Stopping with No Regret

Yu-Jui Huang
(Dublin City University)

Joint work with
Adrien Nguyen Huu (École des ponts ParisTech)
Traian Pirvu (McMaster University)

National Tsing Hua University
Hsinchu, Taiwan
17^{th} December 2014

Consider

- a Markovian process $X : [0,\infty) \times \Omega \mapsto \mathbb{R}^d$.
- a payoff function $g : \mathbb{R}^d \mapsto \mathbb{R}_+$.

Given initial time and state $(t,x) \in \mathbb{X} := [0,\infty) \times \mathbb{R}^d$, an individual faces an **optimal stopping problem**: can I find a $\tau \in \mathcal{T}_t$ such that my expected discounted payoff

$$J(t,x;\tau) := \mathbb{E}_{t,x}[\delta(t,\tau)g(X_\tau^{t,x})] \tag{1}$$

can be maximized?

The discount function $\delta(\cdot,\cdot)$ satisfies: for any $t \geq 0$, $\delta(t,t) = 1$,

$$s \mapsto \delta(t,s) \text{ is decreasingn}, \quad s \mapsto \delta(s,t) \text{ is increasing.}$$

The answer is affirmative by classical results.

Optimal Stopping Times [Karatzas & Shreve (1998)]

For all $(t,x) \in \mathbb{X}$, the stopping times

$$\widetilde{\tau}(t,x) := \inf\left\{ s \geq t \;:\; \delta(t,s)g(X_s^{t,x}) \geq \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_s} \mathbb{E}_{s,X_s^{t,x}}[\delta(t,\tau)g(X_\tau^{t,x})] \right\}, \quad (2)$$

$$\bar{\tau}(t,x) := \inf\left\{ s > t \;:\; \delta(t,s)g(X_s^{t,x}) \geq \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_s} \mathbb{E}_{s,X_s^{t,x}}[\delta(t,\tau)g(X_\tau^{t,x})] \right\}. \quad (3)$$

are optimal, i.e.

$$J(t,x;\widetilde{\tau}(t,x)) = \sup_{\tau \in \mathcal{T}_t} J(t,x;\tau).$$

Observe: $\widetilde{\tau}(t,x)$ is optimal **merely** from the perspective of the individual at time t.

Question: Will the individual regret $\widetilde{\tau}(t,x)$ in the future?

Analyze the choices of the individual over time.

1. At time t: she chooses $\widetilde{\tau}(t,x) \in \mathcal{T}_t$.
2. At time $t' > t$: if she has not stopped yet, $\widetilde{\tau}(t,x)$ now reads

$$\widetilde{\tau}(t,x) = \inf\left\{ s \geq t' \;:\; \delta(t,s)g(X_s^{t,x}) \geq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_s} \mathbb{E}_{s,X_s^{t,x}}[\delta(t,\tau)g(X_\tau^{t,x})] \right\}.$$

But now, the individual wants to maximize

$$\mathbb{E}_{t',X_{t'}^{t,x}}[\delta(t',\tau)g(X_\tau^{t,x})] \text{ over } \tau \in \mathcal{T}_{t'}.$$

Thus, her optimal stopping time is

$$\widetilde{\tau}(t',X_{t'}^{t,x}) = \inf\left\{ s \geq t' \;:\; \delta(t',s)g(X_s^{t,x}) \geq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_s} \mathbb{E}_{s,X_s^{t,x}}[\delta(t',\tau)g(X_\tau^{t,x})] \right\}.$$

At time t':

The individual **regrets** her previous choice $\widetilde{\tau}(t,x)$,
and **deviates** to $\widetilde{\tau}(t', X_{t'}^{t,x})$.

This phenomenon of **regret and deviation over time** is the so-called **time inconsistency**.

