7 Optimal stopping

We show how optimal stopping problems for Markov chains can be treated as dynamic optimization problems.

Let $(X_n)_{n\geqslant 0}$ be a Markov chain on S, with transition matrix P. Suppose given two bounded functions

$$c : S \to \mathbb{R}, \quad f : S \to \mathbb{R},$$

respectively the *continuation cost* and the *stopping cost*. A random variable T, with values in $\mathbb{Z}^+ \cup \{\infty\}$, is a *stopping time* if, for all $n \in \mathbb{Z}^+$, the event $\{T = n\}$ depends only on $X_0, \dots, X_n$. Define the *expected total cost function* V^T by

$$V^T(x) = \mathbb{E}_x\left(\sum_{k=0}^{T-1} c(X_k) + f(X_T)1_{\{T<\infty\}}\right), \quad x \in S,$$

and define for $n \in \mathbb{Z}^+$ and $x \in S$,

$$V_n(x) = \inf_{T\leqslant n} V^T(x), \quad V_*(x) = \inf_{T<\infty} V^T(x), \quad V(x) = \inf_{T} V^T(x),$$

where the infima are taken over all stopping times T, first with the restriction $T \leqslant n$, then with $T < \infty$, and finally unrestricted. Where unbounded stopping times are involved, we assume that c and f are non-negative, so the sums and expectations are well defined. It is clear that $V_n(x) \geqslant V_{n+1}(x) \geqslant V_*(x) \geqslant V(x)$ for all n and x, as the infima are taken over progressively larger sets. The calculation of these functions and the determination, where possible, of minimizing stopping times are known as *optimal stopping problems*[20].

We translate these problems now into dynamic optimization problems, with state-space $S\cup\{\partial\}$ and action space $\{0,1\}$. Action 0 will correspond to continuing, action 1 to stopping. On stopping, we go to ∂ and stay there. Define, for $x \in S$,

$$P(x,0)_y = p_{xy}, \quad P(x,1)_\partial = \delta_{y\partial},$$

and

$$c(x,a) = \begin{cases} c(x), & a = 0, \\ f(x), & a = 1. \end{cases}$$

Given a stopping time T, there exists for each $n \geqslant 0$ a set $B_n \subseteq S^{n+1}$ such that $\{T = n\} = \{(X_0, \dots, X_n) \in B_n\}$. Define a control u by

$$u_n(x_0, \dots, x_n) = \begin{cases} 1, & \text{if } (x_0, \dots, x_n) \in B_n, \\ 0, & \text{otherwise.} \end{cases}$$

Note that we obtain all controls for starting time 0 in this way and that the controlled process is given by

$$\tilde{X}_n = \begin{cases} X_n, & n \leqslant T, \\ \partial, & n \geqslant T+1. \end{cases}$$

[20]We limit our discussion to the time-homogeneous case. If there is a time dependence in the transition matrix or in the costs, a reduction to the time-homogeneous case can be achieved as in footnote 3, specifically, by considering the process $\tilde{X}_n = (k+n, X_{n+k})$.

Hence, V_n is the infimal cost function for the n-horizon problem, with final cost f, so satisfies $V_0(x) = f(x)$ and, for all $n \geqslant 0$,

$$V_{n+1}(x) = \min\{f(x), (c + PV_n)(x)\}, \quad x \in S.$$

Moreover, V is the infimal cost function for the infinite-horizon problem, so, if c and f are non-negative, then V is the minimal non-negative solution to

$$V(x) = \min\{f(x), (c + PV)(x)\}, \quad x \in S.$$

The V_* problem corresponds to a type of restriction on controls which we have not seen before. However the argument of Proposition 2.1 can be adapted to show that V_* also satisfies the optimality equation

