Optimal Stopping and Applications

Alex Cox

May 7, 2019

Abstract

These notes are intended to accompany a Graduate course on Optimal stopping, and in places are a bit brief. They follow the book 'Optimal Stopping and Free-boundary Problems' by Peskir and Shiryaev, and more details can generally be found there.

1 Introduction

1.1 Motivating Examples

Given a stochastic process X_t, an optimal stopping problem is to compute the following:

$$\sup_{\tau} \mathbb{E}F(X_\tau), \tag{1}$$

where the supremum is taken over some set of stopping times. As well as computing the supremum, we will also usually be interested in finding a description of the stopping time. For motivation, we consider three applications where optimal stopping problems arise.

(i) **Stochastic analysis.** Let B_t be a Brownian motion. Then classical results tell us that for any (fixed) $T \geq 0$,

$$\mathbb{E}\left[\sup_{0\leq s\leq T} |B_s|\right] = \sqrt{\frac{\pi T}{2}}$$

What can we say about the left-hand side if we replace the constant T with a stopping time τ? To solve this problem, it turns out that we can consider the (class of) optimal stopping problems:

$$V(\lambda) = \sup_{\tau} \mathbb{E}\left[\sup_{0\leq s\leq \tau} |B_s| - \lambda\tau\right]. \tag{2}$$

Where we take the supremum over all stopping times which have $\mathbb{E}\tau < \infty$, and $\lambda > 0$. (Note that in order to fit into the exact form described in (1), we need to take e.g. $X_t = (B_t, \sup_{s\le t}|B_s|, t))$. Now suppose we can solve this problem for all $\lambda > 0$, then we can use a Lagrangian-style approach to get back to the original question: for any suitable stopping time τ we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le s\le\tau}|B_s|\right] &\le V(\lambda) + \lambda\mathbb{E}\tau \\ &\le \inf_{\lambda>0}\left(V(\lambda) + \lambda\mathbb{E}\tau\right)\end{aligned}$$

Now, if we suppose that the infimum here is attained, and the supremum in (2) is also attained at the optimal λ, then we see that the right-hand side is a function of $\mathbb{E}\tau$, and further, we can attain equality for this bound, so that it is the smallest possible bound.

(ii) **Sequential analysis.** Suppose we observe (in continuous time) a statistical process X_t — perhaps X_t is a measure of the health of patients in a clinical trial. We wish to know whether a particular drug has a positive impact on the patients. If we suppose that the process X_t has:

$$X_t = \mu t + B_t$$

where μ is the unknown effect of the treatment. Then to make inference, we might test the hypotheses:

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu = \mu_1.$$

Our goal might be to minimise the average time we take to make a decision subject to some constraints on the probabilities of making wrong decisions:

$$\begin{aligned}\mathbb{P}(\text{accept } H_0|H_1 \text{ true}) &\le \alpha \\ \mathbb{P}(\text{accept } H_1|H_0 \text{ true}) &\le \beta.\end{aligned}$$

Again, using Lagrangian techniques, we are able to rewrite this as an optimal stopping problem, which we can solve to find the optimal stopping time (together with a 'rule' to decide whether we accept or reject when we stop).

(iii) **Mathematical Finance.** Let S_t be the share price of a company. A common type of option is the American Put option. This contract gives the holder of the option the right, but not the obligation, to sell the stock for a fixed price K to the writer of the contract at any time before some fixed horizon T. In particular, suppose the holder exercises at a stopping time τ, then the discounted[1] value of the payoff is: $\mathrm{e}^{-r\tau}(K - S_\tau)_+$. This is the difference between the amount she receives, and the actual value of the asset, and we take the positive part, since she will only choose to *exercise* the option, and sell the share, if the price she would receive is larger than the price she would get from selling on the market. If the holder of the contract chooses the stopping time τ, the fundamental theorem of asset pricing[2] says that the payoff has initial value

$$\mathbb{E}^{\mathbb{Q}}\mathrm{e}^{-r\tau}(K - S_\tau)_+,$$

[1] We convert time-t money into time-0 money by multiplying the time-t amount by the discount factor e^{-rt}, where r is the interest rate.

[2] If you don't know what this is, ignore the $\mathbb{Q}$ that crops up, and just think of this as pricing by taking an average of the final payoff.

where $\mathbb{Q}$ is the risk-neutral measure since the seller would not want to sell this too cheaply (and in fact, there is a better argument based on hedging that we will see later), and she doesn't know what strategy the buyer will use, so she should take the supremum over all stopping times $\tau \leq T$:

$$\text{Price} = \sup_{\tau \leq T} \mathbb{E}^{\mathbb{Q}} e^{-r\tau}(K - S_\tau)_+.$$

So pricing the American put option is equivalent to solving an optimal stopping problem.

1.2 Simple Example

Once a problem of interest has been set up as an optimal stopping problem, we then need to consider the method of solution. The exact techniques vary depending on the problem, however there are two important approaches — the martingale approach, and the Markov approach, and there are a number of common features that any optimal stopping problem is likely to exhibit.

We will now consider a reasonable simple example where we will be able to explicitly observe many of the common features.

Let X_n be a simple symmetric random walk on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{Z}_+}, (\mathbb{P}_x)_{x\in\mathbb{Z}})$, where $\mathcal{F}_n$ is the natural filtration of X_n, and $\mathbb{P}_x(X_0 = x) = 1$. Let $f(\cdot)$ be a function on $\mathbb{Z}$ with $f(n) = 0$ if $n \geq N$ or $n \leq -N$, for some N. Then the optimal stopping problem we will consider is to find:

$$\sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_0 f(X_\tau)$$

where $\mathcal{M}^{H_{\pm N}}$ is the set of stopping times which are earlier than the hitting time $H_{\pm N} = \inf\{n \geq 0 : |X_n| \geq N\}$. Note that some kind of constraint along these lines is necessary to make the problem non-trivial, since otherwise we can use the recurrence of the process to wait until the process hits the point at which the function $f(\cdot)$ is maximised — similar restrictions are common throughout optimal stopping, although often arise more naturally from the setup.

For two stopping times τ_1, τ_2, introduce the class of stopping times (generalising the above):

$$\mathcal{M}_{\tau_1}^{\tau_2} = \{ \text{ stopping times } \tau : \tau_1 \leq \tau \leq \tau_2\} \tag{3}$$

where we will often omit τ_1 if this is intended to be $\tau_1 \equiv 0$. Then we may use the Markov property of X_n to deduce[3]that

$$\sup_{\tau \in \mathcal{M}_n^{H_{\pm N}}} \mathbb{E}_0 \left[f(X_\tau) | \mathcal{F}_n \right]$$

[3]There is a technical issue here: namely that the set of stopping times over which we take the supremum may be uncountable, and therefore the resulting 'sup,' which is a function from $\Omega \to \mathbb{R}$, may not be measurable, and therefore may not be a random variable. We will introduce the correct notion to resolve this issue later, but for the moment, we interpret this statement informally.

must be a function of X_n only (i.e. independent of $\mathcal{F}_n$ except through X_n), so, on the set $\{n \leq H_{\pm N}\}$, we get:

$$\sup_{\tau \in \mathcal{M}_n^{H_{\pm N}}} \mathbb{E}_0 \left[f(X_\tau) | \mathcal{F}_n \right] = V^*(X_n, n)$$

for some function $V^*(\cdot, \cdot)$, and in principle, we believe that the function should only depend on the spatial, and not the time parameter, so that we introduce as well:

$$V(y) = V^*(y, 0) = \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_y f(X_\tau). \tag{4}$$

Note that we must have[4] $V(y) \leq V^*(y, n)$.

Our aim is now to find $V(0)$ (and the connected optimal strategy), but in fact it will be easiest to do this by characterising the whole function $V(\cdot)$. The function $V(\cdot)$ will be called the **value function**.

In general, we can characterise $V(\cdot)$ as follows:

Theorem 1.1. *The value function, $V(\cdot)$, is the smallest concave function on $\{-N, \ldots, N\}$ which is larger than $f(\cdot)$. Moreover:*

(i) $V(X_n^{H_{\pm N}})$ *is a supermartingale;*

(ii) there exists a stopping time τ^* *which attains the supremum in* (4)*;*

(iii) the stopped process $V(X_{n \wedge \tau^*}^{H_{\pm N}})$ *is a martingale;*

(iv) V *satisfies the 'Bellman equation:'*

$$V(x) = \max\{f(x), \mathbb{E}_x V(X_1)\}; \tag{5}$$

(v) $V(x) = V^*(x, n)$ *for all* $n \geq 0$.

(vi) $V(X_n)$ *is the smallest supermartingale dominating* $f(X_n)$ *(i.e.* $V(X_n) \geq f(X_n)$*).*

Proof. We first note that the smallest concave function dominating f does indeed exist: let $\mathcal{G}$ be the set of concave functions which are larger than f, and set

$$g(x) = \inf_{h \in \mathcal{G}} h(x).$$

The set $\mathcal{G}$ is non-empty, since $h(x) \equiv M$ is in $\mathcal{G}$ for M greater than $\max_{n \in \{-N, \ldots, N\}} f(n)$, so g is finite. Suppose g is not concave. Then we can find $x < y < z$ such that

$$g(y) < \frac{z-y}{z-x} g(x) + \frac{y-x}{z-x} g(z)$$

[4] Any difference between $V(y)$ and $V^*(y, n)$ must come from the fact that there is a stopping time which uses the extra randomness in $\mathcal{F}_n$ (which contains no information about the future of the process); if we have a stopping time which is optimal without the randomness, we can always construct a stopping time at $\mathcal{F}_n$ which has the same expected payoff.

and therefore, there is some $h \in \mathcal{G}$ for which:

$$\begin{aligned} h(y) &< \frac{z-y}{z-x}g(x) + \frac{y-x}{z-x}g(z) \\ &< \frac{z-y}{z-x}h(x) + \frac{y-x}{z-x}h(z), \end{aligned}$$

contradicting the concavity of h.

Now we consider the function $V(\cdot)$. Clearly, from the definition (4), we must have $V(x) \geq f(x)$ (just take $\tau \equiv n$). Moreover, $V(\cdot)$ must be concave. We introduce the following notation: let $\mathbb{P}_x$ (and $\mathbb{E}_x$) denote the probability measure with $X_0 = x$, and define the stopping time

$$H_{a,b} = \inf\{n \geq 0 : X_n = a \text{ or } X_n = b\}$$

then we get

$$\begin{aligned} V(y) &= \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_y f(X_\tau) \\ &\geq \sup_{\tau \in \mathcal{M}^{H_{\pm N}}_{H_{x,z}}} \mathbb{E}_y f(X_\tau) \\ &\geq \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_y f(X_{H_{x,z} + \tau \circ \theta_{H_{x,z}}}) \end{aligned}$$

where θ_T is the usual shift operator. So

$$\begin{aligned} V(y) &\geq \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \left[\mathbb{E}_y \left[f(X_{H_{x,z} + \tau \circ \theta_{H_{x,z}}}) | X_{H_{x,z}} = x \right] \frac{z-y}{z-x} \right. \\ &\qquad \left. + \mathbb{E}_y \left[f(X_{H_{x,z} + \tau \circ \theta_{H_{x,z}}}) | X_{H_{x,z}} = z \right] \frac{y-x}{z-x} \right] \\ &\geq \frac{z-y}{z-x} \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_x f(X_\tau) + \frac{y-x}{z-x} \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_z f(X_\tau) \\ &\geq \frac{z-y}{z-x} V(x) + \frac{y-x}{z-x} V(z) \end{aligned}$$

which gives the required concavity, and so we have $V \in \mathcal{G}$.