In classical literature of Mathematical Finance,

$$\delta(t,s) = e^{-\rho(s-t)} \quad \text{for some } \rho > 0.$$

This implies the identity

$$\delta(t,s)\delta(s,r) = \delta(t,r), \quad \forall\, 0 \le t \le s \le r.$$

Then we see that $\widetilde{\tau}(t,x) = \widetilde{\tau}(t', X_{t'}^{t,x})$ for all $t' \ge t$.
$\Rightarrow$ time inconsistency does not exist under exponential discounting.

Why not stay with exponential discounting?

- The payoff may not be monetary (utility, happiness, health,...).
- Numerous empirical studies show that people do not discount money like an exponential discount function.
  $\Rightarrow$ People admit "decreasing impatience"
  (Laibson (1997), O'Donoghue & Rabin (1999))

Stopping Policies

Since an individual may modify her choice of stopping times over time, her stopping strategy is a stopping policy defined below.

Definition (Stopping Policies)

We say $\tau : \mathbb{X} \mapsto \mathcal{T}_0$ is a **stopping policy** if for any $(t,x) \in \mathbb{X}$,

$$\tau(t,x) \in \mathcal{T}_t.$$

We denote by $\mathcal{T}(\mathbb{X})$ the collection of all stopping policies.

Idea behind: Given $(t,x) \in \mathbb{X}$ and $\tau \in \mathcal{T}(\mathbb{X})$,

- At each time s, we employ the stopping time $\tau(s, X_s^{t,x}) \in \mathcal{T}_s$.
- We follow $\tau(s, X_s^{t,x})$ *only* at the exact moment s, as it will be replaced by $\tau(u, X_u^{t,x}) \in \mathcal{T}_u$ as soon as $u > s$.
- $\tau(s, X_s^{t,x}) \in \mathcal{T}_s$ is used only to decide whether at time s we want to stop:

  if $\tau(s, X_s^{t,x}) = s$, we stop right away; otherwise we continue.

Given initial time and state $(t, x) \in \mathbb{X}$ and $\tau \in \mathcal{T}(\mathbb{X})$, we will eventually **stop at the moment**

$$\mathfrak{T}\tau(t,x) := \inf \left\{ s \geq t \ : \ \tau(s, X_s^{t,x}) = s \right\}. \qquad (4)$$

Goal: Formulate "no regret over time".

Idea: "no regret over time" is like an equilibrium!

Given initial time $t \geq 0$, consider a **leader-follower game** with

"Player t": yourself today "Player s": yourself at time $s \geq t$.

Fix $\tau \in \mathcal{T}(\mathbb{X})$.

- Suppose each "Player s" will employ $\tau(s, X_s^{t,x}) \in \mathcal{T}_s$. "Player t has to decide whether she wants to stick with $\tau(t,x)$.
- "Player t" has only two possible actions: to stop or to continue.
  1. If she stops, she gets $g(x)$ right away.
  2. If she continues, she will eventually stop at the time

$$\mathfrak{T}^*\tau(t,\omega) := \inf\left\{s > t \ : \ \tau(s, X_s^{t,x}) = s\right\}. \tag{5}$$

Her expected gain is therefore

$$\mathbb{E}_{t,x}\left[\delta(t, \mathfrak{T}^*\tau(t,x)) g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right].$$

- $g(x) > \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]$:
  She chooses to **stop** right away at time t.
- $g(x) < \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]$:
  She chooses to **continue** at time t.
  She will eventually stop at the time $\mathfrak{T}^*\tau(t,x)$.
- $g(x) = \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]$:
  She is **indifferent** between to stop and to continue at time t.
  $\Rightarrow$ no incentive to deviate from $\tau(t,x)$, so just follows $\tau(t,x)$.
  She will eventually stop at the time $\mathfrak{T}\tau(t,x)$.

