Applications of Optimal Stopping and Stochastic Control

Saul Jacka

YRM
Warwick
15 April, 2011

WARWICK

Some problems

- The secretary problem
- Bayesian sequential hypothesis testing
- the multi-armed bandit problem (competing treatments in clinical trials)
- pricing American options
- best constants in inequalities between stochastic processes
- steering a diffusion to a goal or keeping a diffusion in a band
- finite fuel control problems
- optimal investment problems
- optimal coupling problems
- change point detection
- minimising shuttle time

This talk will focus on the Bellman approach.There are others, but none so general except the Girsanov change of measure approach.

(Sub- and super-)martingales and stopping times

X is a *martingale* if it represents a gambler's fortune when they play a fair gambling game. Thus even if their bet sizes will vary depending on the history of the game,

$$E[X_{t+s}|\mathcal{F}_t] = X_t,$$

for each $s, t \geq 0$.

X is a sub/supermartingale if the game is advantageous/disadvantageous, so

$$E[X_{t+s}|\mathcal{F}_t] \geq \text{ or } \leq X_t.$$

A *stopping time* is a random time whose occurrence is immediately detectable—thus the first time that X exceeds \$100 is a stopping time, but the last time it exceeds \$100 is NOT.

The key interaction between these concepts is in the Optional Sampling Theorem:

• (under suitable integrability conditions) these inequalities remain valid if t and $t+s$ are replaced by stopping times $\sigma \le \tau$. So, in the submartingale case

$$E[X_\tau|\mathcal{F}_\sigma] \ge X_\sigma.\text{ and, in particular } E[X_\tau] \ge X_0.$$

• if we add an increasing process to a martingale we get a submartingale and if we subtract we get a supermartingale.

Essential infimum and supremum: for any non-empty set $\mathcal{C}$, ess $\inf_{c \in \mathcal{C}} Y_c$ is defined as any r.v. L such that

- $L \leq Y_c$ a.s. for every $c \in \mathcal{C}$;
- if $Z \leq Y_c$ a.s. for every $c \in \mathcal{C}$, then $Z \leq L$ a.s.

ess sup is defined analagously. These are just the appropriate almost sure equivalents of sup and inf.

• Not to be confused with ess $\sup(X) \stackrel{def}{=} \inf\{t : P(X > t) = 0\}$

Ito's lemma and stochastic integrals: suppose

$$X_t = x + \underbrace{\int_0^t H_s dBs}_{\text{martingale under integrability condns.}} + \int_0^t \mu_s ds$$

where B is a 1-d Brownian motion. Further suppose that f is a (piecewise) $C^{2,1,1}$ function and Y is continuous and increasing. Then

$$\begin{aligned} f(X_t, Y_t, t) &= f(X_0, Y_0, 0) + \int_0^t f_x(X_s, Y_s, s) \overbrace{(H_s dB_s + \mu_s ds)}^{dX_s} \\ &+ \int_0^t [\frac{1}{2} f_{xx}(X_s, Y_s, s) H_s^2 + f_t(X_s, Y_s, s)] ds \\ &+ \int_0^t f_y(X_s, Y_s, s) dY_s \end{aligned}$$

• If f_x has some positive jumps then need to add an increasing process to RHS to correct this representation.

Optimal Stopping: The generic problem is as follows: given a stochastic process G,

- find $\sup_{\tau \in \mathcal{T}} E[G_\tau]$,

where $\mathcal{T}$ is the set of all stopping times.

Under suitable integrability assumptions we have the following result:
Theorem: Define

$$S_t = ess\ sup_{\tau \geq t} E[G_\tau | \mathcal{F}_t], \quad (*)$$

then S is the minimal supermartingale W such that $W_t \geq G_t$ a.s. for all t.

S is called the Snell envelope of the (gains) process G

Proof: Exercise:

- use the optional sampling theorem to show that any supermartingale dominating G dominates S.
- use increasing constraint on candidate stopping times to show S a supermartingale.

Q: How do we identify the Snell envelope in practice?

- guess the optimal stopping times in $(*)$, calculate corresponding payoffs and check that they give a supermartingale.
- In a continuous diffusion setting, we can use Ito's Lemma to help. In conjunction, use smooth pasting.
- In discrete and finite time we have following recursive characterisation:

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

This is just proved by backwards induction

Problem: change point detection

- $(Y_n)_{n\geq 1}$ are iid with known density f_∞
- $(Z_n)_{n\geq 1}$ are iid with known density f_0
- θ is a non-negative r.v. (the change point)– not generally observable.
- $X_n = Y_n 1_{(n\leq\theta)} + Z_n 1_{(\theta<n)}$.