We note first that we have $V(-N) = V(N) = 0$, from (4), and the fact that $f(N) = f(-N) = 0$. from this, we deduce that any $g \in \mathcal{G}$ is non-negative. In fact, since for any $g \in \mathcal{G}$, and $x \in \{-N+1, \dots, N-1\}$, concavity implies $g(x) \geq \frac{1}{2}(g(x-1) + g(x+1))$, we have

$$g(X_n) \geq \mathbb{E}_{\cdot} \left[g(X_{n+1}) | \mathcal{F}_n \right]$$

and hence $g(X_n^{H_{\pm N}})$ is a supermartingale. In particular, for all $\tau \in \mathcal{M}^{H_{\pm N}}$, we have

$$\begin{aligned} g(x) &\geq \mathbb{E}_x g(X_\tau) \\ &\geq \mathbb{E}_x f(X_\tau) \end{aligned}$$

and therefore

$$\begin{aligned} g(x) &\geq \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_x f(X_\tau) \\ &\geq V(x), \end{aligned}$$

and we conclude that $V(\cdot)$ is the smallest concave function. Note that the above argument also applies to $V(\cdot)$, since $V \in \mathcal{G}$, so we conclude that $V(X_n^{H_{\pm N}})$ is also a supermartingale.

Now we need to construct the optimal stopping time. Define the sets:

$$C = \{x : V(x) > f(x)\}$$
$$D = \{x : V(x) = f(x)\}.$$

We claim that the function $V(\cdot)$ is linear between points in $D(\cdot)$: recall that the minimum of any two concave functions is also concave, and suppose that there is a point y of $C(\cdot)$ at which $V(\cdot)$ is larger than the line between $(x, V(x))$ and $(z, V(z))$, where x and z are the nearest points below and above y in D. Then there exists a point $y^* \in \{x+1, \dots, z\}$ which maximises $\frac{f(y^*)-f(x)}{y^*-x}$. Taking the minimum of V and the line between $(x, V(x))$ and $(y^*, f(y^*))$ is then a strictly smaller concave function, which remains larger than $V(\cdot)$. Hence $V(\cdot)$ is indeed linear between points of D.

Now consider the stopping time $\tau^* = \inf\{n \geq 0 : X_n \in D\}$. Clearly, if $X_0 = y \in C$, and x and z are the nearest points below and above y in D, we have

$$\mathbb{E}_y f(X_{\tau^*}) = \frac{z-y}{z-x} f(x) + \frac{y-x}{z-x} f(z) = V(y).$$

Alternatively, if $y \in D$, $\tau^* = 0, \mathbb{P}_y$-a.s., and so $\mathbb{E}_y f(X_{\tau^*}) = f(y)$.

Hence τ^* is optimal in the sense:

$$V(y) = \sup_{\tau \in \mathcal{M}^{H_{\pm N}}} \mathbb{E}_y f(X_\tau) = \mathbb{E}_y f(X_{\tau^*}).$$

Consequently, $V(y) = \mathbb{E}_y V_{\tau^*}^{H_{\pm N}}$, and the process $V_{n\wedge\tau^*}^{H_{\pm N}}$ (which we have already shown is a non-negative supermartingale) must be a (uniformly integrable) martingale.

Property *(iv)* now clearly follows — if $f(x) \geq \mathbb{E}_x V(X_1)$ then $\tau^* = 0$, otherwise $\tau^* \geq 1$.

Now consider the function $V^*(x, n)$. Suppose we have $V^*(X_n, n) > V(X_n)$ for some n, X_n, and let $\tau_n \in \mathcal{M}_n^{H_{\pm N}}$ be a stopping time for which:

$$\mathbb{E}_0[f(X_{\tau_n})|\mathcal{F}_n] > V(X_n).$$

Using the supermartingale property of $V(X_n)$, we conclude that

$$V(X_n) \geq \mathbb{E}_0[V(X_{\tau_n})|\mathcal{F}_n] \geq \mathbb{E}_0[f(X_{\tau_n})|\mathcal{F}_n],$$

which contradicts the choice of τ_n.

Finally, suppose there exists a supermartingale Y_n with $Y_n \leq V(X_n)$, but $Y_n \geq f(X_n)$. Write $\tau_n = \inf\{m \geq n : X_m \in D\}$. Then $Y_n \geq \mathbb{E}[Y_{\tau_n}|\mathcal{F}_n] \geq \mathbb{E}[f(X_{\tau_n})|\mathcal{F}_n] = V(X_n)$. Hence $V(X_n)$ is the smallest supermartingale dominating $f(\cdot)$. □

Remark 1.2. (i) The sets C and D are called the **continuation set** and **stopping set** respectively. Note the connection between C and D and the concavity of the value function:

$$V(x) - \frac{1}{2}(V(x+1) + V(x-1)) = 0 \qquad x \in C \tag{6}$$
$$V(x) - \frac{1}{2}(V(x+1) + V(x-1)) \leq 0 \qquad x \in D. \tag{7}$$

In particular, $V(X_n)$ is a martingale in C, and a supermartingale on D (strictly so, if the inequality is strict).

(ii) We have made strong use of the Markov property here: the ability to write the value function as a function of X_n has made things considerably easier; in addition, the fact that there is no time-horizon helps, otherwise, our martingale would have to be a function of both time and space, and correspondingly harder to characterise.

In general, we would like to consider situations where the process is not Markov, and this means we cannot characterise the value function quite so simply. However, part *(vi)* suggests the way out — rather than look for functions on the state space of the Markov process, we will look for the smallest supermartingales which dominates the payoff. The corresponding supermartingale will be called the **Snell envelope**.

(iii) The other question is: how might things change in continuous time and space? A simple way of considering this is to see what happens if we rescale the current problem by letting the space and time steps get small appropriately, so that in the limit, we get a Brownian motion. Thinking about (6)–(7), we might expect the statespace, $[-N, N]$, to break up into two sets, C and D, such that:

$$\lim_{n\to\infty} \frac{\left(V(x) - \frac{1}{2}(V(x+1/n) + V(x-1/n))\right)}{n^2} = V''(x) = 0 \qquad x \in C$$
$$\lim_{n\to\infty} \frac{\left(V(x) - \frac{1}{2}(V(x+1/n) + V(x-1/n))\right)}{n^2} = V''(x) \leq 0 \qquad x \in D$$

and such that $V(x) = f(x)$ in D.

Thinking a bit more, suppose our process starts at $y \in [-N, N]$ — a possible approach to finding $V(y)$ is to find the points $x \leq y \leq z$, and the function V such that:

$$\begin{aligned} V''(w) &= 0, & w &\in (x, z) \\ V(w) &= f(w), & w &\in \{x, z\} \\ V'(w) &= f'(w), & w &\in \{x, z\} \\ V''(w) &\leq 0, & w &\in [-N, N] \end{aligned}$$

then this will be sufficient to characterise the value function at y. Such a formulation is analogous to the **free-boundary problems** that we will see in later lectures (although these will normally be formulated as PDE problems, rather than differential equation problems as it is here). Conditions like those on the value and the gradient on the 'boundary' are **continuous fit** and **smooth fit** conditions.

2 Discrete time

Our approach throughout this section will be to start in the more general case, where we simply assume that there is some gains process G_n, which is the amount we receive at time n if we choose to stop, and then later to specify the situation to the Markov case, where we can say more about the value function and the optimal stopping time.

2.1 Martingale treatment

We begin by assuming simply that $(G_n)_{n\geq 0}$ is a sequence of adapted random variables on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$, and interpret G_n as the gain, or reward we receive for stopping at time n.

Recall the definition of $\mathcal{M}_{\tau_1}^{\tau_2}$ as the set of stopping times which are less than τ_2 and greater than τ_1; we shall omit the lower limit when this is 0, and if we omit the upper limit, we admit any stopping time which is almost surely finite. Our optimisation will be taken over some set of stopping times $\mathcal{N} \subseteq \mathcal{M}$, so the value function is:

$$V = \sup_{\tau\in\mathcal{N}} \mathbb{E}G_\tau.$$

To keep the problem technically simple, we assume that

$$\sup_{\tau\in\mathcal{N}} \mathbb{E}\left[\sup_{n\leq\tau} |G_n|\right] < \infty. \tag{8}$$

Note that this can be rather restrictive, but if $\mathcal{N} = \mathcal{M}^\tau$ for some relatively nice τ — like $\tau = H_{\pm N}$, or some constant — we are usually OK.

We begin by considering the case where we have some finite horizon to the problem: that is, $\mathcal{N} = \mathcal{M}^N$, for a fixed constant N. Clearly, if we arrive at this horizon, we have no choice but to stop, so that the value function and the gain function are the same at time N. From here, we can work backwards — if, conditional on the information at time $N-1$, we are on average better stopping, we should stop. Otherwise we continue, and hence we can calculate the value function at $N-1$. This is exactly the idea expressed in the Bellman equation (5). Formally, we can define the candidate for our Snell envelope[5] S_n inductively as follows:

$$S_n = \begin{cases} G_N, & n = N \\ \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}, & n = N-1, N-2, \dots, 0. \end{cases} \tag{9}$$

Then the following result holds:

Theorem 2.1. *For each n, $V_n = \mathbb{E}S_n$ is the value of the optimal stopping problem:*

$$V_n = \sup_{\tau\in\mathcal{M}_n^N} \mathbb{E}G_\tau, \tag{10}$$

[5] The difference between the value function and the Snell envelope is that the value function is a numerical value, possibly depending on the choice of the state in the Markovian context, whereas the Snell envelope is a random variable. In the example of the previous section $V(\cdot)$ was the value function, $V(X_n)$ the Snell envelope.

assuming (8) *holds. Moreover:*

(i) The stopping time

$$\tau_n = \inf\{n \leq k \leq N : S_k = G_k\}$$

is optimal in (10)*;*

(ii) The process $(S_k)_{0 \leq k \leq N}$ *is the smallest supermartingale which dominates* $(G_k)_{0 \leq k \leq N}$*;*

(iii) The stopped process $(S_{k \wedge \tau_n})_{n \leq k \leq N}$ *is a martingale.*

Proof. Note that condition (8) ensures that all the stochastic processes mentioned are integrable.

We begin by showing that:

$$S_n \geq \mathbb{E}[G_\tau | \mathcal{F}_n], \quad \text{for } \tau \in \mathcal{M}_n^N, \tag{11}$$

$$S_n = \mathbb{E}[G_{\tau_n} | \mathcal{F}_n]. \tag{12}$$

We proceed backwards inductively from $n = N$. Clearly both statements are true for $n = N$. The inductive step follows from the definition of S_n: suppose the statement is true for n and take $\tau \in \mathcal{M}_{n-1}^N$, then

$$\begin{aligned} \mathbb{E}[G_\tau | \mathcal{F}_{n-1}] &= G_{n-1}\mathbf{1}_{\{\tau = n-1\}} + \mathbb{E}[G_{\tau \vee n} | \mathcal{F}_{n-1}]\mathbf{1}_{\{\tau \geq n\}} \\ &\leq S_{n-1}\mathbf{1}_{\{\tau = n-1\}} + \mathbb{E}[\mathbb{E}[G_{\tau \vee n} | \mathcal{F}_n] | \mathcal{F}_{n-1}]\mathbf{1}_{\{\tau \geq n\}}. \end{aligned}$$

Now, $\tau \vee n \in \mathcal{M}_n^N$, so $\mathbb{E}[G_{\tau \vee n} | \mathcal{F}_n] \leq S_n$, and by the definition of S_{n-1},

$$\begin{aligned} \mathbb{E}[G_\tau | \mathcal{F}_{n-1}] &\leq S_{n-1}\mathbf{1}_{\{\tau = n-1\}} + \mathbb{E}[S_n | \mathcal{F}_{n-1}]\mathbf{1}_{\{\tau \geq n\}} \\ &\leq S_{n-1}\mathbf{1}_{\{\tau = n-1\}} + S_{n-1}\mathbf{1}_{\{\tau \geq n\}} \\ &\leq S_{n-1}. \end{aligned}$$

If we now replace τ above with τ_n, and note that $\tau_{n-1} = \tau_n$ and $S_{n-1} = \mathbb{E}[S_n | \mathcal{F}_{n-1}]$ on the set $\{\tau_{n-1} \geq n\}$, all the inequalities above can be seen to be equalities, and thus:

$$S_{n-1} = \mathbb{E}[G_{\tau_n} | \mathcal{F}_{n-1}].$$

Taking expectations in (11) and (12), we conclude that $\mathbb{E}S_n \geq \mathbb{E}G_\tau$ for any $\tau \in \mathcal{M}_n^N$, and further that $\mathbb{E}S_n = \mathbb{E}G_{\tau_n}$. Hence (10) holds, and τ_n is optimal for this equation.

