Algorithmic Trading: Optimal Control

Sebastian Jaimungal
University of Toronto

Jan, 2018

Stochastic Optimal Control

- **Stochastic Optimal Control** is concerned with **maximizing / minimizing** a **performance criteria** where the criteria is affected by **future unknown noise** in the system, as well as the **actions** of the controller / agent

Stochastic Optimal Control

- We aim to solve the problem

$$H(x) = \sup_{u \in \mathcal{A}} \mathbb{E}\left[\underbrace{G(X_T^{x,u})}_{\textbf{terminal reward}} + \underbrace{\int_0^T F(s,\, X_s^{x,u}, u_s)\, ds}_{\textbf{running reward/penalty}} \right]$$

where

- $u = (u_t)_{t \geq 0}$ is the **control process** and the agent chooses it
- $\mathcal{A}$ is the **admissible set** of controls – e.g., exclude doubling strategies
- $X^u = (X_t^u)_{t \geq 0}$ is the **controlled process**, which the agent partially controls, and satisfies the SDE

$$dX_t^u = \mu(t, X_t^u, u_t)\, dt + \sigma(t, X_t^u, u_t)\, dW_t\,, \qquad X_0 = x$$

Stochastic Optimal Control

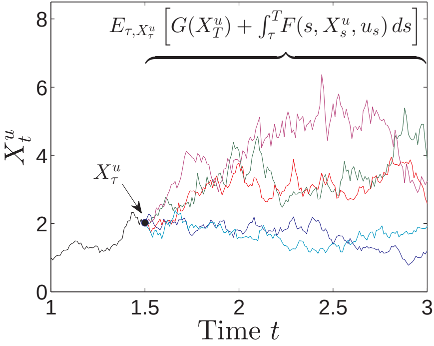
- Trick is to introduce a larger class of problems indexed by time
- The **value function** is defined as

$$H(t,x) = \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x}\left[G(X_T^u) + \int_t^T F(s,\, X_s^u, u_s)\, ds \right]$$

  where $\mathbb{E}_{t,x}[\cdot]$ means expectation conditional on $X_t = x$.
- Prove a **dynamic programming principle** and the corresponding **dynamic programming equation**

Stochastic Optimal Control

We will "flow" an arbitrary admissible control and re-write the value function recursively...



Stochastic Optimal Control

- Take an **arbitrary admissible control** $u \in \mathcal{A}$

$$
\begin{aligned}
&H^u(t,x) \\
&= \mathbb{E}_{t,x}\left[G(X_T^u) + \int_t^T F(s,\, X_s^u, u_s)\, ds \right] \\
&= \mathbb{E}_{t,x}\left[G(X_T^u) + \int_\tau^T F(s,\, X_s^u, u_s)\, ds + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right] \\
&= \mathbb{E}_{t,x}\left[\mathbb{E}_{\tau, X_\tau^u}\left[G(X_T^u) + \int_\tau^T F(s,\, X_s^u, u_s)\, ds \right] + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right] \\
&\qquad \text{(by iterated expectation)} \\
&= \mathbb{E}_{t,x}\left[H^u(\tau, X_\tau^u) + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right] \qquad \text{(by defn)}
\end{aligned}
$$

Stochastic Optimal Control

- However, $H(t,x) \geq H^u(t,x)$, with equality if $u = u^*$, hence

$$
\begin{aligned}
H^u(t,x) &\leq \mathbb{E}_{t,x}\left[H(\tau, X^u_\tau) + \int_t^\tau F(s,\, X^u_s, u_s)\, ds \right] \\
&\leq \sup_{u\in\mathcal{A}} \mathbb{E}_{t,x}\left[H(\tau, X^u_\tau) + \int_t^\tau F(s,\, X^u_s, u_s)\, ds \right]
\end{aligned}
$$

  and so

$$
\boxed{H(t,x) \leq \sup_{u\in\mathcal{A}} \mathbb{E}_{t,x}\left[H(\tau, X^u_\tau) + \int_t^\tau F(s,\, X^u_s, u_s)\, ds \right]}
$$

