Optimal Stopping, Smooth Pasting and the Dual Problem

Saul Jacka and Dominic Norgilas, University of Warwick

Imperial

21 February 2018

The general optimal stopping problem:

Given a filtered probability space $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$ and an adapted gains process G, find

$$S_t \stackrel{def}{=} \operatorname*{ess\,sup}_{\text{optional } \tau \geq t} \mathbb{E}[G_\tau | \mathcal{F}_t]$$

Recall, under very general conditions

- S is the minimal supermartingale dominating G
- $\tau_t \stackrel{def}{=} \inf\{s \geq t : \; S_s = G_s\}$ is optimal
- for any t, S is a martingale on $[t, \tau_t]$
- when

$$G_t = g(X_t) \tag{1}$$

for some (continuous-time) Markov Process X, S_t can be written as a function, $v(X_t)$.

Remark

1.1 Condition $G_t = g(X_t)$ is less restrictive than might appear. With θ being the usual shift operator, can expand statespace of X by appending adapted functionals F with the property that

$$F_{t+s} = f(F_s, (\theta_s \circ X_u;\ 0 \le u \le t)). \tag{2}$$

The resulting process $Y \overset{def}{=} (X, F)$ is still Markovian. If X is strong Markov and F is right-cts then Y is strong Markov.

e.g if X is a BM,

$$Y_t = \left(X_t, L_t^0, \sup_{0 \le s \le t} X_s, \int_0^t \exp\left(-\int_0^s \alpha(X_u)du\right) g(X_s) \right) ds$$

is a Feller process on the filtration of X.

1. *When is v in the domain of the generator, $\mathcal{L}$, of X?* (Surprisingly, unable to find any general results about this.)
2. The dual problem is to find

$$V = \inf_{M \in \mathcal{M}_0} \mathbb{E}[\sup_t (G_t - M_t)],$$

   where $\mathcal{M}_0$ is the collection of uniformly integrable martingales started at 0.
   Is the dual of the Markovian problem a stoch. control problem for a controlled Markov Process?
3. The smooth pasting principle is used to find explicit solutions to optimal stopping problems essentially by "pasting together a martingale (on the continuation region) and the gains process (on the stopping region)"
   Can we say anything about smooth pasting?

Example: Suppose $X_t = (Z_t, L^0_t, t)$, where Z is a one-dimensional Ito diffusion and L^0 is its local time at 0. Let $g : (z, l, t) \mapsto -e^{-\alpha t} l$.

Claim:

$$v : (z, l, t) \mapsto -\psi(z) e^{-\alpha t} l,$$

where, defining τ_0 to be the first hitting time of 0 by Z, ψ is given by

$$\psi(z) \stackrel{def}{=} \mathbb{E}_z e^{-\alpha \tau_0} \quad (**)$$

Proof: first, note that V given by

$$V_t = v(Z_t, L^0_t, t)$$

is the conditional expected payoff obtained by stopping at the first hit of 0 by Z after time t. So $V \le S$. Conversely $V_{t \wedge \tau_0}$ is obviously a uniformly integrable martingale since ψ is bounded, satisfies $(**)$ and L^0 is continuous and only increases when Z is at 0.

Thus V is a supermartingale (since ψ is positive and L^0 is increasing). Consequently $V = S$ $\square$

Notice that ψ (and hence v) is not, in general C^1. For example, taking Z to be a BM and $\alpha = \frac{1}{2}$,

$$\psi : x \mapsto e^{-|x|}.$$

Thus smooth pasting may fail even if g is C^∞.

Define $\mathbb{G} = \left\{ \text{semimartingales such that } \mathbb{E}\left[\sup_{0\le t} |G_t| \right] < \infty \right\}$.

Theorem

If $G \in \mathbb{G}$ then

- *the Snell envelope S of G, admits a right-continuous modification and is the minimal supermartingale that dominates G.*
- *both G and S are class (D).*
- *G and S admit unique decompositions*

$$G = N + D, \qquad S = M - A \tag{3}$$

where $N \in \mathcal{M}_{0,loc}$ and D is a predictable finite-variation process, $M \in \mathcal{M}_0$, and A is a predictable, increasing process of integrable *variation (in IV).*

Remark

It is more normal to assume that the process A in the Doob-Meyer decomposition of S is started at zero. The dual problem is one reason why we do not do so here.

Recall that

$$H^1 = \{\text{special semimartingales } N+D \text{ where } \sup_t |N_t| + \int_0^\infty |dD_t| \in L^1\}.$$

The main assumption in this section is the following:

Assumption

3.1 G is in $\mathbb{G}$ and in H^1_{loc}.

Under Assumption 3.1, the previous theorem's conclusions hold and, in the decomposition $G = N + D$, D is a predictable IV_{loc} process.