The supermartingale and dominance properties follow from (12). To show S_n is the smallest supermartingale dominating S_n, consider an alternative supermartingale U_n which also dominates G_n. Then clearly, $S_N = G_N \leq U_N$, and it follows from (9) that if $S_n \leq U_n$, then also $S_{n-1} \leq U_{n-1}$

The martingale property finally follows from the fact that $(S_{k \wedge \tau_n})_{n \leq k \leq N}$ is a supermartingale with $\mathbb{E}S_{n \wedge \tau_n} = \mathbb{E}S_{N \wedge \tau_n}$. □

We now want to try and move to the infinite-horizon case, but there is a technical issue: ideally, we should have been able to write the value function as

$$S_n = \sup_{\tau \in \mathcal{M}_n^N} \mathbb{E}[G_\tau | \mathcal{F}_n],$$

however the set of stopping times in $\mathcal{M}_n^N$ will typically be uncountable, so that the right-hand side may not be measurable! To get round this, we need to introduce the concept of an **essential supremum**.

Theorem 2.2. *Let $\{Z_\alpha; \alpha \in I\}$ be a collection of random variables in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with I an arbitrary index set. Then there exists a unique (up to $\mathbb{P}$-null sets) random variable $Z^* : \Omega \to \bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ such that:*

(i) $\mathbb{P}(Z_\alpha \leq Z^*) = 1, \quad \forall \alpha \in I,$

(ii) if $Y : \Omega \to \bar{\mathbb{R}}$ is another random variable satisfying (i), then

$$\mathbb{P}(Z^* \leq Y) = 1.$$

*We call Z^* the **essential supremum** of $\{Z_\alpha; \alpha \in I\}$, and write*

$$Z^* = \operatorname*{esssup}_{\alpha \in I} Z_\alpha.$$

Moreover, there exists a countable subset J of I, such that

$$Z^* = \sup_{\alpha \in J} Z_\alpha.$$

Sketch of Proof. Mapping by a suitable $f : \bar{\mathbb{R}} \to [-1, 1]$, we may assume all Z_α are bounded. Let $\mathcal{C}$ be the set of countable subsets of I. Then

$$a = \sup_{C \in \mathcal{C}} \mathbb{E}\left[\sup_{\alpha \in C} Z_\alpha\right] = \sup_{n \geq 1} \mathbb{E}\left[\sup_{\alpha \in C_n} Z_\alpha\right]$$

for some suitable sequence $C_n \subseteq \mathcal{C}$. However, $\bigcup_{n \geq 1} C_n$ is a countable set, so we may define the random variable

$$Z^* = \sup_{\bigcup_n C_n} Z_n.$$

From here, it is relatively easy to check that Z^* has the required properties. □

We're now able to define the Snell envelope for the infinite horizon problem.

Suppose we want to solve a problem of the form:

$$\sup_{\tau \in \mathcal{M}} \mathbb{E} G_\tau$$

then our candidate Snell envelope at time n is:

$$S_n = \operatorname*{esssup}_{\tau \in \mathcal{M}_n} \mathbb{E}\left[G_\tau | \mathcal{F}_n\right], \tag{13}$$

with the associated optimal strategy

$$\tau_n = \inf\{k \geq n : S_k = G_k\}. \tag{14}$$

Theorem 2.3. *Suppose* (8) *holds with* $\mathcal{N} = \mathcal{M}$*, and let* S_n*,* τ_n *be as defined in* (13) *and* (14)*. Fix* n*, and suppose that* $\mathbb{P}(\tau_n < \infty) = 1$*. Then* $V_n = \mathbb{E}S_n$ *is the value of the optimal stopping problem:*

$$V_n = \sup_{\tau \in \mathcal{M}_n} \mathbb{E}G_\tau. \tag{15}$$

Moreover,

(i) The stopping time τ_n *is optimal in* (15)*;*

(ii) The process $(S_k)_{k \geq n}$ *is the smallest supermartingale which dominates* $(G_k)_{k \geq n}$*;*

(iii) The stopped process $(S_{k \wedge \tau_n})_{k \geq n}$ *is a martingale.*

Finally, if $\mathbb{P}(\tau_n < \infty) < 1$*, there is no optimal stopping time in* (15)*.*

Proof. We begin by showing that $S_n \geq \mathbb{E}[S_{n+1}|\mathcal{F}_n]$. Suppose that $\sigma_1, \sigma_2 \in \mathcal{M}_{n+1}$ and let $A = \{\mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}] \geq \mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}]\} \in \mathcal{F}_{n+1}$. Then we can define a stopping time

$$\sigma_3 = \sigma_1 \mathbf{1}_A + \sigma_2 \mathbf{1}_{A^C} \in \mathcal{M}_{n+1}.$$

Hence:

$$\begin{aligned} \mathbb{E}[G_{\sigma_3}|\mathcal{F}_{n+1}] &= \mathbf{1}_A \mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}] + \mathbf{1}_{A^C} \mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}] \\ &= \max\left\{\mathbb{E}[G_{\sigma_1}|\mathcal{F}_{n+1}], \mathbb{E}[G_{\sigma_2}|\mathcal{F}_{n+1}]\right\}. \end{aligned}$$

Now, by Theorem 2.2, we know there exists a countable subset $J \subseteq \mathcal{M}_{n+1}$ such that

$$S_{n+1} = \sup_{\tau \in J} \mathbb{E}[G_\tau|\mathcal{F}_{n+1}].$$

Moreover, we can assume that J is closed under the operation described above (i.e. if $\sigma_1, \sigma_2 \in J$ then σ_3 is also in J) without losing the countability of J. Hence, there is a sequence of stopping times σ_k such that

$$S_{n+1} = \lim_{k \to \infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]$$

where $\mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]$ is a sequence of increasing random variables. So

$$\begin{aligned} \mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}\left[\lim_{k \to \infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]|\mathcal{F}_n\right] \\ &= \lim_{k \to \infty} \mathbb{E}\left[\mathbb{E}[G_{\sigma_k}|\mathcal{F}_{n+1}]|\mathcal{F}_n\right] \\ &= \lim_{k \to \infty} \mathbb{E}[G_{\sigma_k}|\mathcal{F}_n] \\ &\leq S_n \end{aligned} \tag{16}$$

In addition, for all $\tau \in \mathcal{M}_n$, we get (by conditioning on $\{\tau = n\}, \{\tau > n\}$)

$$\mathbb{E}[G_\tau|\mathcal{F}_n] \leq \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\},$$

so $S_n \leq \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}$, however clearly also $S_n \geq G_n$, and (16) imply

$$S_n = \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}. \tag{17}$$

Then, for $n \leq k < \tau_n$, $S_k = \mathbb{E}[S_{k+1}|\mathcal{F}_k]$ a.s., that is:

$$S_{k\wedge\tau_n} = \mathbb{E}[S_{(k+1)\wedge\tau_n}|\mathcal{F}_k], \quad \forall k \geq n, \tag{18}$$

and so, for $N > n$,

$$S_n = \mathbb{E}[G_{\tau_n}\mathbf{1}_{\{\tau_n<N\}} + S_N\mathbf{1}_{\{\tau_n\geq N\}}|\mathcal{F}_n].$$

In particular, if $\mathbb{P}(\tau_n < \infty)$, we can let $N \to \infty$, and use (8) and dominated convergence to conclude:

$$S_n = \mathbb{E}[G_{\tau_n}|\mathcal{F}_n] = \operatorname*{esssup}_{\tau\in\mathcal{M}_n} \mathbb{E}[G_{\tau_n}|\mathcal{F}_n].$$

So $V_n = \mathbb{E}S_n = \sup_{\tau\in\mathcal{M}_n}\mathbb{E}[G_\tau]$, and τ_n is optimal.

Suppose $(U_k)_{k\geq n}$ is also a supermartingale dominating $(G_k)_{k\geq n}$. Then

$$S_k = \mathbb{E}[G_{\tau_k}|\mathcal{F}_k] \leq \mathbb{E}[U_{\tau_k}|\mathcal{F}_k] \leq U_k,$$

where the final inequality follows from conditional Fatou, applied to $U_{m\wedge\tau_k} \geq -\sup_{n\leq r\leq\tau_k}|G_r|$ which (by (8)) is an integrable random variable, independent of m.

We can conclude that the stopped process is a martingale from (18) and the integrability condition.

For the final statement in the theorem, suppose that there is another stopping time $\tau^* \in \mathcal{M}_n$, then if $\mathbb{P}(S_{\tau^*} < G_{\tau^*}) > 0$, we get:

$$\mathbb{E}G_{\tau^*} < \mathbb{E}S_{\tau^*} \leq \mathbb{E}S_n = V_n$$

so that any optimal τ^* has $\mathbb{P}(\tau^* \geq \tau_n) = 1$. In particular, if $\mathbb{P}(\tau_n < \infty) < 1$, there is no optimal τ^* with $\mathbb{P}(\tau^* < \infty) = 1$. □

2.2 Markovian setting

Now consider a time-homogeneous Markov chain, $X = (X_n)_{n\geq 0}$, $X_n \in E$ for some measurable space $(E, \mathcal{B})$, defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P}_x)$, with $\mathbb{P}_x(X_0 = x)$. Further, for simplicity, we will assume that $(\Omega, \mathcal{F}) = (E^{\mathbb{Z}_+}, \mathcal{B}^{\mathbb{Z}_+})$, so that the shift operators $\theta_n : \Omega \to \Omega$ can be defined by $(\theta_n(w))_k = w_{n+k}$, and $\mathcal{F}_0$ is trivial.

Of course, this is just a special case of the setting considered in Section 2.1, so can we say anything extra? Consider an infinite horizon problem:

$$V(x) = \sup_{\tau\in\mathcal{M}} \mathbb{E}_x G(X_\tau) \tag{19}$$

where we note that we have replaced the general process G_n by a process depending only on the current state $G(X_n)$. Assume that there is an optimal strategy for the problem. Before, the solution was expressed in terms of the Snell envelope. The next result shows that we can make a more explicit connection between the Snell envelope:

Lemma 2.4. *Let S_n be the Snell envelope as given by* (13)*. Then we have:*

$$S_n = V(X_n), \tag{20}$$

$\mathbb{P}_x$-a.s., for all $x \in E, n \geq 0$.

Proof. We first note that $V(X_n) \leq S_n$, since the latter is taken over all stopping times in $\mathcal{M}_n$, whereas the former is equal to the essential supremum where the supremum is taken over stopping times in $\mathcal{M}_n$ of the form $\tau_n = n + \tau \circ \theta_n$. We also note that

$$\begin{aligned} V(x) &\geq \sup_{\tau \in \mathcal{M}_1} \mathbb{E}_x[G(X_\tau)] \\ &\geq \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x[G(X_{1+\tau\circ\theta_1})] \\ &\geq \sup_{\tau \in \mathcal{M}_0} \mathbb{E}_x[\mathbb{E}_{X_1} G(\tilde{X}_\tau)]] \end{aligned}$$

where X and $\tilde{X}$ are equal in law

$$\begin{aligned} &\geq \mathbb{E}_x\left[\operatorname*{esssup}_{\tau \in \mathcal{M}_0} \mathbb{E}_{X_1}[G(\tilde{X}_\tau)]\right] \\ &\geq \mathbb{E}_x[V(X_1)]. \end{aligned}$$

In particular, we can conclude that $V(X_n)$ is a supermartingale, and dominates $G(\cdot)$, but is smaller than S_n. By property *(ii)* of Theorem 2.3, we deduce that $S_n = V(X_n)$. □

So, whereas before we needed to work with the Snell envelope, now we need only consider $V(\cdot)$.

