6.7950 Fall 2022: - Recitation 1 Handout

1 Finite-Horizon Inventory Control

We consider a retailer who sells a single item (e.g., laser pointers for forgetful instructors) and makes periodic decisions on the item's inventory strategies over T periods. At the start of each period $t \in [1, \cdots, T]$, the inventory level is s_t and the retailer makes the decision to order $a_t \geq 0$ new items, which arrive immediately. At the end of each period, D_t orders are placed and realized to the extent that there is stock available considering both current and recently ordered items. That is, if $D_t \leq s_t + a_t$, then all orders are satisfied and the inventory for the next period becomes $s_{t+1} = s_t + a_t - D_t$, while incurring a holding cost of $h_t > 0$ per remaining item. If instead there aren't enough items to satisfy all orders, then the retailer backlogs all unsatisfied demand while incurring a backlogging cost $b_t > 0$ for each of the $D_t - s_t - a_t$ unfulfilled orders.

We assume the distribution of D_t is independent across time (but not necessarily identical). Further, for all t, $\mathbb{E}[|D_t|] < +\infty$. At the end of the T periods, whatever stock remains at s_{T+1} is of no importance. For notational convenience, we define the operator $x^+ := \max\{x, 0\}$ to clamp a value x to the nearest non-negative number.

The cost C in a single period t is thus

$$C_t(s_t, a_t, D_t) := h_t(s_t + a_t - D_t)^+ + b_t(D_t - a_t - s_t)^+$$

The dynamic programming recursion to compute the optimal total cost V^* can then be transcribed as

$$\begin{cases} V^*_{T+1}(s_{T+1}) := 0, \\ V^*_t(s_t) := \min\limits_{a_t \geq 0} \mathbb{E}_{D_t}\left[C_t(s_t, a_t, D_t) + V^*_{t+1}\left((s_t + a_t - D_t)^+\right)\right], \forall\, t = T, \ldots, 1. \end{cases}$$

So, the optimal cost-to-go function $V^*_t(s_t)$ is the optimal expected cost over the periods from t to T, given that the starting inventory level in period t is s_t.

1. Prove that $V^*_t(\cdot)$ is convex. Hint: you can use induction,

2. Prove the optimal strategy in this problem is a "base-stock" policy, i.e., there exists real numbers $\overline{s}_t \geq 0$, such that the optimal action a^*_t is to maintain a stock of $\overline{s}_t$, that is

$$a^*_t(s_t) = \begin{cases} \overline{s}_t - s_t, & s_t \leq \overline{s}_t, \\ 0, & s_t > \overline{s}_t, \end{cases}$$

   Hint: use the result from the previous question.

Solution:

1. Let's first name the quantities being minimized for convenience

$$V_t^*(s_t) = \min_{a_t \geq 0} V_t^{s,a}(s_t, a_t) = \min_{a_t \geq 0} \mathbb{E}_{D_t}\left[V_t^{\text{diff}}(s_t + a_t - D_t)\right]$$

   If $V_t^{s,a}$ is convex and well defined, then so is V^* (see HW0 Problem 8.c). Similarly, since $\mathbb{E}[|D_t|] < +\infty$ and the expectation computed over D_t above is for a function of the form $f(s_t + a_t - D_t)$, then $V_t^{s,a}$ will also be convex if V_t^{diff} is convex (see HW0 P. 8.b and note that convexity is preserved in general for $g(x) = \mathbb{E}_y f(x, y)$, so that structure wasn't strictly necessary). If it could be proven that V_{t+1}^* is convex, then V_t^{diff} would be a sum of convex functions and thus we could conclude that it is convex (see HW0 P. 8.a) and by the proposed chain of consequences, V_t^* would be convex. In other words, these results serve as the induction step for proving the convexity of V_t^*. All that remains to complete the proof is to prove the base case, which is trivial since V_{T+1}^* is the null function, which is convex.

2. As convexity was proven, we first note that $V_t^*(s) \to \infty$ as $\|s\| \to \infty$, so the function has a finite minimum. Let's say this minimum is achieved at $\hat{s}_t$ with action $\hat{a}_t$. Then, by the structure of the cost function, it will also be achieved by any state-action pair (s_t, a_t) where $s_t + a_t = \hat{s}_t + \hat{a}_t = \bar{s}_t$, so long as the constraint $a_t \geq 0$ is respected (i.e., when $s_t \leq \bar{s}_t$). So whenever $s_t \leq \bar{s}_t$, an optimal action is: $a_t = \bar{s}_t - s_t$.

   For $s_t > \bar{s}_t$, the global minimum of $V_t^{s,a}$ would be achieved if a_t could be negative. Since it cannot, all valid choices of a_t would be larger than the minimizer of a convex function and thus would be in a domain where it is non-decreasing. Therefore, as $V_t^{s,a}$ never decreases for $a_t \geq 0$ the minimum constrained value would be achieved when $a_t = 0$. This procedure can be visualized for a hypothetical value function in Figure 1.

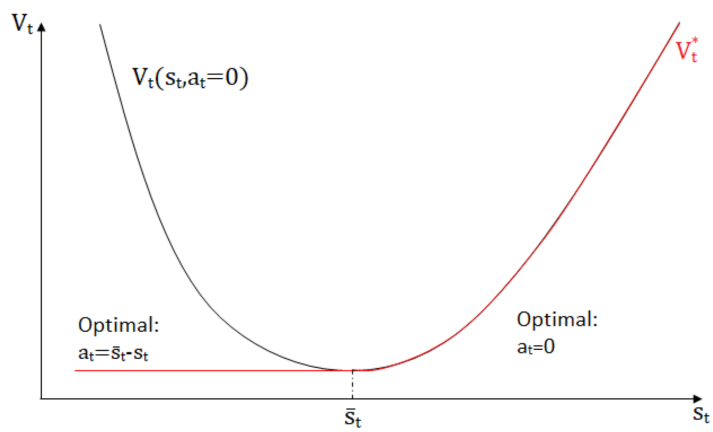

Figure 1: Visualization of the value function under different choices of $a_t(s_t)$, while using an optimal policy afterwards: $a_t = 0$ (black); $a_t = a_t^*$, the base-stock policy, which is optimal for the problem (red)