We finally arrive to the main result:

Theorem

3.2 Suppose Assumption 3.1 holds. Let D^- (D^+) denote the decreasing (increasing) components of D. Then $A << D^-$, and μ, defined by

$$\mu_t := \frac{dA_t}{dD_t^-}, \quad 0 \leq t,$$

satisfies $0 \leq \mu_t \leq 1$.

Remark

As is usual in semimartingale calculus, we treat a process of bounded variation and its corresponding Lebesgue-Stiltjes signed measure as synonymous.

Proof First localise G and S so they are both in H^1. Recall the characterisation of a predictable IV process V: we have:

$$V_t - V_s = \lim_{\delta \downarrow 0} \sum_{i=0}^{\lfloor (t-s)/\delta \rfloor} \mathbb{E}[V_{s+(i+1)\delta} - V_{s+i\delta} | \mathcal{F}_{s+i\delta}], \qquad (4)$$

with limit being in L^1 (taking a subsequence[1] if necessary).

[1] this part may be why result is novel; result by Beiglböck et al. from 2012 is needed

Now, set

$$\Delta \stackrel{def}{=} \mathbb{E}[A_v - A_u|\mathcal{F}_u] = \mathbb{E}[S_u - S_v|\mathcal{F}_u]$$
$$= \mathbb{E}\left[\mathbb{E}[G_{\tau_u}|\mathcal{F}_u] - \operatorname*{ess\,sup}_{\sigma \geq v} \mathbb{E}[G_\sigma|\mathcal{F}_v]\middle|\mathcal{F}_u\right] \tag{5}$$

Taking $\sigma = \tau_u \vee v$ in (5), we obtain

$$\Delta \leq \mathbb{E}[G_{\tau_u} - G_{\tau_u \vee v}|\mathcal{F}_u] = \mathbb{E}[D_{\tau_u} - D_{\tau_u \vee v}|\mathcal{F}_u]$$
$$= \mathbb{E}[(D^+_{\tau_u} - D^+_{\tau_u \vee v}) + D^-_{\tau_u \vee v} - D^-_{\tau_u}|\mathcal{F}_u] \leq \mathbb{E}[D^-_{\tau_u \vee v} - D^-_{\tau_u}|\mathcal{F}_u]$$
$$\leq \mathbb{E}[D^-_v - D^-_u|\mathcal{F}_u]. \tag{6}$$

The last inequalities following since: D^+ and D^- are increasing; $\tau_u \geq u$; and, on the event that $\tau_u \geq v$, the term inside the penultimate expectation vanishes. Applying (4) to inequality (6) we get that $0 \leq A_t - A_s \leq D^-_t - D^-_s$ for all $s \leq t$, giving the result $\square$

Assumption

3.4 X is a right process with quasi-continuous filtration.

Remark

Note that if X satisfies the assumption then expanding the state by a right-continuous functional F of the form in Remark 1.1, (X, F) also satisfies Assumption 3.4. If X is Feller then it satisfies the assumption.

Finally,

Assumption

3.6 $\sup_t |g(X_t)| \in L^1$ *and* $g \in \mathbb{D}(\mathcal{L})$*, i.e.*

$$g(X_t) = g(x) + M_t^g + \int_0^t \mathcal{L}g(X_s)ds, \quad 0 \leq t, \, x \in E, \qquad (7)$$

so that G is a semimartingale and the FV process in the semimartingale decomposition of $G = g(X)$ is absolutely continuous with respect to Lebesgue measure, and therefore predictable. Moreover, we deduce that $g(X)$ satisfies Assumption 3.1.

The result of this section is the following:

Theorem

Suppose X and g satisfy Assumptions 3.4 and 3.6, then $v \in \mathbb{D}(\mathcal{L})$.

Proof Since $D_t := g(X_0) + \int_0^t \mathcal{L}g(X_s)ds, \quad 0 \leq t$, (ignoring initial values) D^+ and D^- are explicitly given by

$$D_t^+ := \int_0^t \mathcal{L}g(X_s)^+ ds,$$
$$D_t^- := \int_0^t \mathcal{L}g(X_s)^- ds,$$

so D^- is absolutely continuous with respect to Lebesgue measure.

Applying Theorem 3.2, we conclude that

$$v(X_t) = v(x) + M_t - \int_0^t \mu_s \mathcal{L}g(X_s)^- ds, \quad 0 \leq t, \tag{8}$$

where μ is a non-negative Radon-Nikodym derivative with $0 \leq \mu_s \leq 1$.