To help matters, we note the key equation (17) becomes:

$$\begin{aligned} S_n = V(X_n) &= \max\{G_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]\} \\ &= \max\{G_n, \mathbb{E}[V(X_{n+1})|\mathcal{F}_n]\} \\ &= \max\{G_n, \mathbb{E}_{X_n}[V(X_1)]\}, \end{aligned} \tag{21}$$

and we introduce an operator T which maps functions $F : E \to \mathbb{R}$ by:

$$(TF)(x) = \mathbb{E}_x F(X_1).$$

We also introduce an important concept from harmonic analysis:

Definition 2.5. We say that $F : E \to \mathbb{R}$ is **superharmonic** if $TF(x)$ is well defined, and

$$TF(x) \leq F(x)$$

for all $x \in E$.

Then, provided $\mathbb{E}_x|F(X_n)| < \infty$ for all $x \in E$, we get:

F is superharmonic iff $(F(X_n))_{n\geq 0}$ is a supermartingale for all $\mathbb{P}_x, x \in E$.

Note: for the example in Section 1.2, we had $TF(x) = \frac{1}{2}(F(x+1) + F(x-1))$, so $TF(x) \leq F(x)$ if and only if F is concave.

We are now in a position to state an analogue of Theorem 2.3. To update certain notions, we state the corresponding versions of (8):

$$\sup_{\tau \in \mathcal{N}} \mathbb{E}\left[\sup_{n \leq \tau} |G(X_n)|\right] < \infty. \tag{22}$$

Further, given the value function V, we define the continuation region:

$$C = \{x \in E : V(x) > G(x)\}$$

and the stopping region

$$D = \{x \in E : V(x) = G(x)\}.$$

The candidate optimal stopping time is then:

$$\tau_D = \inf\{n \geq 0 : X_n \in D\}. \tag{23}$$

Theorem 2.6. *Suppose* (22) *holds, with* $\mathcal{N} = \mathcal{M}$, *and let* $V(\cdot)$ *be defined by* (19), *and* τ_D *be given by* (23). *Suppose that* $\mathbb{P}(\tau_D < \infty) = 1$. *Then*

(i) the stopping time τ_D *is optimal in* (19);

(ii) the value function is the smallest superharmonic function which dominates the gain function G *on* E;

(iii) the stopped process $(V(X_{n \wedge \tau_D}))_{n \geq 0}$ *is a* $\mathbb{P}_x$*-martingale for every* $x \in E$

Finally, if $\mathbb{P}(\tau_D < \infty) < 1$, *there is no optimal stopping time in* (19).

Proof. This follows directly from Theorem 2.3, having made the identification $S_n = V(X_n)$ in (20) □

Note that we may rephrase the Bellman equation (21) in terms of the operator T as:

$$V(x) = \max\{G(x), TV(x)\},$$

and the right hand side is really just an operator, so if we define a new operator:

$$QF(x) = \max\{G(x), TF(x)\},$$

we see that the Bellman equation becomes: $V = QV$.

This gives us a relatively easy starting point to look for candidate value functions, but note that solving $V = QV$ is not generally sufficient: consider the 'Markov' process $X_n = n$, and let $G(x) = (1 - \frac{1}{x})$. Then $TV(x) = V(x+1)$, so the superharmonic functions are simply decreasing functions, and $QV(x) = \max\{1 - \frac{1}{x}, V(x+1)\}$. Clearly, $V \equiv 1$, and $QV = V$ for this choice, *but* $QF = F$ whenever $F \equiv$ *const* ≥ 1.

We are, however able to say some things about the value function in terms of the operator Q:

Lemma 2.7. *Under the hypotheses of Theorem 2.6,*

$$V(x) = \lim_{n\to\infty} Q^n G(x), \quad \forall x \in E.$$

Moreover, suppose that F satisfies $F = QF$ and

$$\mathbb{E}\left(\sup_{n\geq 0} F(X_n)\right) < \infty.$$

Then $F = V$ iff:

$$\limsup_{n\to\infty} F(X_n) = \limsup_{n\to\infty} G(X_n), \quad \mathbb{P}_x\text{-a.s.}, \forall x \in E,$$

In which case, $F(X_n)$ converges $\mathbb{P}_x$-a.s. for all $x \in E$.

We won't prove this, but see Peskir and Shiryaev, Corollary 1.12 and Theorem 1.13, and the comments at the end of this chapter.

One of the important features of this result, and one that proves to be important in practice, is that we now have a set of conditions that we can check to confirm that a candidate for the value function is indeed the value function — a common feature of many practical problems is that the solution is obtained by 'guessing' a plausible V, and checking that the function solves an appropriate version of $V = QV$, along with any other suitable technical conditions.

We end the discrete time theory with a quick discussion on the time-inhomogeneous case: we can move from a time-inhomogeneous Markov process Z_n to the time-homogeneous case X_n by considering the process

$$X_n = (Z_n, n),$$

and e.g. $TF(z,n) = \mathbb{E}_{z,n}F(Z_{n+1}, n+1)$, and the same results typically hold.

Also, we can think of the finite-horizon, time-homogeneous case as a special case of the time-inhomogeneous case, in particular, if T is the operator associated with a time-homogeneous process X_n:

$$TF(x) = \mathbb{E}_x[F(X_1)],$$

and N is the time-horizon, if we write $V_n(x)$ for

$$V_n(x) = \sup_{\tau\in\mathcal{M}_n^N} \mathbb{E}[G(X_\tau)|\mathcal{F}_n],$$

then

$$\begin{aligned} V_n(x) &= \max\{G(x), TV_{n+1}(x)\} \\ &= QV_{n+1}(x). \end{aligned}$$

Using $V_N(x) = G(x)$, we then get:

$$\begin{aligned} V_n(x) &= Q^{N-n}V_N(x) \\ &= Q^{N-n}G(x). \end{aligned}$$

From which we get:

$$V_0(x) = Q^N G(x).$$

Note that there is no issue with multiple solutions when we have a finite horizon.

Finally, we comment on the proof of Lemma 2.7: the term $\lim_{n\to\infty} Q^n G(x)$ can now be interpreted as the limit of the finite horizon problem as we let the horizon go to infinity. If the stopped gain function is well behaved (in the sense of (22)), and the stopping time τ_D is almost surely finite, we can conclude the convergence. Now if $F = QF$, and the integrability condition holds, we see that $F(X_n)$ is a supermartingale dominating the value function. The final condition can then be used to show that the candidate value function is indeed the smallest such supermartingale.

3 Continuous Time

3.1 Martingale treatment

Mostly, the ideas in continuous time remain identical to the ideas in the discrete time setting. To fill in the details, we need to address a few technical issues, but mostly the proofs will be similar to the discrete case.

Let our gains process $(G_t)_{t\geq 0}$ be an adapted process, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, satisfying the usual conditions[6]. Moreover, we will suppose that the process is right-continuous, and *left-continuous over stopping times*: if τ_n is a sequence of stopping times, increasing to a stopping time τ, then $G_{\tau_n} \to G_\tau$, $\mathbb{P}$-a.s.. Note that this is a weaker condition than left-continuity. Our new integrability condition will be:

$$\sup_{\tau\in\mathcal{N}} \mathbb{E}\left[\sup_{0\leq t\leq\tau} |G_\tau|\right] < \infty. \tag{24}$$

Note in particular that with such an integrability condition, we may apply the optional sampling theorem to any supermartingale which dominates the gains process: let $Y_t \geq G_t$ be a supermartingale, and $\tau \in \mathcal{N}, \tau \geq t$ an almost-surely finite stopping time. Then $Y_t \geq \mathbb{E}[Y_{\tau\wedge N}|\mathcal{F}_t]$, and $\inf_{N\in\mathbb{N}} Y_{\tau\wedge N} \geq -\sup_{0\leq t\leq\tau} |G_\tau|$, so dominated convergence allows us to deduce that $Y_t \geq \mathbb{E}[Y_\tau|\mathcal{F}_t]$.

We suppose $\mathcal{N} = \mathcal{M}_t^T$, for some $t < T$, (including the case $T = \infty$, which we interpret as the a.s. continuous stopping times) and consider the problem:

$$V_t = \sup_{\tau\in\mathcal{M}_t^T} \mathbb{E}G_\tau. \tag{25}$$

Now, introduce the process

$$S_t^* = \operatorname*{esssup}_{\tau\in\mathcal{M}_t^T} \mathbb{E}[G_\tau|\mathcal{F}_t], \tag{26}$$

[6] That is, $(\mathcal{F}_t)_{t\geq 0}$ is complete ($\mathcal{F}_0$ contains all $\mathbb{P}$-negligible sets), and right-continuous ($\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$).

and we would hope that such a process could be our Snell envelope, with the stopping time:

$$\tau_t = \inf\{t \le r \le T : S^*_r = G_r\}. \tag{27}$$

There is however a technical issue here: although our gains process and the filtration are both right continuous, the process S^*_t is not *a priori* right-continuous: notably, we could, for fixed t, redefine S^*_t on a set of probability zero without contradicting the definition (26). So pick any such S^*_t, and suppose further that the probability space admits an independent, $U(0,1)$ random variable U. Then choose a version of (S^*_t), say $(\tilde{S}_t)$ which has $S^*_U = G_U$, but otherwise has $\tilde{S}_t = S^*_t$. This is a possible alternative choice according to (26), but note that, if we stop both processes according to their respective versions of (27), we get $\tilde{S}_{\tilde{\tau}_t} = S^*_{\tau^*_t \wedge U}$, which will in general be substantially different!

The solution to this is from the following:

Lemma 3.1. *Let S^*_t be as given in (26). Then S^*_t is a supermartingale and $\mathbb{E}S^*_t = V_t$. In addition, S^*_t admits a right-continuous version, S_t, which is also a supermartingale, and for which $\mathbb{E}S_t = V_t$.*

Proof. The supermartingale property follows by using an identical proof to that used in Theorem 2.3, to derive (16). The equality $\mathbb{E}S^*_t = V_t$ then follows from the definition of an essential supremum.

For the final statement, we note that the existence of S_t will follow if we can show that $t \mapsto \mathbb{E}S^*_t$ is right-continuous (e.g. Revuz & Yor, Theorem II.2.9). From the martingale property, we get $\mathbb{E}S^*_t \ge \mathbb{E}S^*_{t+\delta}$, so that $\lim_{t_n \downarrow t} \mathbb{E}S^*_{t_n} \le \mathbb{E}S^*_t$. Now choose $\tau \in \mathcal{M}^T_t$, and consider $\tau_n = \tau \vee (t + \frac{1}{n}) \in \mathcal{M}_{t+\frac{1}{n}}$. Then right continuity of G_t implies $\lim_{n\to\infty} G_{\tau_n} = G_\tau$, and so we get $\mathbb{E}G_\tau \le \liminf_{n\to\infty} \mathbb{E}G_{\tau_n} \le \liminf_{n\to\infty} \mathbb{E}S^*_{t+\frac{1}{n}}$. □

Now we have a suitably defined Snell envelope, we are in a position to state the main theorem:

Theorem 3.2. *Let S_t be the process defined in Lemma 3.1, and define the stopping time*

$$\tau_t = \inf\{t \le r \le T : S_r = G_r\}. \tag{28}$$

Suppose (24) holds with $\mathcal{N} = \mathcal{M}$. Fix t, and suppose that $\mathbb{P}(\tau_t < \infty) = 1$. Then $V_t = \mathbb{E}S_t$ is the value of the optimal stopping problem:

$$V_t = \sup_{\tau \in \mathcal{M}^T_t} \mathbb{E}G_\tau. \tag{29}$$

Moreover,

(i) The stopping time τ_t is optimal in (29);

(ii) The process $(S_r)_{t \le r \le T}$ is the smallest right-continuous supermartingale which dominates $(G_r)_{t \le r \le T}$;

(iii) The stopped process $(S_{r \wedge \tau_t})_{t \le r \le T}$ is a right-continuous martingale.

Finally, if $\mathbb{P}(\tau_t < \infty) < 1$*, there is no optimal stopping time in* (29).