The above can be summarized as

$$\Theta\tau(t,x) := t\ 1_{S_\tau(t,x)} + \mathfrak{T}\tau(t,x)1_{I_\tau(t,x)} + \mathfrak{T}^*\tau(t,x)1_{C_\tau(t,x)}, \quad (6)$$

where

$$\begin{aligned} S_\tau(t,x) &:= \{g(x) > \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]\},\\ I_\tau(t,x) &:= \{g(x) = \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]\}, \quad (7)\\ C_\tau(t,x) &:= \{g(x) < \mathbb{E}_{t,x}\left[\delta(t,\mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right]\}. \end{aligned}$$

the individual at time t would like to stick with $\tau(t,x)$ if and only if

$$\tau(t,x) = \Theta\tau(t,x).$$

Definition

We say $\tau \in \mathcal{T}(\mathbb{X})$ is an **equilibrium stopping policy** if

$$\Theta\tau(t,x) = \tau(t,x), \quad (t,x) \in \mathbb{X}.$$

Denote by $\mathcal{E}(\mathbb{X})$ the collection of all equilibrium stopping policies.

If $\tau \in \mathcal{T}(\mathbb{X})$ is an equilibrium policy, then no incentive to deviate from τ over time. $\Rightarrow$ This characterizes "no regret over time".

Note that:

- Several studies in stochastic **control** (Ekeland & Pirvu (2008), Ekeland et al (2012), Björk & Murgoci (2014),...) used equilibrium concept to formulate time-consistency.
- It has been unclear how to extend the equilibrium idea to **stopping** problems (see the discussion in Xu & Zhou (2013)).
- Our definition above *just* does this job!

Given initial time and state $(t, x) \in \mathbb{X}$, an individual wants to maximize her current expected payoff by choosing a $\tau \in \mathcal{E}(\mathbb{X})$.

$$\sup_{\tau \in \mathcal{E}(\mathbb{X})} \mathbb{E}_{t,x} \left[\delta(t, \mathfrak{T}\tau(t,x)) g(X^{t,x}_{\mathfrak{T}\tau(t,x)}) \right].$$

LEMMA

For any $\tau \in \mathcal{E}(\mathbb{X})$, $\mathfrak{T}\tau(t,x) = \tau(t,x)$ for all $(t,x) \in \mathbb{X}$.

Thus, the optimal time-consistent stopping problem reduces to

$$\sup_{\tau \in \mathcal{E}(\mathbb{X})} \mathbb{E}_{t,x} \left[\delta(t, \tau(t,x)) g(X^{t,x}_{\tau(t,x)}) \right].$$

$$\sup_{\tau \in \mathcal{E}(\mathbb{X})} \mathbb{E}_{t,x}\left[\delta(t, \tau(t,x)) g(X^{t,x}_{\tau(t,x)})\right]. \qquad (8)$$

- Is the problem **well-defined**? ($\mathcal{E}(\mathbb{X})$ nonempty?)

  Consider $\tau_{\text{tr}} \in \mathcal{T}(\mathbb{X})$ defined by

  $$\tau_{\text{tr}}(t,x) \equiv t \quad \text{for all } (t,x) \in \mathbb{X}.$$

  Then $\mathfrak{T}\tau_{\text{tr}}(t,x) = \mathfrak{T}^*\tau_{\text{tr}}(t,x) = t$ for all $(t,x) \in \mathbb{X}$.
  $\Rightarrow \tau_{\text{tr}}$ is an equilibrium policy.
  We call τ_{tr} the **trivial** equilibrium policy.

- Is the problem **non-trivial**? ($\mathcal{E}(\mathbb{X})$ in not a singleton?)
  If $\mathcal{E}(\mathbb{X}) = \{\tau_{\text{tr}}\}$, the problem (8) is trivial and boring.

  Question: Can we find **non-trivial** equilibrium policies?