Setting $\lambda_t = \mu_t \mathcal{L}g(X_t)^-$, all that remains is to show that λ_t is $\sigma(X_t)$-measurable (since then there exists $\beta : E \to \mathbb{R}_+$, such that $\lambda = \beta(X)$).

This is fairly elementary (by the Markov property and quasi-continuity of the filtration) and thus $v \in \mathbb{D}(\mathcal{L})$.

$\square$

Recall that the dual problem is to find

$$V_0 = \inf_{m \in \mathcal{M}_0} \mathbb{E}[\sup_t (G_t - m_t)]. \tag{9}$$

To see this, first take $m = M$, the martingale in the decomposition

$$V_t = M_t - A_t.$$

Then $G_t - M_t = G_t - V_t - A_t \leq -A_t$ and A is an increasing process, so $\sup_t (G_t - M_t) = -A_0 = V_0$ and the RHS of (9) is at most V_0.

Conversely, for any stopping time τ, $\sup_t (G_t - m_t) \geq (G_\tau - m_\tau)$ and so

$$\mathbb{E}[\sup_t (G_t - m_t)] \geq \sup_\tau \mathbb{E}[(G_\tau - m_\tau)] = \sup_\tau \mathbb{E}[G_\tau] = V_0.$$

So the RHS of (9) is at least V_0.

We know that the optimal m is the martingale appearing in the decomposition of V. Since $v \in \mathbb{D}(\mathcal{L})$, this is $M_t^v \stackrel{def}{=} v(X_t) - v(X_0) - \int_0^t \mathcal{L}v(X_s)ds$. It follows that the dual problem is

$$V(x) = \inf_{h \in \mathbb{D}(\mathcal{L})} \mathbb{E}_x[\sup_t \Big(g(X_t) - (h(X_t) - \int_0^t \mathcal{L}h(X_s)ds) \Big)]$$

and a little thought shows that this is a controlled Markov process problem, with controlled MP Y^h given by $Y^h = (X, F^h)$ where

$$F_t^h = (\int_0^t \mathcal{L}h(X_s)ds, \sup_{s \le t} \Big(g(X_s) - h(X_s) + \int_0^s \mathcal{L}h(X_u)du \Big))$$

and cost function given by $\sup_s \Big(g(X_s) - h(X_s) + \int_0^s \mathcal{L}h(X_u)du \Big)$

We assume that X is a one-dimensional regular diffusion on E, a possibly infinite interval. Let $s(\cdot)$ denote a scale function of X.

Theorem

Suppose Assumption 3.6 holds, then $v \in \mathbb{D}(\mathcal{L})$. Let $Y = s(X)$ and let L^z_t denote its local time at z up to time t.

If

- $s \in \mathcal{C}^1$,
- $< Y, Y >_t$ *is absolutely continuous with respect to Lebesgue measure*

and

- *each $L^z_\cdot$ is either singular with respect to Lebesgue measure or identically zero,*

then $v(\cdot)$ is $\mathcal{C}^1$.

Proof Note that $Y = s(X)$ is a Markov process, and let $\mathcal{G}$ denote its martingale generator. Then $v(x) = W(s(x))$, where

$$W(y) = \sup_{\tau} \mathbb{E}_{s^{-1}(y)}[g \circ s^{-1}(Y_\tau)]. \tag{10}$$

Then, since $v \in \mathbb{D}(\mathcal{L})$,

$$v(X_t) = v(x) + M_t^v + \int_0^t \mathcal{L}v(X_s)ds,$$

and thus

$$W(Y_t) = W(y) + M_t^v + \int_0^t (\mathcal{L}v) \circ s^{-1}(Y_s)ds, \quad 0 \leq t.$$

Therefore, $W \in \mathbb{D}(\mathcal{G})$, i.e.

$$W(Y_t) = W(y) + M_t^W + \int_0^t \mathcal{G}W(Y_s)ds, \qquad (11)$$

with $\mathcal{G}W(\cdot) \leq 0$.

Y is a local martingale and so it's easy to show that $W(\cdot)$ is a concave function – in fact it's the least concave majorant of g. Using the generalised Ito formula we have

$$W(Y_t) = W(y) + \int_0^t W_-^{'}(Y_s)dY_s + \frac{1}{2}\int L_t^z \nu(dz), \qquad (12)$$

where L_t^z is the local time of Y at z, and ν is a non-negative, σ-finite measure corresponding to the derivative $W^{''}$ in the sense of distributions.

