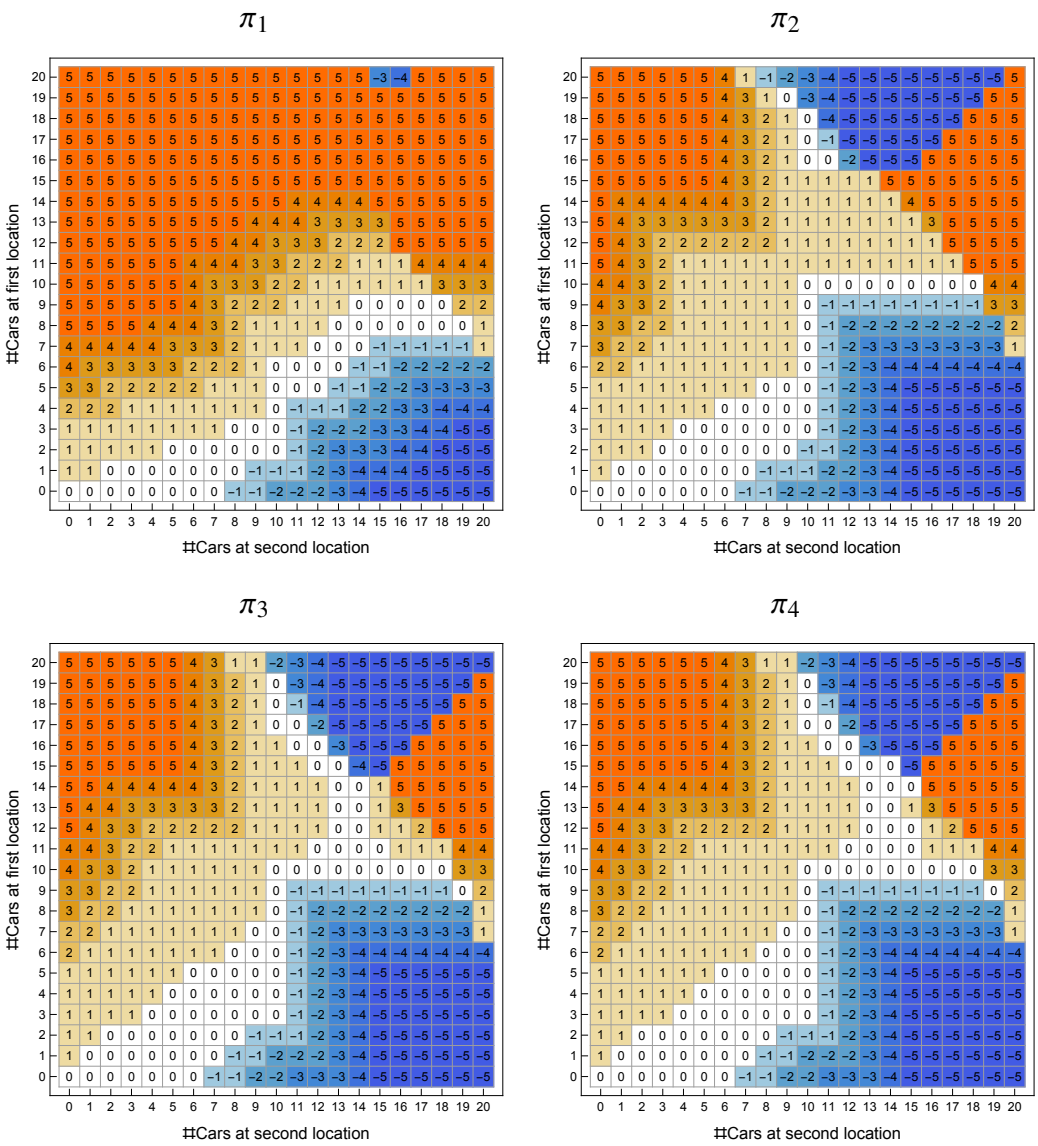

Figure 4: *Exercise 4.5*: The sequence of policies found by policy iteration on the modified Jack's car rental problem. The policy iteration algorithm was initialized by the policy that never moves any cars π_0, depicted in Fig. 1.

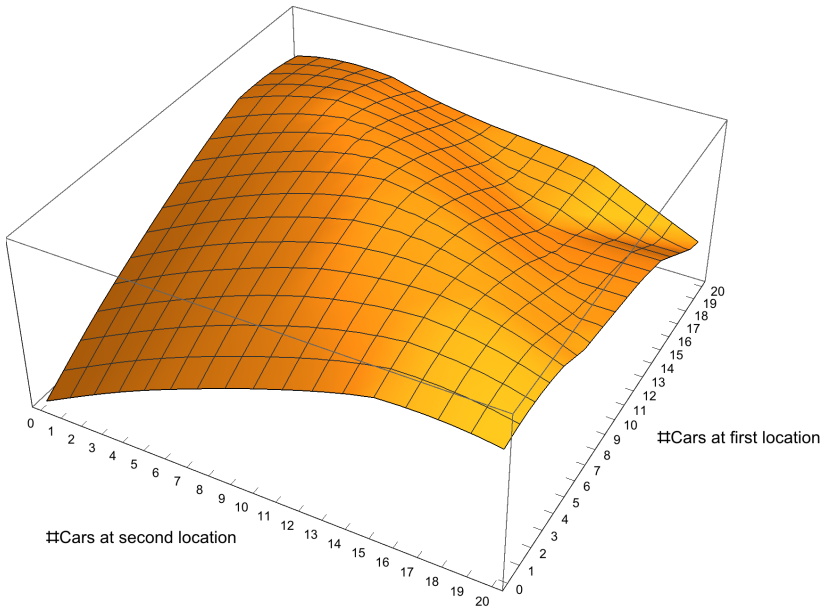

Figure 5: *Exercise 4.5*: Final state-value function for the Exercise 4.5.