Stochastic Optimal Control

- Take an ε-**optimal control** υ^ε – a control that performs better than $H(t,x) - \varepsilon$, but of course not as good as $H(t,x)$, i.e., such that

$$H(t,x) \geq H^{\upsilon^\varepsilon}(t,x) \geq H(t,x) - \varepsilon$$

- Modify the ε-optimal control between t and τ by an arbitrary control u, i.e., define $\tilde{\upsilon}^\varepsilon$ by

$$\tilde{\upsilon}^\varepsilon_t := u_t \, \mathbb{1}_{t \leq \tau} + \upsilon^\varepsilon_t \, \mathbb{1}_{t > \tau}$$

Stochastic Optimal Control

- Then,

$$
\begin{aligned}
H(t,x) &\geq H^{\tilde{\upsilon}^\varepsilon}(t,x) \\
&= \mathbb{E}_{t,x}\left[H^{\tilde{\upsilon}^\varepsilon}(\tau, X_\tau^{\tilde{\upsilon}^\varepsilon}) + \int_t^\tau F(s,\, X_s^{\tilde{\upsilon}^\varepsilon}, \tilde{\upsilon}_s^\varepsilon)\, ds \right] && \text{by iterated expectations} \\
&= \mathbb{E}_{t,x}\left[H^{\upsilon^\varepsilon}(\tau, X_\tau^u) + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right] && \text{by the modified strategy} \\
&\geq \mathbb{E}_{t,x}\left[H(\tau, X_\tau^u) - \varepsilon + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right] && \varepsilon\text{-optimal control}
\end{aligned}
$$

- Hence, taking $\varepsilon \downarrow 0$, and since u is arbitrary

$$
\boxed{H(t,x) \geq \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x}\left[H(\tau, X_\tau^u) + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right]}
$$

Stochastic Optimal Control

- Putting both inequalities together we arrive at the **dynamic programming principle (DPP)**

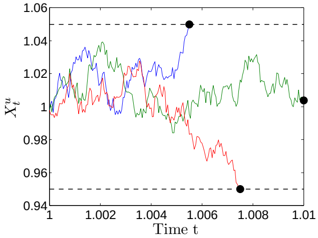
$$\boxed{H(t,x) = \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x}\left[H(\tau, X^u_\tau) + \int_t^\tau F(s,\, X^u_s, u_s)\, ds \right]}$$

- The **dynamic programming equation (DPE)** is the infinitesimal version of this principle

Stochastic Optimal Control

- Take τ in the DPP to be the following

$$\tau = T \wedge \inf\left\{s > t : (s - t,\ |X_s^u - x|) \notin [0, h) \times [0, \epsilon)\right\}.$$



- that is, either
  1. an amount of time h passes, and the processes has deviated by less than ϵ, or
  2. the process deviates by ϵ and we stop

Stochastic Optimal Control

- From the **DPP**

$$H(t,x) \geq \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x}\left[H(\tau, X_\tau^u) + \int_t^\tau F(s,\, X_s^u, u_s)\, ds \right]$$

- Hence, for a **constant strategy** υ on the interval $[t,\tau]$

$$H(t,x) \geq \mathbb{E}_{t,x}\left[H(\tau, X_\tau^\upsilon) + \int_t^\tau F(s,\, X_s^\upsilon, \upsilon)\, ds \right]$$

- Applying **Itô's lemma** to the value function

$$\begin{aligned} H(\tau, X_\tau^\upsilon) = H(t, X_t) &+ \int_t^\tau (\partial_t + \mathscr{L}_s^\upsilon)\, H(s, X_s^\upsilon)\, ds \\ &+ \int_t^\tau \partial_x H(s, X_s^\upsilon)\, \sigma(s, X_s^\upsilon, \upsilon)\, dW_s, \end{aligned}$$

where the generator is

$$\mathscr{L}_t^\upsilon = \mu(t,x,\upsilon)\,\partial_x + \tfrac{1}{2}\sigma^2(t,x,\upsilon)\,\partial_{xx}$$

Stochastic Optimal Control

- We then have,

$$H(t,x) \geq \mathbb{E}_{t,x}\Bigg[H(t,X_t) + \int_t^\tau \left(\partial_t + \mathscr{L}_s^\upsilon\right) H(s,X_s^\upsilon)\,ds + \int_t^\tau \partial_x H(s,X_s^\upsilon)\,\sigma(s,X_s^\upsilon,\upsilon)\,dW_s + \int_t^\tau F(s,\,X_s^\upsilon,\upsilon)\,ds\Bigg]$$