Proof. That (29) holds, and S_t is the smallest right-continuous martingale dominating G_t follow from the previous lemma, and the usual argument: if $Y_t \geq G_t$ is another right-continuous supermartingale, $Y_t \geq \mathbb{E}[Y_\tau|\mathcal{F}_t] \geq \mathbb{E}[G_\tau|\mathcal{F}_t]$, so $Y_t \geq S_t$. Note that right-continuity in fact implies $\mathbb{P}(Y_r \geq S_r \text{ for all } r \geq t) = 1$.

The major new issue concerns the optimality of τ. Previously, we were able to prove optimality using the Bellman equation, but this no longer makes sense when we move to the continuous time setting. Instead, we need to find a new argument.

Suppose initially that the gains process satisfies the additional condition $G_t \geq 0$. Then, for $\lambda \in (0, 1]$ introduce

$$\tau_t^\lambda = \inf\{r \geq t : \lambda S_r \leq G_r\},$$

so $\tau_t^1 \equiv \tau_t$. Right continuity of S_t, G_t implies

$$\lambda S_{\tau_t^\lambda} \leq G_{\tau_t^\lambda},$$
$$\tau_{t+}^\lambda = \tau_t^\lambda.$$

Moreover, $S_t \geq \mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_t]$, since S_t is a supermartingale. We want to show that $S_t = \mathbb{E}[S_{\tau_t^1}|\mathcal{F}_t]$.

Suppose $\lambda \in (0, 1)$, and define

$$R_t = \mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_t],$$

and then, for $s < t$,

$$\begin{aligned}\mathbb{E}[R_t|\mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_t]|\mathcal{F}_s] \\ &= \mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_s] \\ &\leq \mathbb{E}[S_{\tau_s^\lambda}|\mathcal{F}_s] \\ &\leq R_s.\end{aligned}$$

Then R_s is a supermartingale, and an identical argument to the proof of Lemma 3.1 allows us to deduce $t \mapsto \mathbb{E}S_{\tau_t^\lambda}$ is decreasing, and therefore we can find a right-continuous version of R_t. In particular, since $G_t \geq 0$, we also have $R_t \geq 0$.

Now consider the (right-continuous) process

$$L_t = \lambda S_t + (1-\lambda) R_t.$$

Then, for $\lambda \in (0, 1)$:

$$\begin{aligned}L_t &= \lambda S_t + (1-\lambda) R_t \mathbf{1}_{\{\tau_t^\lambda = t\}} + (1-\lambda) R_t \mathbf{1}_{\{\tau_t^\lambda > t\}} \\ &\geq \lambda S_t + (1-\lambda)\mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_t]\mathbf{1}_{\{\tau_t^\lambda = t\}} \\ &\geq \lambda S_t + (1-\lambda) S_t \mathbf{1}_{\{\tau_t^\lambda = t\}} \\ &\geq S_t \mathbf{1}_{\{\tau_t^\lambda = t\}} + \lambda S_t \mathbf{1}_{\{\tau_t^\lambda > t\}} \\ &\geq G_t,\end{aligned}$$

where we have used: $\tau_t^\lambda > t$ implies $\lambda S_t > G_t$. Since L_t is now a supermartingale dominating G_t, we must also have $L_t \geq S_t$, and therefore $S_t \leq R_t$. So in fact $S_t = R_t$, and rearranging, we get

$$S_t = \mathbb{E}[S_{\tau_t^\lambda}|\mathcal{F}_t] \leq \frac{1}{\lambda}\mathbb{E}[G_{\tau_t^\lambda}|\mathcal{F}_t]. \tag{30}$$

But as $\lambda \uparrow 1$, τ_t^λ increases to some limiting (possibly infinite) stopping time σ, and by the left continuity for stopping times of G_t, we get $S_t \leq \mathbb{E}[G_\sigma|\mathcal{F}_t]$.

So it just remains to show that $\sigma = \tau_t^1$ (since we already have $S_t \geq \mathbb{E}[G_\sigma|\mathcal{F}_t]$).

Since $\tau_t^1 \geq \tau_t^\lambda$ for $\lambda \in (0,1)$, we have $\tau_t^1 \geq \sigma$, but if $\{\sigma < \tau_t^1\}$ has positive probability, then also with positive probability we have $S_\sigma > G_\sigma$, and hence

$$S_t = \mathbb{E}[G_\sigma|\mathcal{F}_t] < \mathbb{E}[S_\sigma|\mathcal{F}_t]$$

with positive probability, but this contradicts the supermartingale property of S_t. Hence we can conclude that $S_t = \mathbb{E}[G_{\tau_t^1}|\mathcal{F}_t] = \mathbb{E}[S_{\tau_t^1}|\mathcal{F}_t]$.

Now we note that we may remove the assumption that $G_t \geq 0$. Since we are assuming (24),

$$M_t = \mathbb{E}\left[\inf_{0\leq t\leq T} G_t|\mathcal{F}_t\right]$$

is a right-continuous, uniformly integrable martingale. Now define a new problem with gain process $\tilde{G}_t = G_t - M_t$, and note that this is now non-negative, and right-continuous (although not necessarily left-continuous for stopping times). Then if we solve the problem for $\tilde{G}_t$, we get:

$$\tilde{S}_t = \operatorname*{esssup}_{\tau\in\mathcal{N}} \mathbb{E}[\tilde{G}_\tau|\mathcal{F}_t] = \operatorname*{esssup}_{\tau\in\mathcal{N}} \mathbb{E}[G_\tau - M_\tau|\mathcal{F}_t] = S_t - M_t.$$

We can now apply the above argument to the processes $(\tilde{G}_t, \tilde{S}_t)$ up to the step at (30), and from there reverting to the processes (G_t, S_t), we get the desired conclusion.

The remaining claims in the theorem follow with the usual arguments. □

3.2 Markovian setting

Take $(X_t)_{t\geq 0}$ a Markov process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_x)$ on a measurable space $(E, \mathcal{B})$ where, for simplicity, we suppose $E = \mathbb{R}^d, \mathcal{B} = Borel(\mathbb{R}^d)$, and $(\Omega, \mathcal{F}) = (E^{[0,\infty)}, \mathcal{B}^{[0,\infty)})$, so that the shift operators are measurable. In addition, we again suppose the sample paths of $(X_t)_{t\geq 0}$ are right-continuous, and left-continuous for stopping times. As usual, we also introduce an integrability condition:

$$\sup_{\tau\in\mathcal{N}} \mathbb{E}_x\left[\sup_{t\leq\tau}|G(X_t)|\right] < \infty. \tag{31}$$

Our problem of interest is:

$$V(x) = \sup_{\tau\in\mathcal{N}} \mathbb{E}_x G(X_\tau). \tag{32}$$

In this section, we will in general consider the case where $\mathcal{N} = \mathcal{M}$, although we will also at times consider the case $\mathcal{N} = \mathcal{M}^T$, for which all the results we state will also hold (and in cases, will be simpler).

The argument we used in Lemma 2.4 can be used essentially unmodified to see that

$$V(x) \geq \mathbb{E}_x V(X_s)$$

for all $x \in E, s \geq 0$, and therefore $V(X_t) = S_t, \mathbb{P}_x$-a.s., using the same reasoning.

Again, we define the continuation and stopping regions by:

$$C = \{x \in E : V(x) > G(x)\}$$
$$D = \{x \in E : V(x) = G(x)\}.$$

We want $\tau_D = \inf\{t \geq 0 : X_t \in D\}$ to be a stopping time, and $X_{\tau_D} \in D$. For a right-continuous process in a right-continuous filtration, this is true if D is closed. Recall also that a function is *lower semi-continuous (upper semi-continuous)* if $f(x) \leq \liminf_{x_n \to x} f(x_n)$ $(f(x) \geq \limsup_{x_n \to x} f(x_n))$ for all $x \in E$. Equivalently, f is lower semi-continuous (lsc) if $\{x \in E : f(x) \leq A\}$ is closed for any $A \in \mathbb{R}$. Similarly, f is upper semi-continuous (usc) if $\{x \in E : f(x) \geq A\}$ is closed. Note that if f is usc, $-f$ is lsc, and if f, g are usc/lsc, $f + g$ is usc/lsc. Then D is closed if $\{V - G \leq 0\}$ is closed, which occurs if $V - G$ is lsc, and therefore if V is lsc and G is usc. It is under these conditions that we will generally work.

The restriction to lsc V allows us to rule out certain pathological examples (e.g. $V(x) = \mathbf{1}_{\{x=0\}} = G(x)$, when $X_t = BM(\mathbb{R}^2)$), but is also not a strong constraint: if $x \mapsto \mathbb{E}_x G(X_\tau)$ is continuous (or lsc) for each τ, then $x \mapsto \mathbb{E}_x G(X_\tau)$ is also lsc. The usc of G ensures that we don't stop outside the stopping region: e.g. $G(x) = \mathbf{1}_{\{x \neq 0\}}$, then $\mathbb{P}_0(\tau_D = 0) = 1$, but $G(X_{\tau_D}) = 0$.

We can also extend Definition 2.5 to the current context:

Definition 3.3. A measurable function $F : E \to \mathbb{R}$ is **superharmonic** if

$$\mathbb{E}_x F(X_\sigma) \leq F(x)$$

for all stopping times σ and all $x \in E$. (Note that this requires $F(X_\sigma) \in \mathcal{L}^1(\mathbb{P}_x)$ for all $x \in E$.)

Lemma 3.4. *Suppose F is lsc and $(F(X_t))_{t \geq 0}$ is uniformly integrable. Then F is superharmonic if and only if $(F(X_t))_{t \geq 0}$ is a right-continuous supermartingale under $\mathbb{P}_x$, for each $x \in E$.*

The 'if' direction is clear. For the 'only if' direction, if F is superharmonic,

$$F(X_t) \geq \mathbb{E}_{X_t} F(X_s) = \mathbb{E}_x \left[F(X_{t+s}) | \mathcal{F}_t\right].$$

So F is a super-martingale. The hard part is checking right-continuity. We leave the proof of this fact to Peskir & Shiryaev, Proposition I.2.5.

From this, we get the following theorem:

Theorem 3.5. *Suppose there exists an optimal stopping time τ^* with $\mathbb{P}(\tau^* < \infty) = 1$, so that*

$$V(x) = \mathbb{E}_x G(X_{\tau^*})$$

for all $x \in E$. Then

(i) V is the smallest superharmonic function dominating G on E;

Moreover, if V is lsc, and G is usc, then also

(ii) $\tau_D \leq \tau^*$ for all $x \in E$, and τ_D is optimal;

(iii) the stopped process $(V(X_{t\wedge\tau_D}))_{t\geq 0}$ is a right-continuous $\mathbb{P}_x$-martingale.

Proof.

$$\begin{aligned}\mathbb{E}_x V(X_\sigma) = \mathbb{E}_x \mathbb{E}_{X_\sigma} G(X_{\tau^*}) &= \mathbb{E}_x \mathbb{E}_x \left[G(X_{\tau^*}) \circ \theta_\sigma | \mathcal{F}_\sigma\right] \\ &= \mathbb{E}_x G(X_{\sigma+\tau^*\circ\theta_\sigma}) \leq \sup_\tau \mathbb{E}_x G(X_\tau) = V(x).\end{aligned}$$

So V is a super-harmonic function, dominating G. For any other super-harmonic F we get

$$\mathbb{E}_x G(X_\tau) \leq \mathbb{E}_x F(X_\tau) \leq F(x),$$

and taking the supremum over τ, we get $V(x) \leq F(x)$. For any optimal stopping time, if $\mathbb{P}_x(V(X_{\tau^*}) > G(X_{\tau^*})) > 0$, we get

$$\mathbb{E}_x G(X_{\tau^*}) < \mathbb{E}_x V(X_{\tau^*}) \leq V(x),$$

so τ^* optimal implies $\tau_D \leq \tau^*$. By Lemma 3.4, $V(X_t)$ is a right-continuous supermartingale, and further $V(X_{\tau_D}) = G(X_{\tau_D})$ by lsc/usc. Then

$$\begin{aligned}V(x) = \mathbb{E}_x G(X_{\tau^*}) &= \mathbb{E}_x V(X_{\tau^*}) \\ &\leq \mathbb{E}_x V(X_{\tau_D}) = \mathbb{E}_x G(X_{\tau_D}) \leq V(x),\end{aligned}$$

where the inequality follows from the supermartingale property.