We seek τ, the best estimate of θ under the following criteria

- the Average Detection Delay, $\mathrm{ADD}(\tau) \overset{def}{=} E[\tau - \theta | \tau > \theta]$ is minimised
- subject to the constraint that the Probability of False Alarm, $\mathrm{PFA}(\tau) \overset{def}{=} P(\tau \leq \theta) \leq \alpha$.

So we seek

$$\inf_{\mu \le \alpha,\ \tau:\ PFA(\tau)=\mu} \frac{E[(\tau-\theta)^+]}{1-\mu},$$

which gives rise to the Lagrangian

$$\inf_\tau E[(\tau-\theta)^+] + \lambda P(\tau \le \theta).$$

This is not yet an optimal stopping problem.

Trick: $E[(\tau-\theta)^+] = E[\sum_{m=1}^{\tau} P(\theta < m|\mathcal{F}_m)]$

Proof:

$$
\begin{aligned}
E[(\tau-\theta)^+] &= E[\sum_{m=1}^{\infty} 1_{\theta<m\le\tau}] \\
&= E[\sum_{m=1}^{\infty} P(\theta<m\le\tau|\mathcal{F}_m)] \\
&= E[\sum_{m=1}^{\infty} P(\theta<m|\mathcal{F}_m)1_{m\le\tau}] \\
&= E[\sum_{m=1}^{\tau} P(\theta<m|\mathcal{F}_m)]
\end{aligned}
$$

□

It follows that we seek to solve the optimal stopping problem with gains process

$$G_t = \sum_{m=1}^{t} P(\theta < m|\mathcal{F}_m) - \lambda P(\theta < t|\mathcal{F}_t).$$

- The general solution is unknown.
- The case where θ is geometric and independent of the Ys and Zs was solved by Shiryaev (see [6]). The optimum policy is to stop the first time $P(\theta < n|\mathcal{F}_n) \geq k$, for a suitable k.

Good Lambda Inequalities: Here's a simpler problem with the same sort of structure for the gains process:

$$G_t = 1_{(X_t \geq x \cap Y_t \leq y)} - \lambda 1_{(X_t > z)},$$

where X and Y are non-negative, continuous processes strictly increasing to ∞, $0 < z < x$ and y is positive.

- problem might as well be discrete: setting $T_w \stackrel{def}{=} \inf\{s : X_s \geq w\}$, the only times at which we might wish to stop are T_z and T_x (since G is 0 on $[0, T_z]$ and is constant on $(T_z, T_x]$ and decreasing after T_x).

Thus

$$S_t = \begin{cases} G_t & \text{for } t \geq T_x, \\ E[G_{T_x}|\mathcal{F}_t] = P(Y_{T_x} \leq y|\mathcal{F}_t) - \lambda & \text{for } T_z < t \leq T_x, \\ E[\max(0, P(Y_{T_x} \leq y|\mathcal{F}_{T_z}) - \lambda)|\mathcal{F}_t] & \text{for } t \leq T_z. \end{cases}$$

In particular, $S_0 = E[(P(Y_{T_x} \leq y|\mathcal{F}_{T_z}) - \lambda)^+]$ and so the best constant λ appearing in the inequality

$$P(X_\tau \geq x \cap Y_\tau \leq y) \leq \lambda P(X_\tau \leq z)$$

is ess sup$\big(P(Y_{T_x} \leq y|\mathcal{F}_{T_z})\big)$. See ([2]) for details and applications.

Stochastic Control.

Example:

- $X^{\mu}_t = x + B_t + \int_0^t \mu_s ds$, where B is a standard Brownian Motion
- we may choose the process μ under the constraint that $|\mu_t| \leq 1$ for all t.
- seek to minimise

$$E[\int_0^{\infty} e^{-\alpha t}(X^{\mu}_t)^2 dt].$$

.

We generalise as follows:

• *Dynamics*: For each $c \in \boldsymbol{C}$, a collection of control processes, X^c is a process taking values in some path space S.

• *Problem*: Our problem is to minimise $E[J(X^c, c)]$ for a given cost function $J : S \times \boldsymbol{C} \rightarrow \mathbb{R}$.