Plan A: Consider an iterative procedure as follows:

1. At first, the individual wants to follow $\tau_1 := \widetilde{\tau}$. Then, she realizes that, at each moment t, her best stopping strategy is $\Theta\tau_1(t, x)$. She therefore switches from τ_1 to $\tau_2 := \Theta\tau_1$.
2. With the intention to follow τ_2, the individual realizes that, at each moment t, her best stopping strategy is $\Theta\tau_2(t, x)$. She therefore switches from τ_2 to $\tau_3 := \Theta\tau_2$.
3. The individual continues this procedure until

$$\Theta\tau_n = \tau_n \qquad \text{for some } n \text{ large enough.}$$

In short, we take

$$\tau_0 := \lim_{n\to\infty} \Theta^n \widetilde{\tau} \tag{9}$$

as a **candidate** equilibrium stopping policy.

- how to define the convergence in (11)?
- The limit τ_0 exist? If τ_0 exists, is it really an equilibrium policy?

Consider the operator $\Theta_{\mathrm{h}} : \mathcal{T}(\mathbb{X}) \mapsto \mathcal{T}(\mathbb{X})$ defined by

$$\Theta_{\mathrm{h}}\tau(t,x) := t\ 1_{S_\tau(t,x)\cup I_\tau(t,x)} + \mathfrak{T}^*\tau(t,x)1_{C_\tau(t,x)}, \quad \forall \tau \in \mathcal{T}(\mathbb{X}). \tag{10}$$

This describes an agent with "haste": she stops as soon as "to stop" is no worse than "to continue", i.e.

$$g(x) = \mathbb{E}_{t,x}\left[\delta(t, \mathfrak{T}^*\tau(t,x))g\left(X_{\mathfrak{T}^*\tau(t,x)}\right)\right].$$

Plan B: do the iteration in Plan A, but with Θ_{h}.

$$\tau_1^{\mathrm{h}} := \widetilde{\tau}, \quad \tau_2^{\mathrm{h}} := \Theta_{\mathrm{h}}\tau_1^{\mathrm{h}}, \quad \tau_3^{\mathrm{h}} := \Theta_{\mathrm{h}}\tau_2^{\mathrm{h}}, \quad \cdots$$

In short, we take

$$\tau_0^{\mathrm{h}} := \lim_{n\to\infty} \Theta_{\mathrm{h}}^n \widetilde{\tau} \tag{11}$$

as a **candidate** equilibrium stopping policy.

Assuming the discount function δ satisfies

$$\delta(t,s)\delta(s,r) \le \delta(t,r) \quad \forall\, 0 \le t \le s \le r. \tag{12}$$

Then, we have

Theorem

$\tau_0 = \lim_n \Theta^n \widetilde{\tau}$ and $\tau_0^{\mathsf{h}} = \lim_n \Theta_{\mathsf{h}}^n \widetilde{\tau}$ are equilibrium policies, i.e.

$$\Theta\tau_0 = \tau_0, \qquad \Theta_{\mathsf{h}}\tau_0^{\mathsf{h}} = \tau_0^{\mathsf{h}}.$$

Question: How strong is (12)?

Empirical studies indicate people has **"decreasin impatience"**, i.e.

People are more willing to wait (more patient),
when time horizon is longer.

Definition

The discount function $\delta(t,s) = h(s-t)$ induces **"decreasin impatience"** if for any $s \geq 0$,

$$\frac{h(t+s)}{h(t)} \text{ is increasing in } t.$$

The above condition obviously implies (12).

Conclusion: (12) already includes all discount functions with "decreasing impatience".

Consider

- $\{X_t\}_{t\geq 0}$: a one-dimensional Brownian motion
- Hyperbolic discount function

$$\delta(t,s) = \frac{1}{1+(s-t)}.$$

- payoff function as $g(x) = |x|$.