For *(iii)*:

$$\begin{aligned}\mathbb{E}_x[V(X_{\tau_D})|\mathcal{F}_t] &= \mathbb{E}_x[G(X_{\tau_D})|\mathcal{F}_t] \\ &= \mathbb{E}_{X_t}[G(X_{\tau_D})|\mathcal{F}_t]\mathbf{1}_{\{t<\tau_D\}} + V(X_{\tau_D})\mathbf{1}_{\{t\geq\tau_D\}} \\ &= V(X_{t\wedge\tau_D})\end{aligned}$$

□

This theorem constitutes necessary conditions for the existence of a solution. Now we consider sufficient conditions:

Theorem 3.6. *Suppose there exists a function $\hat{V}$, which is the smallest superharmonic function dominating G on E. Suppose also $\hat{V}$ is lsc and G is usc. Let $D = \{x \in E : \hat{V}(x) = G(x)\}$, and suppose τ_D is as above. Then*

(i) if $\mathbb{P}_x(\tau_D < \infty) = 1$, for all $x \in E$, then $\hat{V} = V$ and τ_D is optimal;

(ii) if $\mathbb{P}_x(\tau_D < \infty) < 1$, for some $x \in E$, then there is no optimal stopping time τ.

Proof. For all τ, x:

$$\mathbb{E}_x G(X_\tau) \leq \mathbb{E}_x \hat{V}(X_\tau) \leq \hat{V}(x).$$

Taking the supremum over all τ, we get $G(x) \leq V(x) \leq \hat{V}(x)$. Suppose in addition $G(\cdot) \geq 0$, and introduce for $\lambda \in (0,1)$,

$$C_\lambda = \{x \in E : \lambda \hat{V}(x) > G(x)\}$$
$$D_\lambda = \{x \in E : \lambda \hat{V}(x) \leq G(x)\}.$$

Then the usc/lsc of $\hat{V}, G$ imply C_λ is open and D_λ is closed. Also, $C_\lambda \uparrow C, D_\lambda \downarrow D$ as $\lambda \uparrow 1$. Taking $\tau_{D_\lambda} = \inf\{t \geq 0 : X_t \in D_\lambda\}$ we get $\tau_{D_\lambda} \leq \tau_D$ a.s., and so $\mathbb{P}_x(\tau_{D_\lambda} < \infty) = 1$ for all $x \in E$.

Note that $x \mapsto \mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}})$ is superharmonic:

$$\begin{aligned}
\mathbb{E}_x\left[\mathbb{E}_{X_\sigma}[\hat{V}(X_{\tau_{D_\lambda}})]\right] &= \mathbb{E}_x\left[\mathbb{E}_x\left[\hat{V}(X_{\sigma+\tau_{D_\lambda}\circ\theta_\sigma})|\mathcal{F}_\sigma\right]\right] \\
&= \mathbb{E}_x \hat{V}(X_{\sigma+\tau_{D_\lambda}\circ\theta_\sigma}) \\
&\leq \mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}})
\end{aligned}$$

since $\hat{V}$ is superharmonic, and $\sigma + \tau_{D_\lambda} \circ \theta_\sigma \geq \tau_{D_\lambda}$.

Now, if $x \in C_\lambda$, then $G(x) < \lambda \hat{V}(x) \leq \lambda \hat{V}(x) + (1-\lambda)\mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}})$ since the last term is non-negative. If $x \in D_\lambda$, $G(x) \leq \hat{V}(x) = \lambda \hat{V}(x) + (1-\lambda)\mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}})$, so for all $x \in E$, we get

$$G(x) \leq \lambda \hat{V}(x) + (1-\lambda)\mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}})$$

But since both terms on the right are superharmonic, the whole expression is, and therefore:

$$\hat{V}(x) \leq \lambda \hat{V}(x) + (1-\lambda)\mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}}),$$

which implies

$$\hat{V} \leq \mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}}).$$

Again using superharmonicity, we conclude:

$$\hat{V}(x) = \mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}}).$$

Since $\hat{V}$ is lsc and G is usc, we deduce

$$\hat{V}(X_{\tau_{D_\lambda}}) \leq \frac{1}{\lambda} G(X_{\tau_{D_\lambda}}), \tag{33}$$

and we get
$$\hat{V}(x) = \mathbb{E}_x \hat{V}(X_{\tau_{D_\lambda}}) \leq \frac{1}{\lambda}\mathbb{E}_x G(X_{\tau_{D_\lambda}}) \leq \frac{1}{\lambda}V(x).$$
Letting $\lambda \uparrow 1$, we conclude that $\hat{V}(x) \leq V(x)$, and therefore $\hat{V} = V$, $V(x) \leq \frac{1}{\lambda}G(X_{\tau_{D_\lambda}})$.

Now consider the sequence τ_{D_λ} as $\lambda \uparrow 1$. Then $\tau_{D_\lambda} \uparrow \sigma$, for some stopping time σ. We need to show $\sigma = \tau_D$. Since $\tau_{D_\lambda} \leq \tau_D$ for $\lambda < 1$, we must have $\sigma \leq \tau_D$. Using (33) and the left-continuity of X_t over stopping times, and lsc of $\hat{V}$, usc of G, we must have:
$$V(X_\sigma) \leq G(X_\sigma),$$
and hence $V(X_\sigma) = G(X_\sigma)$, so that $\sigma \geq \tau_D$ (recall the definition of τ_D.)

Finally, to see that τ_D is indeed optimal, using (33) and Fatou:
$$\begin{aligned}
V(X) &\leq \limsup_{\lambda\uparrow 1} \mathbb{E}_x G(X_{\tau_{D_\lambda}}) \\
&\leq \mathbb{E}_x \limsup_{\lambda\uparrow 1} G(X_{\tau_{D_\lambda}}) \\
&\leq \mathbb{E}_x G(\limsup_{\lambda\uparrow 1} X_{\tau_{D_\lambda}}) \\
&= \mathbb{E}_x G(X_{\tau_{D_\lambda}})
\end{aligned}$$
where in the last two lines, we have used the usc of G, and the left-continuity for stopping times of X_t.

The final statement in the theorem follows from Theorem 3.5.

We will not prove the final step (that we can drop the condition $G \geq 0$), but refer instead to Peskir & Shiryaev. □

This result allows us to find the value function by looking for the smallest superharmonic function dominating G

In fact, if we can directly (via other methods) prove that the value function must be lsc, we can get the following result:

Theorem 3.7. *If V is lsc, G is usc, and $\mathbb{P}_x(\tau_D < 1) = 1$, then τ_D is optimal.*

Of course, in the finite horizon setting, we don't need to check the final condition! As mentioned above, V lsc can also be relatively easy to check if e.g. $x \mapsto \mathbb{E}_x G(X_\tau)$ is continuous for each τ.

Sketch proof. To deduce this from the previous result, we need to show that V is superharmonic. Since V is lsc, it is measurable, and thus, by the Strong Markov property, we have:
$$V(X_\sigma) = \operatorname*{esssup}_{\tau} \mathbb{E}_x[G(X_{\sigma+\tau\circ\theta_\sigma})|\mathcal{F}_\sigma].$$
Using similar arguments to e.g. the proof of Theorem 2.3, we can finds a sequence of stopping times such that
$$V(X_\sigma) = \lim_{n\to\infty} \mathbb{E}_x[G(X_{\sigma+\tau_n\circ\theta_\sigma})|\mathcal{F}_\sigma],$$

where the right hand side is an increasing sequence. Then:

$$\mathbb{E}_x V(X_\sigma) = \lim_{n\to\infty} \mathbb{E}_x[G(X_{\sigma+\tau_n\circ\theta_\sigma})] \leq V(x).$$

Since any other superharmonic function F has

$$\mathbb{E}_x G(X_\sigma) \leq \mathbb{E}_x F(X_\sigma) \leq F(x),$$

V must be the smallest superharmonic function, and now everything follows from Theorem 3.6. □

4 Free Boundary Problems

4.1 Infinitesimal Generators

Let X_t be a suitably nice[7] time-homogeneous Markov process, and write P_t for the transition semi-group of the process — that is, $P_t f(x) = \mathbb{E}_x f(X_t)$.

Suppose $f \in C_0$ (the set of continuous functions with limit 0 at infinity) is sufficiently smooth that

$$\mathbb{L}_X f = \lim_{t\downarrow 0} \frac{1}{t}(P_t f - f)$$

exists in C_0. Then $\mathbb{L}_X : \mathcal{D}_X \to C_0$ is the **infinitesimal generator** of X, where we write $\mathcal{D}_X \subseteq C_0$ for the set of functions where $\mathbb{L}_X$ is defined.

Definition 4.1. If X is a Markov process, a Borel function f belongs to the domain $\mathbb{D}_X$ of the **extended infinitesimal generator** if there exists a Borel function g such that

$$f(X_t) - f(X_0) - \int_0^t g(X_s)\,ds$$

is a right-continuous martingale for every $x \in E$; in which case, we write $g = \mathbb{L}_X f$. It can then be shown that $\mathcal{D}_X \subseteq \mathbb{D}_X$.

Theorem 4.2 (Revuz & Yor, VII.1.13)**.** *If P_t is a 'nice' semi-group on $\mathbb{R}^d$, and $C_K^\infty \subseteq \mathcal{D}_X$,*

(i) $C_K^2 \subseteq \mathcal{D}_X$;

(ii) for every relatively compact open set U, there exist functions a_{ij}, b_i, c on U and a kernel N such that for all $f \in C_K^2$ and $x \in U$:

$$\mathbb{L}_X f(x) = c(x)f(x) + \sum b_i(x)\frac{\partial f}{\partial x_i}(x) + \sum a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \int_{\mathbb{R}\setminus\{x\}} \left[f(y) - f(x) - \mathbf{1}_U(y)\sum(y_i - x_i)\frac{\partial f}{\partial x_i}(x)\right] N(x,dy). \quad (34)$$

$N(x,\cdot)$ is a Radon measure on $\mathbb{R}^d \setminus \{x\}$, a is a symmetric, non-negative matrix, $c \leq 0$ and a, c do not depend on U.

[7]Feller — i.e. if P_t is the transition semi-group, and $P_t f(x) = \mathbb{E}_x f(X_t)$, then for $f \in C_0$, the set of continuous functions tending to zero at infinity, then $||P_t f|| \leq ||f||$, where $||\cdot||$ is the uniform norm, for all $f \in C_0$, $P_t \circ P_s = P_{t+s}$ and $\lim_{t\downarrow 0} ||P_t f - f|| = 0$ for every $f \in C_0$.

We can interpret some of the terms in the operator as follows: the $c(x)$ term encodes the rate at which the process is 'killed' at x, $b(x)$ is the drift, and $(a_{ij}(x))$ the diffusion matrix. $N(x, \cdot)$ is the jump measure. In the case where the measure N is identically zero, the process X_t is continuous.

4.2 Boundary problems

Consider now our value function. If D exists, with τ_D almost surely finite, then $V(X_{t\wedge\tau_D})$ is a right-continuous martingale, and $V(X_t)$ is a supermartingale. Provided $V \in \mathcal{D}_X$, in terms of the operator $\mathbb{L}_X$, this implies we have

$$\mathbb{L}_X V \leq 0 \quad x \in E \tag{35}$$
$$V \geq G \quad x \in E \tag{36}$$
$$\mathbb{L}_X V = 0 \quad x \in C \tag{37}$$
$$V = G \quad x \in D \tag{38}$$

Reversing these arguments, we can try to exploit this as a solution method: suppose that we can find a closed set $\hat{D}$ such that $\mathbb{P}(\tau_{\hat{D}} < \infty) = 1$, and a function $\hat{V}$ such that (35)–(38) holds for this set. Then

$$\hat{V}(X_t) - \hat{V}(X_0) - \int_0^t \mathbb{L}_X \hat{V}(X_s)\, ds$$

is a martingale, and hence $\hat{V}(X_{t\wedge\tau_{\hat{D}}})$ is a martingale, and $\hat{V}(X_t)$ is a supermartingale. Moreover, since $\hat{D}$ is closed, $X_{\tau_{\hat{D}}} \in \hat{D}$, so

$$\hat{V}(x) = \mathbb{E}_x \hat{V}(X_{\tau_{\hat{D}}}) = \mathbb{E}_x G(X_{\tau_{\hat{D}}}),$$

and for any stopping time $\tau \in \mathcal{M}$,

$$\hat{V}(x) \geq \mathbb{E}_x \hat{V}(X_\tau) \geq \mathbb{E}_x G(X_\tau)$$

and therefore $\tau_{\hat{D}}$ must be optimal.

