Course Notes On Dynamic Optimization (Fall 2023)
Lecture 1B: Optimal Stopping

Instructor: Daniel Russo

Email: djr2174@gsb.columbia.edu

Graduate Instructor: David Cheikhi

Email: d.cheikhi@columbia.edu

September, 2023

These notes are based of scribed notes from a previous edition of the class. I have done some follow up light editing, but there may be typos or errors.

The optimal stopping problem is one of the core problems in the decision sciences. In all variants of the problem a decision-maker faces uncertainty about the future — whether a financial asset will in value, whether a parking spot will be available– and faces a tension between stopping now or taking a risk by continuing (to hold the asset, or drive closer to a destination without parking.)

We look at an example in which a decision-maker (DM) has an asset and a fixed selling horizon. In each period they receive an offer for the asset and may choose to sell or wait to collect more offers. We consider two variants. In one, the DM must either accept an offer or permanently forego it. In another, the DM can collect many offers, then return to accept the highest one made so far.

Learning objectives.

1. Understand how to express a decision-problem as a special case of the general problem described in class.
2. Get a feel for how problem-specific structures can lead to special solutions. In both versions of our problem the optimal policy is a *threshold policy*; in the case where offers are retained the policy is also *myopic*.

1 Asset selling with irrevocable decisions

Consider the following problem:

- A person has an asset they must sell within a liquidation deadline of N periods. In each period, they receive single offer, modeled as a non-negative random variable drawn independently from a distribution. We assume the offers are also identically distributed, though many results do not depend on this choice.

- They may accept the offer and invest the money at fixed rate of interest $r > 0$, or reject it and wait to receive another offer.
- Their goal is to maximize the expected final revenue from the sale and the accrued interest.

Casting this as a special case of our general problem. We use x_k's to denote the states, $\forall k = 0, \ldots, N-1$. The state should have the information of whether the asset has sold as well as the current offer. We incorporate these information in the following way: $\forall k = 0, \cdots, N-1$

$$x_{k+1} = \begin{cases} \text{sold} & \text{if } U_k = \text{Accept or } x_k = \text{sold} \\ w_k & \text{o.w.} \end{cases}$$

Set $x_0 = 0$ as a dummy variable. Hence the state space $S_k \subset \mathbb{R} \cup \{\text{sold}\}, \forall k = 0, \ldots, N-1$. The action space, $\forall k = 0, \cdots, N-1$,

$$U_k(x_k) = \begin{cases} \text{Reject} & \text{if } x_k = \text{sold} \\ \{\text{Accept, Reject}\} & \text{o.w. ,} \end{cases}.$$

Accounting for compounding interest, the revenue for each period is defined as, $\forall k = 1, \ldots, N-1$,

$$g_k(x_k, u_k, w_k) = \begin{cases} 0 & \text{if } U_k \neq \text{Accept} \\ (1+r)^{N-k} x_k & \text{if } U_k = \text{Accept} \end{cases}$$

with terminal revenue

$$g_N(x_N) = \begin{cases} 0 & \text{if } x_N = \text{sold} \\ x_N & \text{o.w.} \end{cases}$$

encoding that an unsold asset is automatically sold in period N, regardless of what offer is made.

Remark. There are at least two tricky aspects of this formulation. One is that the stopping problem seems to have an indefinite horizon — one which could last N periods but might terminate earlier if the asset is sold. We cast it as a special case of our formulation by introducing a special absorbing state (sold). This modeling trick is common in DP.

A second 'trick' has to with indexing. We index offers starting from 0, as $w_0, w_1, \ldots, w_{N-1}$. In period k, the state x_k reflects the offer w_{k-1}. To fit our formalism, the disturbance w_k must influence *the next state* x_{k+1}.

DP Algorithm: $J_N^*(x) = g_N(x)$ for $x \neq$ sold. For $k = N-1, N-2, \ldots, 0$,

$$J_k^*(x) = \begin{cases} \max\{(1+r)^{N-k} x, \mathbb{E}[J_{k+1}^*(w_k)]\} & \text{if } x \neq \text{sold} \\ 0 & \text{if } x = \text{sold} \end{cases}$$

It is cleaner to introduce the normalized functions $V_k(x) \equiv J_k(x)/(1+r)^{N-k}$. Rather than up-weight earnings in early periods, these "discount" earnings far away into the future. Under these normalized value functions, the recursion becomes

$$V_k^*(x) = \begin{cases} \max\{x, \mathbb{E}[V_{k+1}^*(w_k)]/(1+r)\} & \text{if } x \neq \text{sold} \\ 0 & \text{if } x = \text{sold.} \end{cases}$$

Optimal Policy: The optimal policy is a threshold policy. For $x_k \neq$ sold,

$$\mu_k^*(x_k) = \begin{cases} \text{Accept} & \text{if } x_k \geqslant \alpha_k \\ \text{Reject} & \text{if } x_k \leqslant \alpha_k \end{cases},$$

where $\alpha_k = \mathbb{E}[V_k^*(w_k)]/(1+r)$.