The Bellman principle

We'll need the following definitions:

Definition: For each $t \geq 0$ and each $c \in \boldsymbol{C}$, define

$$\boldsymbol{C}_t^c = \{d \in \boldsymbol{C} : d_s = c_s \text{ for all } s \leq t\}.$$

Definition For each $t \geq 0$ and each $c \in \boldsymbol{C}$, define

$$V_t^c = \text{ess inf}_{d \in \boldsymbol{C}_t^c} E[J(X^d, d)|\mathcal{F}_t].$$

Now suppose that we have a functional $L(\cdot,\cdot,\cdot)$ acting on the triples $((X^c_s)_{0\le s\le t},(c_s)_{0\le s\le t},t)$ with the following properties:

- $\overbrace{L^c_t}^{\text{constant over } \boldsymbol{C}^c_t} \stackrel{def}{=} L((X^c_s)_{0\le s\le t},(c_s)_{0\le s\le t},t) \stackrel{L^1}{\to} L^c_\infty \le J(X^c,c)$
  and
- $L^c_\cdot$ is a submartingale

then

$$L^c_t \le V^c_t.$$

Moreover, if, in addition, there exists $\hat{c}$ with $\hat{c}\in \boldsymbol{C}^c_t$ such that

- $(L^{\hat{c}}_{t+s})_{s\ge 0}$ is a martingale with $L^{\hat{c}}_\infty = J(X^{\hat{c}},\hat{c})$,

then

$$L^c_t = V^c_t.$$

Problem: Bayesian Sequential Hypothesis Testing: (see [4] for discrete case)

- $(X_n)_{n\geq 1}$ are iid r.v.s with common density f.
- Know that $f = f_0$ with prior probability p and $f = f_1$ with probability $1 - p$.
- At each time point t we may stop and declare that the density is f_0 or f_1 or we may pay c to sample one more of the Xs.
- If we declare density f_0 incorrectly, we lose L_0 and if we declare density f_1 incorrectly, we lose L_1.
- problem is to minimise our expected cost.

Dynamics A quick check shows that $p_t \stackrel{def}{=} P(f = f_0 | X_1, \dots, X_t)$ satisfies

$$p_t = \frac{p}{p + (1-p)\Lambda_t},$$

where

$$\Lambda_t = \prod_1^t \frac{f_1(X_s)}{f_0(X_s)}$$

is the Likelihood Ratio for the first t observations.

The performance functional J satisfies

$$J(X, \tau, D) = c\tau + p_\tau L_1 1_{(D=1)} + (1 - p_\tau) L_0 1_{(D=0)}.$$

where τ is the stopping time and D is the decision.

A randomisation argument shows that V_0 is a concave function of p which implies that the optimal strategy must be of the form

- stop and opt for f_1 $(D = 1)$ if $p_.$ has fallen below p_*,
- stop and opt for f_0 $(D = 0)$ if $p_.$ has risen above p^*,
- otherwise continue sampling.

Determination of p_* and p^* is, in general, an open problem.

• look at the continuous analogue, where X_t is a Brownian motion with, under f_0, no drift, and, under f_1, constant drift μ.
Exercise: The process p_t satisfies

$$dp_t = -\mu p_t(1 - p_t)dW_t,$$

where W is the (conditional) BM $X_t - \mu \int_0^t (1 - p_s)ds$.

Theorem:

$$V_t^{\tau,D} = W_t \overset{def}{=} \begin{cases} \psi(p_t) + ct & t < \tau, \\ p_\tau L_1 1_{(D=1)} + (1-p_\tau) L_0 1_{(D=0)} + c\tau & \tau \leq t \end{cases}$$

where

$$\psi(p) = \begin{cases} pL_1 & \text{for } p \leq a, \\ f(p) \overset{def}{=} K(2p-1)\ln(\frac{1-p}{p}) + C - Dp & \text{for } a < p < b, \\ (1-p)L_0 & \text{for } b \leq p, \end{cases}$$

and a, b, C and D are chosen so that

$$f(a) = aL_1,\ f'(a) = L_1,\ f(b) = (1-b)L_0,\ f'(b) = -L_0.$$

and $K = \frac{2c}{\mu^2}$. These values of a and b give p_* and p^*.

Proof: • Easy to check that $f'' = -\frac{K}{p^2(1-p)^2}$, so is strictly concave with $f'(0+) = \infty$ and $f'(1-) = -\infty$.

• so find unique values of C and D so that lines $y = L_1 p - C + Dp$ and $y = L_0(1-p) - C + Dp$ are tangential to graph of $K(2p-1)\ln(\frac{1-p}{p})$.

• tangency and concavity imply $f(p) \leq \min(L_0(1-p), (1-p)L_1)$.

• Now, ψ is C^1 and piecewise C^2 and by Ito's Lemma,

$$\begin{aligned}
dW_t &= 1_{t<\tau}(\psi'(p_t)dp_t + \frac{1}{2}\psi''(p_t)\mu^2 p_t^2(1-p_t)^2 dt + cdt) \\
&\quad +1_{t=\tau}(1_{D=0}((1-p_\tau)L_0 - \psi(p_\tau)) + 1_{D=1}(p_\tau L_1 - \psi(p_\tau)) \\
&= 1_{t<\tau}(\psi'(p_t)dp_t + c1_{p\notin(a,b)}dt) \\
&\quad +1_{t=\tau}(1_{D=0}((1-p_\tau)L_0 - \psi(p_\tau)) + 1_{D=1}(p_\tau L_1 - \psi(p_\tau))
\end{aligned}$$

which implies that W is a submartingale and is a martingale when the control is as above.