So a common approach will be to 'guess' what D might be, and then define

$$V(x) = \begin{cases} \mathbb{E}_x G(X_{\tau_D}) & : x \in C \\ G(x) & : x \in D \end{cases}$$

and then verify that

$$\mathbb{L}_X V \leq 0 \qquad x \in E$$
$$V \geq G \qquad x \in E$$

In general, we may want to consider slightly more complicated payoffs than a simple function of X_t, where X_t has the generator given by (34). To consider how we might do this, consider first the simple example where X_t is continuous, and so the solution is given by

$$V(x) = \mathbb{E}_x G(X_{\tau_D})$$

so that V is a solution to the Dirichlet problem:

$$\begin{aligned} \mathbb{L}_X V(x) &= 0, \quad x \in C \\ V(x) &= G(x), \quad x \in \partial C, \end{aligned}$$

for the correct choice of D.

Now, consider some other problems — our aim is to derive the correct analogue of the Dirichlet problem for more complicated payoff functions. Again, for all these examples, we shall assume that the process X_t is continuous, and there exists an optimal stopping time for the respective problem which is a.s. finite. Where necessary, we will assume any needed finiteness conditions — the arguments here should be seen as heuristic, rather than exact.

(i) Consider the problem:

$$V(x) = \sup_\tau \mathbb{E}_x \left[\mathrm{e}^{-\lambda_\tau} H(X_\tau) \right]$$

where $\lambda_t = \int_0^t \lambda(X_s)\, ds$, for a measurable function $\lambda(\cdot)$. We can account for this by considering instead the killed process $\tilde{X}_t$, killed at rate $\lambda(X_t)$, which in turn has generator $\mathbb{L}_{\tilde{X}} = \mathbb{L}_X - \lambda I$, to conclude that

$$V(x) = \sup_\tau \mathbb{E}_x \left[H(\tilde{X}_\tau) \right],$$

and so V satisfies the PDE

$$\begin{aligned} \mathbb{L}_X V(x) &= \lambda V, \quad x \in C \\ V(x) &= H(x), \quad x \in \partial C. \end{aligned}$$

(ii) Consider the finite horizon problem:

$$V(x,t) = \sup_{\tau \in \mathcal{M}_t^T} \mathbb{E}_{(x,t)} M(X_\tau)$$

where the appropriate Markov process is $Z_t = (X_t, t)$. Then we know $V(X_t, t)$ is a martingale for $(x,t) \in C \subseteq \mathbb{R} \times \mathbb{R}_+$. Assuming suitable smoothness on f, we can apply Itô to derive the operator of Z_t:

$$\begin{aligned} \mathbb{L}_Z f(x,s) &= \lim_{t \downarrow 0} \left\{ \frac{\mathbb{E}_{(x,s)} f(X_t, t+s) - f(x,s)}{t} \right\} \\ &= \lim_{t \downarrow 0} \left\{ \frac{\mathbb{E}_{(x,s)} \left[\int_0^t \mathbb{L}_X f(X_r, s+r)\, ds + \int_0^t \frac{\partial f}{\partial t}(X_r, s+r)\, dr \right]}{t} \right\} \\ &= \mathbb{L}_X f(x,s) + \frac{\partial f}{\partial t}(x,s). \end{aligned}$$

So the condition on V being a martingale now implies we get the PDE:

$$\begin{aligned} \mathbb{L}_X V(x) &= -\frac{\partial V}{\partial t}, \quad (x,t) \in C \\ V(x) &= M(x), \quad (x,t) \in \partial C. \end{aligned} \tag{39}$$

Note that this is the Cauchy problem in PDE theory. There is also killed version of this problem, where the right hand side of (39) becomes: $-\frac{\partial V}{\partial t} + \lambda V$.

(iii) Let L be a continuous function and consider

$$V(x) = \sup_\tau \mathbb{E}_x \left[\int_0^\tau L(X_s)\, ds \right].$$

Note that this problem does not immediately lend itself to a Markov setting, so we need to expand the state space: consider the process $Z_t = (X_t, L_t)$, where $L_t = l + \int_0^t L(X_s)\, ds$, and l is the starting point of the process L_t. Then for a function h, applying Itô as above, we see that the generator of the new process Z is given by

$$\mathbb{L}_Z h(x,l) = \mathbb{L}_X h + \frac{\partial h}{\partial l} L(x),$$

where the first term is the operator $\mathbb{L}_X$ applied only to the first co-ordinate of h. Now consider the more general problem

$$\tilde{V}(x,l) = \sup_\tau \mathbb{E}_x \left[l + \int_0^\tau L(X_s)\, ds \right].$$

Then $\frac{\partial \tilde{V}}{\partial l} = 1$, and $\tilde{V}(x,l) = V(x) + l$. So the condition $\mathcal{L}_Z \tilde{V}(x,l) = 0$ implies

$$\mathbb{L}_X V + \frac{\partial \tilde{V}}{\partial l} L(x) = 0$$

and therefore, we deduce that V solves the PDE

$$\begin{aligned} \mathbb{L}_X V(x) &= -L(x), \quad x \in C \\ V(x) &= 0, \quad x \in \partial C. \end{aligned} \tag{40}$$

Once again, there is a killed version, where the right hand side of (40) becomes $-L(x) + \lambda V$.

(iv) Consider $S_t = \max_{0 \le r \le t} X_r \vee S_0$, and let our underlying Markov process be $Z_t = (X_t, S_t)$ on $\tilde{E} = \{(x,s) : x \le s\}$, where $(X_0, S_0) = (x,s)$ under $\mathbb{P}_{x,s}$ for $(x,s) \in \tilde{E}$. We wish to consider a problem of the form:

$$V(x,s) = \sup_\tau \mathbb{E}_{x,s} \left[M(X_\tau, S_\tau) \right],$$

where $M(x,s)$ is a continuous function on $\tilde{E}$.

Now we again consider the generator of the process Z. If $x < s$, $\mathbb{L}_Z = \mathbb{L}_X$. Applying Itô at (s,s), since S_t is a finite variation process, we get

$$\begin{aligned}\mathbb{L}_Z f(s,s) &= \lim_{t\downarrow 0}\left\{\frac{\mathbb{E}_{(s,s)} f(X_t,S_t) - f(s,s)}{t}\right\}\\ &= \lim_{t\downarrow 0}\left\{\frac{\mathbb{E}_{(s,s)}\left[\int_0^t \mathbb{L}_X f(X_r,S_r)\,ds + \int_0^t \frac{\partial f}{\partial s}(X_r,S_r)\,dS_r\right]}{t}\right\}\\ &= \mathbb{L}_X f(s,s) + \frac{\partial f}{\partial s}(s,s)\lim_{t\downarrow 0}\left[\frac{\mathbb{E}_{s,s}(S_t - s)}{t}\right].\end{aligned}$$

Now, for small t, and a Brownian motion B_t started at 0, we have $\sup_{r\le t} B_t \sim \sqrt{t}$, and thus in the limit, $\frac{\mathbb{E}_{s,s}(S_t-s)}{t} = \infty$. In order for the generator to be well defined, the function f must therefore satisfy: $\frac{\partial f}{\partial s}(s,s) = 0$. Hence, $\mathbb{L}_Z V = 0$ implies:

$$\begin{aligned}\mathbb{L}_X V(x,s) &= 0, \quad (x,s)\in C\\ \frac{\partial V}{\partial s}(x,s) &= 0, \quad x = s\\ V(x,s) &= M(x,s), \quad (x,t)\in\partial C.\end{aligned} \tag{41}$$

Note that the set C is now a subset of $\tilde{E}$. Again, there is also a killed version, where an additional λV appears on the right hand side of (41).

4.3 A simple free boundary problem

We now consider how these results may become a free boundary problem, by considering a simple example. Suppose $X_t = x + \mu t + \sigma B_t$, for some parameters $\mu \ge 0, \sigma > 0$, and consider the function

$$G(x) = \begin{cases} 1 & : x \le 0\\ 0 & : x\in(0,1)\\ \alpha & : x \ge 1\end{cases}$$

where $\alpha \in (0,1)$. Then, for $\lambda > 0$ consider the problem:

$$V(x) = \sup_\tau \mathbb{E}_x\left[\mathrm{e}^{-\lambda t} G(X_\tau)\right].$$

Clearly, we cannot do any better than a reward of 1, so it will be optimal to stop immediately if we ever find ourselves below 0. Moreover, if it is optimal to stop at $x \ge 1$, it is clearly also optimal to stop at any other $y > x$, since these points will take even longer to get down to 0, and this is the only place where we can score higher than α. Hence the continuation region should be of the form: $(0,z)$ for some $z \ge 1$.

for such a region C, we can now use the previous results to compute the corresponding value function. Let V_z be the solution to:

$$\begin{aligned}\mathbb{L}_X V &= \lambda V, \quad x\in C\\ V(0) &= 1\\ V(z) &= \alpha\end{aligned}$$

Since

$$\mathbb{L}_X f = \frac{1}{2}\sigma^2 \frac{d^2 f}{dx^2} + \mu \frac{df}{dx},$$

$V_z(x)$ satisfies

$$\frac{1}{2}\sigma^2 \frac{d^2 V_z}{dx^2} + \mu \frac{dV_z}{dx} - \lambda V_z = 0$$

which we can solve to get

$$V_z(x) = \beta_1 \mathrm{e}^{d_1 x} + \beta_2 \mathrm{e}^{d_2 x}$$

where $d_1, d_2 = \frac{-\mu \pm \sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2}$. Applying the boundary condition at $x = 0$, we see that $\beta_1 = 1 - \beta_2$, and that

$$V_z(x) = \mathrm{e}^{d_1 x} + \beta_2(\mathrm{e}^{d_2 x} - \mathrm{e}^{d_1 x}).$$

If we introduce $\gamma = \frac{\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2}, \delta = -\frac{\mu}{\sigma^2}$, then $\gamma \geq -\delta \geq 0$, and we can rewrite the solution as

$$V_z(x) = \mathrm{e}^{\delta x}\left[\mathrm{e}^{\gamma x} - 2\beta_2 \sinh(\gamma x)\right].$$

Applying the second boundary condition, $V_z(z) = \alpha$, we see that

$$\beta_2 = \frac{1}{2}\frac{\mathrm{e}^{\gamma z} - \alpha \mathrm{e}^{-\delta z}}{\sinh(\gamma z)},$$

and therefore:

$$V_z(x) = \mathrm{e}^{\delta z}\left[\mathrm{e}^{\gamma x} - \frac{\sinh(\gamma x)}{\sinh(\gamma z)}(\mathrm{e}^{\gamma z} - \alpha \mathrm{e}^{-\delta z})\right].$$

Note that for $x \leq z, \mathrm{e}^{\gamma z}\sinh(\gamma x) \leq \mathrm{e}^{\gamma x}\sinh(\gamma z)$, and therefore the expression is positive.

