

# Jack's Car Rental

AAIS in PKU 陈伟杰 1901111420

November 13, 2019

## 1 Problem Setting

Jack 有两个租车场 A,B, 各自能容纳 20 辆车, 每天出租和收回的车满足 Poisson 分布, 分别的期望值  $\lambda$  如下

$$\lambda_A(\text{rental}, \text{return}) = (3, 3) \quad \lambda_B(\text{rental}, \text{return}) = (4, 2) \quad (1)$$

A, B 两个停车场之间可以调度最多 5 辆车, 调度一辆需要支付 2¥, 而每出租一辆车可以收获 10¥。此外, 规定刚收回的车需要第二天才能出租, 调度的车在每天的开始送到并停进车场。

定义每天结束时 A, B 的车辆数为状态  $s = (n_A, n_B)$ , 调度车辆为行动  $a$ , 对应的值函数为  $V(s)$ , 那么有如下状态转移规则  $(s, a, s', r)$

$$s \xRightarrow[a]{\quad} s_0 \xRightarrow[\text{return}]{\text{rental}} s_1 \xRightarrow[\text{return}]{\quad} s' \quad (2)$$

## 2 Experiment Result

采用策略迭代算法, 评估和改进交替进行, 当策略不变时终止。其中策略评估的值函数  $V(s)$  和策略改进的策略函数  $\pi(s)$  按如下规则迭代

$$\begin{aligned} V(s) &\leftarrow \sum_{s', r} p(s', r | s, \pi(s)) [r + \gamma V(s')] \\ \pi(s) &\leftarrow \underset{a}{\operatorname{argmax}} \sum_{s', r} p(s', r | s, a) [r + \gamma V(s')] \end{aligned} \quad (3)$$

实验中设置衰减系数  $\gamma = 0.9$ , 迭代得到值函数  $V(s)$ 。

注意到, 由于概率转移矩阵可以显式写出, 因此这个问题有两种实现方法, 两者用时相差悬殊。

- 编写模拟器, 运行模拟器做迭代
- 直接将概率转移矩阵存下, 向量化更新值函数

第一种方法运行慢, 但是具有扩展性, 可以扩展到非显式概率转移的 Markov 问题; 第二种方法运行快, 但仅局限于具有显式概率转移的离散有限 Markov 问题。

本次作业使用第二种方法, 绘制图像如下

Episode:5

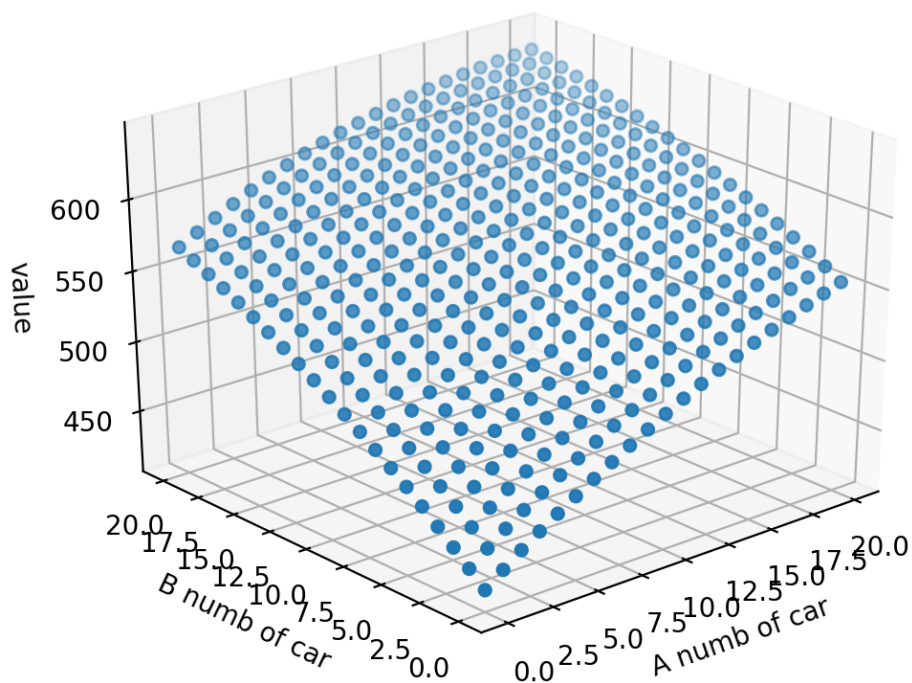


Figure 1: 值函数的状态分布图

### 3 README

代码包括 CarRental.py 和 test.py 两个程序，需要 python3 环境。直接运行 test.py 可以得到上述结果，在下列配置下用时 1min。

- 电脑：MacBook Pro(15-inch, 2019)
- 处理器：2.6 GHz 6-Core Intel Core i7
- 内存：16 GB 2400 MHz DDR4
- 操作系统：macOS Catalina version 10.15.1
- 语言：python3.6