By following the logic of the DP algorithm, it is possible to show, α_k's obey their own recursion. Since we must accept the last offer, $\alpha_N = -\infty$. For $k = N-1, \cdots, 0$, $\alpha_k = \frac{1}{1+r}\mathbb{E}[\max\{w_k, \alpha_{k+1}\}]$.

Preview of infinite horizon DP. When N is large the DM should never wait until the end of the selling horizon, since they would forego the benefit of exponentially compounding interest. If the selling deadline is far away, it is cleaner to take $N \to \infty$. In this case, one cas show that optimal threshold does not depend on the current period (since the selling deadline is always "infinitely far way"); that optimal threshold is the the solution to the fixed-point equation $\alpha = \frac{1}{1+r}\mathbb{E}[\max\{w_1, \alpha\}]$.

2 Asset selling with offers retained

We modify the previous setting such that, upon stopping, the DM can accept the highest offer made so far. The result in this section relies strongly on the offers being i.i.d.

To accommodate this setting, we define the state such that $\forall k = 0, \cdots, N-1$

$$x_{k+1} = \begin{cases} \text{sold} & \text{if } U_k = \text{Accept or } x_k = \text{sold} \\ \max\{x_k, w_k\} & \text{o.w.} \end{cases}$$

Action space and functions g_k's stay the same.

Below we show the optimal policy is **myopic**: the DM should compare 1) stopping immediately and 2) continue for one more period and accepting the next offer. This one-step-lookahead heuristic is optimal even in problems with a long selling horizon.

Proposition 1. *An optimal policy for asset selling with offers retained is $\pi^* = (\mu^*, \mu^*, \cdots, \mu^*)$, where for $x \neq$ sold,*

$$\mu^*(x) = \begin{cases} \textit{Accept} & \textit{if } x \geqslant \frac{1}{1+r}\mathbb{E}[\max\{x, w_k\}] \\ \textit{Reject} & \textit{o.w.} \end{cases}.$$

Note this shows that lookahead one step is optimal when the offers can be retained.

proof sketch. Define the set of optimal stopping region,

$$S_k^* = \{x | x \geqslant \mathbb{E}[V_{k+1}^*(\max\{x, w_k\})]/(1+r)\}$$

Since $V_N^*(x) = x$ for $x \neq \text{sold}$, the optimal stopping region is

$$S_{N-1}^* = \{x | x \geqslant \underbrace{\mathbb{E}_w[\max\{x, w\}]/(1+r)}_{:=\bar{\alpha}}$$

We call this the one-step lookahead stopping region.

Using the DP algorithm, it is easy to show that $V_1^* \succeq V_2^* \succeq V_3^*$ and so on. That is, the decision maker weakly prefers be in period k with offer $x_k = x$ than to be in some period $k' > k$ with the same offer ($x_{k'} = x$), because k' is associated with a shorter remaining selling horizon and hence less ability to collect future offers. This implies that

$$S_0^* \subset S_2^* \cdots \subset S_{N-1}^*.$$

We proceed by backward induction. Suppose that $S_{k+1}^* = S_{N-1}^* = \{x | x \geqslant \bar{\alpha}\}$ (with the base of induction being $k+1 = N-1$). First, suppose $x_k \leqslant \bar{\alpha}$. Since $S_k^* \subset S_{N-1}^*$, we know $\mu_k^*(x_k) = \text{Reject}$.

Now suppose $x_k \geqslant \bar{\alpha}$. The decision-maker must choose whether to stop and earn x_k or continue and earn $(1-r)^{-1}\mathbb{E}[V_{n+1}(x_{n+1}) \mid x_n]$. The key to the argument is to observe that $x_{n+1} \geqslant x_n \geqslant \bar{\alpha}$ almost surely, and we so we know $\mu_{n+1}^*(x_{n+1}) = \text{Accept}$; Since we know the decision-maker will accept in the next period, regardless of the next offer, we know $V_{n+1}^*(x_{n+1}) = x_{n+1} = V_{N-1}^*(x_n)$; the DM can plan as if they will stop in the next period. □