Want to compute explicitly

$$\tau_0 = \lim_{n\to\infty} \Theta^n \widetilde{\tau}, \qquad \tau_0^{\mathrm{h}} = \lim_{n\to\infty} \Theta_{\mathrm{h}}^n \widetilde{\tau}.$$

To solve for $\widetilde{\tau}$, the associated value function is

$$U(t,x) := \sup_{\tau\in\mathcal{T}_t} \mathbb{E}_{t,x}\left[\delta(t,\tau)g(X_\tau^{t,x})\right]. \tag{13}$$

Note that $\{\delta(0,t)U(t,X_t)\}_{t\geq 0}$ need not be a supermartingale, as $\delta(0,t)\delta(t,\tau)\neq\delta(0,\tau)$.

Auxiliary value function: for any fixed $t\geq 0$, we define

$$\begin{aligned} V(t,s,x) &:= \sup_{\tau\in\mathcal{T}_s} \mathbb{E}_{s,x}[\delta(t,\tau)g(X_\tau^{s,x})] \\ &= \sup_{\tau\in\mathcal{T}_s} \mathbb{E}_{s,x}\left[\frac{|X_\tau^{s,x}|}{1+(\tau-t)}\right], \quad \text{for } (s,x)\in[t,\infty)\times\mathbb{R}. \end{aligned}$$

We have $V(t,t,x)=U(t,x)$ and $\{V(t,s,X_s^{t,x})\}_{s\in[t,\infty)}$ is a supermartingale. $\Rightarrow$ the associated PDE: for $(s,x)\in[t,\infty)\times\mathbb{R}$,

$$\min\left\{v_s(t,s,x)+\frac{1}{2}v_{xx}(t,s,x),\ v(t,s,x)-\frac{|x|}{1+(s-t)}\right\}=0.$$

The above free-boundary PDE can be solved explicitly:

$$V(t,s,x) = \begin{cases} \frac{e^{-\frac{1}{2}}}{\sqrt{1+(s-t)}} \exp\left(\frac{x^2}{2(1+(s-t))}\right), & \text{for } |x| < \sqrt{1+(s-t)}, \\ \frac{|x|}{1+(s-t)}, & \text{for } |x| \geq \sqrt{1+(s-t)}, \end{cases}$$

It follows that

$$\widetilde{\tau}(t,x) := \inf\left\{s \geq t : |X_s^{t,x}| \geq \sqrt{1+(s-t)}\right\}.$$

The optimal stopping time $\widetilde{\tau}$ depends on initial time t
$\Rightarrow$ induces **time inconsistency**.

For any $(t, x) \in \mathbb{X}$,

$$\Theta\widetilde{\tau}(t,x) = t\ 1_{\{|x|>x^*\}} + \tau_1^{t,x}\ 1_{\{|x|\leq x^*\}}, \tag{14}$$

where $x^* \in (0,1)$ solves

$$\int_0^\infty e^{-s} \cosh(x^*\sqrt{2s})\, \mathrm{sech}(\sqrt{2s}) ds = x^* \ (x^* \approx 0.924),$$

and

$$\tau_a^{t,x} := \inf\{s \geq t : |X_s^{t,x}| \geq a\} \quad \text{for all } a \geq 0.$$

It is obvious that $\Theta\widetilde{\tau}(t,x) \neq \widetilde{\tau} \Rightarrow \widetilde{\tau}$ is not an equilibrium policy.

An Illuminating Example

Theorem

(I) $\tau_0 = \tau_{\text{tr}}$.

(II) $\tau_0^{\text{h}}(t,x) = t\ 1_{\{|x| \geq x^*\}} + \tau_{x^*}^{t,x}\ 1_{\{|x| < x^*\}}$.

$\tau_0^{\text{h}}(t,x)$ is indeed a non-trivial equilibrium policy.
Also note that

$$\mathbb{E}_{s,x}\left[\frac{\left|X^{s,x}_{\tau_0^{\text{h}}(t,x)}\right|}{1+(\tau_0^{\text{h}}(t,x) - t)}\right] \geq \mathbb{E}_{s,x}\left[\frac{\left|X^{s,x}_{\tau_{\text{tr}}}\right|}{1+(\tau_{\text{tr}} - t)}\right]$$