$$V_*(x) = \min\{f(x), (c + PV_*)(x)\}, \quad x \in S.$$

Example. Consider a simple symmetric random walk on the integers with continuation cost $c(x) = 0$ and stopping cost $f(x) = 1 + e^{-x}$. Since f is convex, specifically since $f(x) \leqslant \frac{1}{2}f(x+1) + \frac{1}{2}f(x-1)$ for all x, a simple inductive argument[21] using the finite-horizon optimality equations shows that $V_n = f$ for all n. Since $(X_n)_{n\geqslant 0}$ is recurrent, the stopping time $T_n = \inf\{n \geqslant 0 : X_n = N\}$ is finite for all N, for every starting point x. So $V_*(x) \leqslant V^{T_n}(x) = 1 + e^{-N}$. Obviously, $V_*(x) \geqslant 1$, so $V_*(x) = 1$ for all x. Finally, $V = V^\infty = 0$. We note that $\inf_n V_n(x) > V_*(x) > V(x)$ for all x.

Proposition 7.1 (One step look ahead rule). *Suppose that $(X_n)_{n\geqslant 0}$ cannot escape from the set*

$$S_0 = \{x \in S : f(x) \leqslant (c + Pf)(x)\}.$$

Then, for all $n \geqslant 0$, the following stopping time is optimal for the n-horizon problem

$$T_n = \inf\{k \geqslant 0 : X_k \in S_0\} \wedge n.$$

Proof. The case $n = 0$ is trivially true. Suppose inductively that the claim holds for n. Then $V_n = f$ on S_0, so $PV_n = Pf$ on S_0 as we cannot escape. So, for $x \in S_0$,

$$V_{n+1}(x) = \min\{f(x), (c + PV_n)(x)\} = f(x)$$

and it is optimal to stop immediately. But, for $x \notin S_0$, it is better to wait, if we can. Hence the claim holds for $n+1$ and the induction proceeds. □

[21] An alternative analysis of this example may be based on the *optional stopping theorem*, which is a fundamental result of martingale theory. This is introduced in the course Stochastic Financial Models and in the Part III course Advanced Probability. The random walk is a martingale, so, since f is convex, $(f(X_n))_{n\geqslant 0}$ is a submartingale. By optional stopping, $\mathbb{E}_x(f(X_T)) \geqslant \mathbb{E}_x(f(X_0)) = f(x)$ for all bounded stopping times T, so $V_n(x) = f(x)$ for all x. The fact that the conclusion of the optional stopping theorem does not extend to T_N is a well known sort of counterexample in martingale theory.

Example (Optimal parking). Suppose that you intend to park on the Backs, and wish to minimize the expected distance you will have to walk to Garrett Hostel Lane, and that a proportion p of the parking spaces are free. Assume that each parking space is free or occupied independently, that a queue of cars behind you take up immediately any space you pass by, and that no new spaces are vacated. Where should you park?

If you reach Garrett Hostel Lane without parking, then you should park in the next available space. This lies at a random distance (in spaces) D, with $\mathbb{P}(D = n) = (1 - p)p^n$, for $n \geqslant 0$, so the expected distance to walk is $\mathbb{E}(D) = q/p$, where $q = 1 - p$. Here we have made the simplifying assumptions that Queen's Road is infinitely long and that there are no gaps between the spaces.

Write V_n for the minimal expected distance starting from n spaces before Garrett Hostel Lane. Then $V_0 = q/p$ and, for $n \geqslant 1$, $V_n = qV_{n-1} + p \min\{n, V_{n-1}\}$. Set $n^* = \inf\{n \geqslant 0 : V_n < n\}$. For $n \leqslant n^*$, we have $V_n = qV_{n-1} + pn$, so $V_n = n + (2q^n - 1)q/p$. Hence $n^* = \inf\{n \geqslant 0 : 2q^n < 1\}$. For $n \geqslant n^*$, we have $V_n = V_{n^*}$. The optimal time to stop is thus the first free space no more than n^* spaces before the Lane. We leave as an exercise the to express this argument in terms of the general framework described above.