By the Lebesgue decomposition theorem, $\nu = \nu_c + \nu_s$, where ν_c and ν_s are measures, absolutely continuous and singular (with respect to Lebesgue measure in the spatial variable), respectively. Denoting the Radon-Nykodym derivative of ν_c by ν_c', the occupation time formula gives

$$W(Y_t) - W(y) = \int_0^t W'_-(Y_s)dY_s + \frac{1}{2}\int L_t^z \nu_c'(z)dz + \frac{1}{2}\int L_t^z \nu_s(dz)$$
$$= \int_0^t W'_-(Y_s)dY_s + \frac{1}{2}\int_0^t \nu_c'(Y_s)d < Y, Y >_s + \frac{1}{2}\int L_t^z \nu_s(dz). \tag{13}$$

By hypothesis, the quadratic variation process $(< Y, Y >_t)_{t \geq 0}$ is absolutely continuous with respect to Lebesgue measure.

Since Y is a continuous semimartingale, $L^z_\cdot$ is carried by the set $\{t : Y_t = z\}$ and is, by assumption, singular with respect to Lebesgue measure. We conclude that ν_s does not charge points, and therefore, since $\nu(z) = W'_+(z) - W'_-(z)$, left and right derivatives of $W(\cdot)$ must be equal. So $W \in \mathcal{C}^1$, and since $s \in \mathcal{C}^1$ by assumption and is strictly increasing, $v \in \mathcal{C}^1$. $\square$

Example

Suppose X is an Itô diffusion, i.e. X is a diffusion with infinitesimal generator (on $\mathcal{C}^2$)

$$\mathcal{L} = \frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}, \quad x \in E,$$

where $\sigma(\cdot)$ and $b(\cdot)$ are continuous functions and $\sigma(\cdot)$ does not vanish, and the endpoints of E are either inaccessible or absorbing. Then the scale function $s \in \mathcal{C}^2$, and since $< X, X >$ is absolutely continuous with respect to Lebesgue measure, so is $< s(X), s(X) >$. The singularity of the local times follows from that of BM. It follows that if g is C^2 then v is C^1.

Question: can we extend this to the discounted case?

Answer: yes, using a measure change argument.

Take ϕ and ψ: decreasing and increasing solutions to the expected discounted hitting time problem $(**)$.

This is most easily understood as a killed problem. So, if the diffusion is X then we kill X at rate α. Denote the killed diffusion by $\tilde{X}$, then define

$$\phi(x) = \begin{cases} \mathbb{P}_x(\tilde{X} \text{ hits } a) & \text{if } x > a \\ 1/\mathbb{P}_a(\tilde{X} \text{ hits } x) & \text{if } x \leq a \end{cases}$$

and

$$\phi(x) = \begin{cases} \mathbb{P}_x(\tilde{X} \text{ hits } a) & \text{if } x < a \\ 1/\mathbb{P}_a(\tilde{X} \text{ hits } x) & \text{if } x \geq a. \end{cases}$$

It is easy to see, by the usual conditional expectation arguments that $e^{-\alpha T}\phi(X_t)$ and $e^{-\alpha T}\psi(X_t)$ are strictly positive local martingales. Now define Λ by

$$\Lambda_t = e^{-\alpha T}\phi(X_t)/\phi(X_0),$$

then in the usual way (Girsanov), Λ defines a new (sub-)probability measure $\mathbb{Q}$ by $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = \Lambda_t$. Defining $s = \psi/\phi$ and Y by $Y_t = s(X_t)$, we see that Y is a $\mathbb{Q}-$local martingale.

Now

$$
\begin{aligned}
e^{-\alpha t} v(X_t) &= \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}_{\mathbb{P}}[e^{-\alpha\tau} g(X_\tau)|\mathcal{F}_t] \\
&= \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}_{\mathbb{P}}[\Lambda_\tau \frac{e^{-\alpha\tau} g(X_\tau)}{\Lambda_\tau}|\mathcal{F}_t] \\
&= \Lambda_t \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}_{\mathbb{Q}}[\frac{e^{-\alpha\tau} g(X_\tau)}{\Lambda_\tau}|\mathcal{F}_t] \\
&= \phi(X_0)\Lambda_t \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}_{\mathbb{Q}}[\frac{g}{\phi}(X_\tau)|\mathcal{F}_t] \\
&= \phi(X_0)\Lambda_t \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}_{\mathbb{Q}}[h(Y_\tau)|\mathcal{F}_t], \;\; (\text{where } h = \frac{g}{\phi} \circ s^{-1}) \\
&= e^{-\alpha t} \phi(X_t) W(Y_t),
\end{aligned}
$$

where W is the payoff (under $\mathbb{Q}$) for the optimal stopping problem for $h(Y.)$. It follows that W is concave and if Y (or X) satisfies the assumptions above and ϕ and ψ are C^1 then v is C^1.

Technical results used are all in Kallenberg, Protter, Revuz & Yor and Rogers and Williams.