Problem: Shuttling in minimal time:

- problem is to control the drift of a Brownian motion (which reflects at 0 and 1) so as to minimise the time it takes to travel from 0 to 1 and back again. **Can only choose drift once at each level**. The problem models one arising in MCMC-with the level corresponding to temperature in a "heat bath".
- suppose that $X_t^\mu = x + B_t + \int_0^t \mu(X_s^\mu)ds$, for each μ.
- we seek a function μ to minimise $E_0[\tau_1] + E_1[\tau_0]$, where τ_z denotes the first hitting time of x and the subscript on the expectations denotes the starting point of X.

- There's no solution to the (dynamic) problem in the "classical" sense but there is an obvious guess for the problem as stated: $\mu \equiv 0$!
- To formulate problem, we allow ourselves to choose μ dynamically — but only once for each level. I.e., letting X^* denote the running supremum of the controlled process, we let

$$X_t^\mu = x + B_t + \int_0^t \mu_{\tau_{X_u^*}} du.$$

Reparameterize using the scale function s, so for each control μ, we define

- $s'_\mu(x) = \exp(-2 \int_0^x \mu_{\tau_z} dz)$,
- and define $s_\mu(x) = \int_0^x s'_\mu(u)du$ and $I_\mu(x) = \int_0^x \frac{du}{s'_\mu(u)}$.

Theorem: The optimal payoff process is given by

$$V_t^{\mu} = \phi(X_t, X_t^*, t),$$

where

$$\phi(x, y, t) = t + 2\left(\sqrt{s(y)I(y)} + (1 - y)\right)^2 - 2\int_{v=0}^{x}\int_{u=0}^{v} \frac{s'(v)}{s'(u)} du dv.$$

Proof:

- observation: if $a, b > 0$, then

$$\inf_{x>0} ax + \frac{b}{x} = 2\sqrt{ab} \text{ attained at } x = \sqrt{\frac{b}{a}}.$$

- Using Ito's lemma, see that

$$d\phi(X_t, X_t^*, t) = \frac{1}{2}\phi_{xx}dt + \phi_x dX_t + \phi_y dX_t^* + \phi_t dt.$$

Now

$$\begin{aligned} \phi_x &= -2s'(x)I(x) \\ \text{and}\phi_{xx} &= -2s''(x)I(x) - 2 = \frac{s''(x)}{s'(x)}\phi_x - 2 \\ &= -2\mu(x)\phi_x - 2, \end{aligned}$$

while

$$\phi_t = 1.$$

So

$$\begin{aligned}\frac{1}{2}\phi_{xx}dt &+ \phi_x dX_t + \phi_t dt \\ &= -(\mu(\cdot)\phi_x + 1)dt + \mu(\cdot)\phi_x dt + \mu(\cdot)\phi_x dB_t + dt \\ &= \mu(\cdot)\phi_x dB_t.\end{aligned}$$

Now, recalling that $I'(x) = \frac{1}{s'(x)}$,

$$\phi_y = 2\big(\sqrt{s(y)I(y)} + (1-y)\big) \times \Big(\frac{s'(y)I(y) + \frac{1}{s'(y)}s(y)}{2\sqrt{s(y)I(y)}} - 1\Big)$$

Using the observation, we see that $\phi_y \geq 0$ and

$$\phi_y = 0 \text{ if } s'(y) = \sqrt{\frac{s(y)}{I(y)}}.$$

The result follows, from Bellman's principle. $\square$

- in general, the corresponding control has a jump in s' so the optimal control will have a singular drift.
- However, if we control optimally from time 0 then we can calculate that s' is constant which corresponds to $\mu = 0$.
- same general form works for discounted time to shuttle
- minimising probability shuttle time exceeds T is open problem

- D Bertsekas (2005-07): Dynamic Programming and Optimal Control, Vol.s I and II, Athena Scientific
- S Jacka (1989): The good lambda inequalities revisited, *Sem. de Probab. XXIII*, 57–65.
- B Oksendal (2003): Stochastic Differential Equations with Applications (6th Edn.), Springer
- S Ross (1983): Introduction to Stochastic Dynamic Programming, Academic Press.
- A Shiryaev (1963): On optimum methods in quickest detection problems , *Theory of Prob and Appl.*, **8**, 2, 22-46.
- A G Tartakovsky and G V Moustakides (2010):State-of-the-Art in Bayesian Changepoint Detection, *Sequential Analysis*, **29**, 125–145.