If we now differentiate this expression in x, we get

$$V_z'(x) = \delta V_z(x) + \mathrm{e}^{\delta x}\left[\gamma \mathrm{e}^{\gamma x} - \gamma \frac{\cosh(\gamma x)}{\sinh(\gamma z)}(\mathrm{e}^{\gamma z} - \alpha \mathrm{e}^{-\delta z})\right]$$

and therefore at the right boundary,

$$V_z'(z) = \delta\alpha + \mathrm{e}^{\delta z}\left[\gamma \mathrm{e}^{\gamma z} - \gamma \frac{\cosh(\gamma z)}{\sinh(\gamma z)}(\mathrm{e}^{\gamma z} - \alpha \mathrm{e}^{-\delta z})\right].$$

If we differentiate this whole expression again with respect to z and simplify, we get:

$$\frac{d}{dz}\left[V_z'(z)\right] = \frac{\gamma}{\sinh^2(\gamma z)}\left[\gamma(\cosh((\delta + \gamma)z) - \alpha) - \delta \mathrm{e}^{-\delta z}\sinh(\gamma z)\right].$$

It follows that this expression is non-negative (since $\gamma \geq -\delta \geq 0$), and therefore that $V_z'(z)$ is increasing as a function of z.

This allows us to think of the solutions to the free boundary problem for given z as follows:

$V_1'(1) \leq 0$: In this case, since the gradient increases as z increases, there exists a unique $z^* \geq 1$ such that $V_{z^*}'(z^*) = 0$, and for all $z < z^*$, $V_z'(z) < 0$ and for all $z > z^*$, $V_z'(z) > 0$. This is represented in Figure 1.

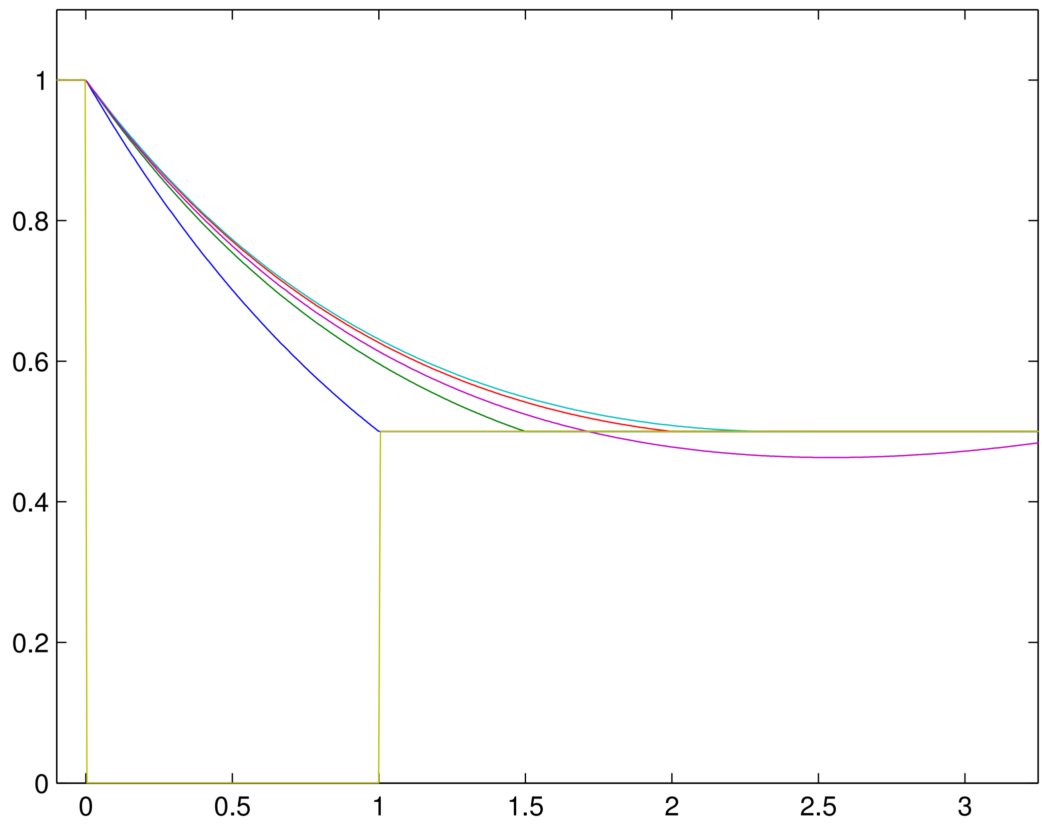

Figure 1: Possible functions $V_z(x)$ for different values of z. In this case, $V_1(1) < 0$. The parameter values are $\alpha = \frac{1}{2}, \mu = 0.25, \lambda = 0.1, \sigma = 1$. Note that there is a 'largest' function which has $V'_{z^*}(z^*) = 0$.

$V'_1(1) > 0$**:** In this case, again, since the gradient increases as z increases, $V'_z(z) > 0$ for all $z \geq 1$. This is represented in Figure 2.

In the first case, where there is a $z^* \geq 1$ such that $V'_{z^*}(z^*) = 0$, we can deduce that this must be the only plausible solution to (35)–(38) — any smaller choice of z gives a function $V_z(x)$ (extended to $\mathbb{R}_+$ with $V_z(x) = \alpha$ for $x > z$) which has a kink at z with left derivative which is negative and right derivative zero. Consequently, the value function is (strictly) convex at this point, and $V_z(X_s)$ would be a (strict) submartingale at this point, which is not possible. Conversely, if we choose a $z < z^*$, $V'_z(z) > 0$, and the function V_z would not remain above the gain function. As a result (see the argument following (35)–(38)), $V_{z^*}(x)$ is the value function. Note that in this case, z^* can be recovered by looking for the solution to the free-boundary problem: find V, z such that

$$
\begin{aligned}
\mathbb{L}_X V &= \lambda V, \quad x \in (0, z),\\
V(0) &= 1,\\
V(z) &= \alpha,\\
V'(z) &= 0.
\end{aligned}
$$

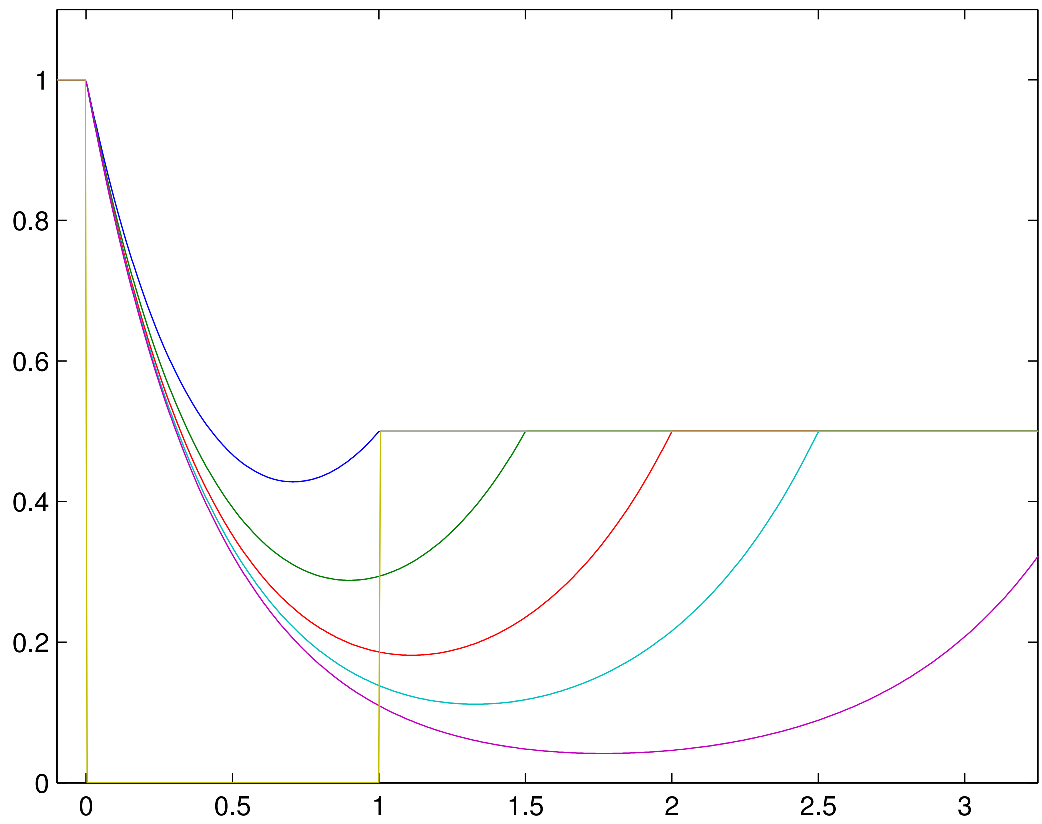

Figure 2: Possible functions $V_z(x)$ for different values of z. In this case, $V_1(1) > 0$. The parameter values are $\alpha = \frac{1}{2}, \mu = 0.25, \lambda = 2, \sigma = 1$. Note that there is no function which has $V'_{z^*}(z^*) = 0$.

The final condition in this case being the **smooth-fit condition**.

In the second case, where $V_1'(1) > 0$, we see again that we cannot take larger values of z since the resulting value function passes below the value function. It is easy to see however that (35)–(38) holds for $z^* = 1$, and this is therefore the optimal choice for the stopping region. In this case, there is no smooth-fit criterion. There are no hard-and-fast rules regarding when there is or is not a smooth-fit principle, but in this example, it is a consequence of the discontinuity in the gain function. In examples where the gain function is smoother, one would expect the smooth-fit principle to hold — see Figure 3, where a gain function $G(x) = \tanh(x)$ for $x > 0$ is used.

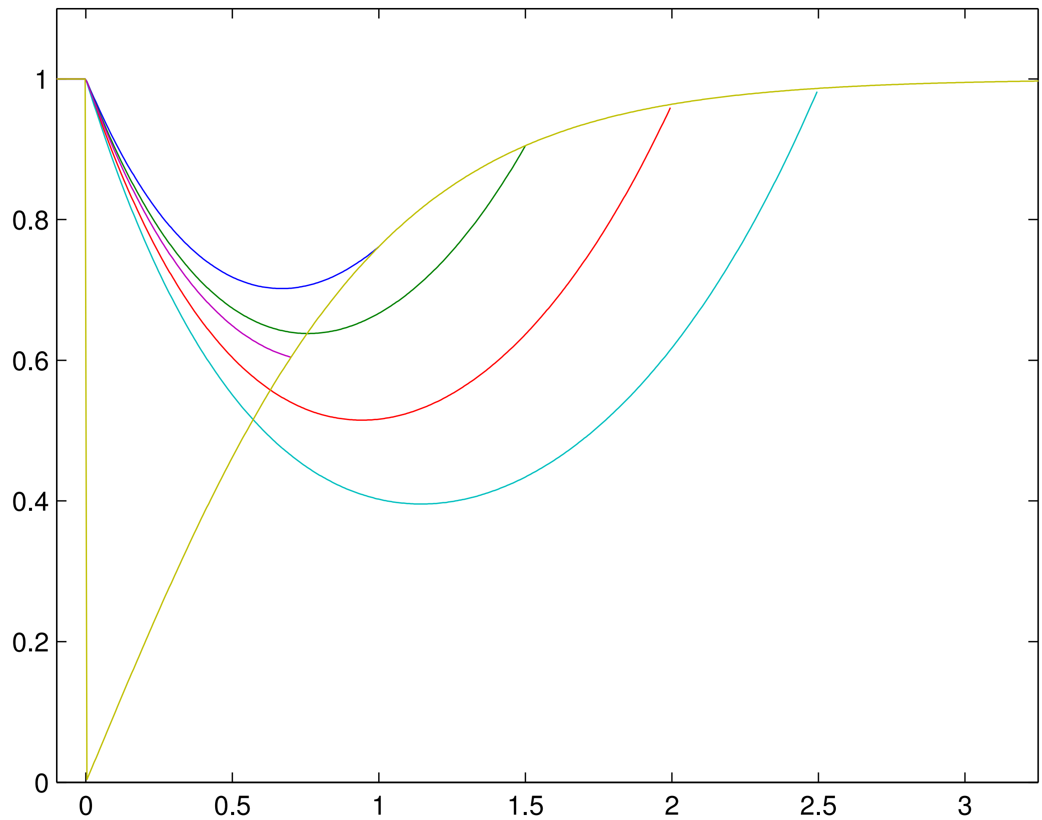

Figure 3: An alternative payoff function, again exhibiting similar behaviour as above, but this time a smooth-fit criterion can always be applied.