TP Introduction to Dynamic Programming

Lily Amsellem

Wednesday, 22nd May 2019

1 Formulating the optimization problem

Question 1. To solve this problem, the following conventions are adopted:

- x represents the stock, d the demand, o the orders
- It is assumed there is no demand on day 8 ($d_8 = 0$)
- The vendor wants to maximize his profit, which is the problem to be studied in what follows

At date t, given the current stock x_t, the stock on the next day will depend on the quantity v_t sold on day t, the orders o_t made that day, and the demand d_{t+1} between t and $t+1$. We have that $v_t = min(d_{t+1}, s_t + o_t)$ since the vendor cannot sell more than his own stock+orders. This yields the following dynamic equation:

$$\forall t \in \{0, ..., T-1\}, x_{t+1} = x_t + o_t - v_t = x_t + o_t - min(d_{t+1}, x_t + o_t) = f_t(x_t, o_t, d_{t+1})$$

The objective function represents the vendor's profit, considering that he buys each product at price 1$ and sells it at 2$:

$$L_t(x_t, o_t, d_{t+1}) = 2 * \min(d_{t+1}, x_t + o_t) - o_t$$

For this problem, the constraints are the following:

- The stock cannot exceed 50 :

  $\forall t \in \{1, ..., T\}, 0 \leq x_t \leq 50$
- Up to 10 new products can be ordered each day, and current stock and orders combined musn't exceed 50 :

  $\forall t \in \{1, ..., T\}, o_t \leq min(10, 50 - x_t)$

The optimization problem to be solved is therefore the following:

$$
\begin{array}{ll}
max & \mathbb{E}[\sum_{t=0}^{T-1} L_t(x_t, o_t, d_{t+1}) + K(x_T)] \\
\text{s.t.} & \forall t \in [\![0, T-1]\!], \quad x_{t+1} = f_t(x_t, o_t, d_{t+1}), \quad x_0 = x0 \\
& \forall t \in [\![1, T]\!], \quad o_t \in U_t(s_t) \\
& \forall t \in [\![1, T]\!], \quad \sigma(o_t) \subset \sigma(d_0, ..., d_t)
\end{array}
$$

Question 2. The dynamic programming equation is, therefore:

$V_T(x) = K(x)$ and $\forall t \in [\![0, T-1]\!], V_t(x) = \max \mathbb{E}[L_t(x_t, o_t, d_{t+1}) + V_{t+1} \circ f_t(x_t, o_t, d_{t+1})]$

Question 3. An upper bound to this problem can correspond to the ideal case where the vendor exactly satisfies the demand every day and sells all of his stock (he ordered the exact right amount). In this case, the upper bound is given by:

$$\sum_{t=1}^{T} D_t = \sum_{t=1}^{T} n_t = 137\$$$

2 Solving the problem

Question 4. Following the dynamic programming algorithm, we can plot the value of the optimisation problem $V_0(x, t = 0)$ as a function of $x0$.

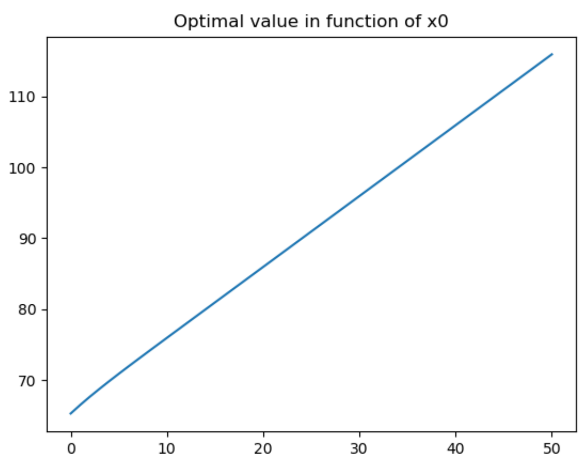


Figure 1: Plot of optimal value in function of initial stock $x0$

Question 5. In the previous questions, we considered that the initial stock was given, and that the vendor hadn't payed for it. Actually, this initial stock has a cost, which must be substracted to the optimal value that we found previously. According to the price of the initial stock, the quantities that should be bought to maximize the profit are shown below. The corresponding profits are (approximately): 66$ if each unit of the initial stock is bought for 1$, 65$ for a buying price of 1.25$ and 78$ for a buying price of 0.75$.

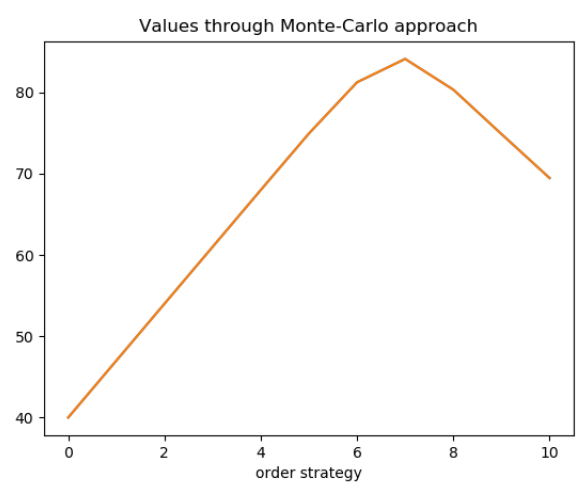
```
Stock to be bought at

1$ :  28

0.75$ :  50

1.25$ :  0
```

Figure 2: Initial stocks to be bought

3 Simulating the problem

Question 6. The optimal value can also be estimated through Monte-Carlo computations. With a strategy of order $o = 5$, the profit is of approximately 75$.

According to the order strategy of the vendor, we can plot the optimal profit, yielding the following results:



Ordering evry day $min(7, 50 - x_t)$ seems like the best strategy to maximize the vendor's profit.

Question 7. We can also check that Monte-Carlo simulations give us the same results than the dynamic programming approach. Computing the results for $x0 = 20$, we see that they are close.

```
Expected value of the optimal policy :  85.95156105987711
Value computed through Monte-Carlo :  84.22199999999945
```

Figure 4: Comparison of results through Dynamic Programming and Monte-Carlo

4 Going Further

Question 8. In the previous questions, we had considered that the final stock was just "lost", without making the vendor lose or gain any profit. Here, we suppose that the vendor can sell each unity of his final stock at 1\$. Therefore, instead of $K(x_T) = 0$, our new final value function is $K(x_T) = x_T$. Making this change, we get the following results:

Figure 5: Comparison of profit with (orange) and without (blue) selling the final stock