- Since $|X_t^\upsilon - x| < \epsilon$ on $[t,\tau]$, the stochastic integral is indeed a martingale and we have

$$H(t,x) \geq \mathbb{E}_{t,x}\left[H(t,x) + \int_t^\tau \left\{ \left(\partial_t + \mathscr{L}_s^\upsilon\right) H(s,X_s^\upsilon) + F(s,\,X_s^\upsilon,\upsilon)\right\} ds\right]$$

Stochastic Optimal Control

- Then,

$$0 \geq \lim_{h \downarrow 0} \mathbb{E}_{t,x} \left[\frac{1}{h} \int_t^{\tau} \left\{ (\partial_t + \mathscr{L}_s^{\upsilon}) \, H(s, X_s^{\upsilon}) + F(s, \, X_s^{\upsilon}, \upsilon) \right\} ds \right]$$
$$= (\partial_t + \mathscr{L}_t^{\upsilon}) \, H(t, x) + F(t, \, x, \upsilon)$$

which follows b/c

(i) as $h \searrow 0$, $\tau = t + h$ a.s. since the process will not hit the barrier of ϵ in extremely short periods of time,

(ii) the condition that $|X_\tau^u - x| \leq \epsilon$, which implies that if the process does hit the barrier it is bounded,

(iii) the Mean-Value Theorem allows us to write $\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \omega_s \, ds = \omega_t$, and

(iv) the process starts at $X_t^{\upsilon} = x$.

Stochastic Optimal Control

- This inequality holds for every constant v, and therefore

$$\boxed{\partial_t H(t,x) + \sup_{v} \{\mathscr{L}_t^v H(t,x) + F(t,\, x, v)\} \leq 0}$$

Stochastic Optimal Control

- Next, we show the opposite inequality, i.e., that

$$\partial_t H(t,x) + \sup_{\upsilon} \{\mathscr{L}_t^{\upsilon} H(t,x) + F(t,\, x, \upsilon)\} \geq 0$$

- We do this by **contradiction**... assume $\exists\,(t_0, x_0)$ s.t.,

$$\partial_t H(t_0,x_0) + \sup_{\upsilon} \{\mathscr{L}_t^{\upsilon} H(t_0,x_0) + F(t_0,\, x_0, \upsilon)\} < 0 \qquad (1)$$

- Then, define a modification φ of the value function H via

$$\varphi(t,x) = H(t,x) + \epsilon\left((t-t_0)^2 + (x-x_0)^4\right)$$

  this lies above the value function, but equals it at (t_0, x_0), and is differentiable enough to apply Itô's lemma

Stochastic Optimal Control

- Note also that

$$\begin{aligned}\partial_t \varphi(t_0, x_0) &= \partial_t H(t_0, x_0)\,,\\ \partial_x \varphi(t_0, x_0) &= \partial_x H(t_0, x_0)\,,\\ \partial_{xx} \varphi(t_0, x_0) &= \partial_{xx} H(t_0, x_0)\end{aligned}$$

- Therefore, from (1), we have that

$$\partial_t \varphi(t_0, x_0) + \sup_{\upsilon} \{\mathscr{L}_t^{\upsilon} \varphi(t_0, x_0) + F(t_0,\, x_0, \upsilon)\} < 0$$

  and if the Hamiltonian $\sup_{\upsilon} \{\mathscr{L}_t^{\upsilon} H(t, x) + F(t,\, x, \upsilon)\}$ is continuous, then

$$\exists \text{ a neighbourhood } \mathcal{N}_r = (t_0 - r, t_0 + r) \times (x_0 - r, x_0 + r)$$

  s.t.

$$\partial_t \varphi(t, x) + \sup_{\upsilon} \{\mathscr{L}_t^{\upsilon} \varphi(t, x) + F(t,\, x, \upsilon)\} < 0 \qquad (2)$$

  for all $(t, x) \in \mathcal{N}_r$

Stochastic Optimal Control

- Define
$$\eta = \max_{(t,x)\in\partial\mathcal{N}_r} (\varphi - H)(t,x) > 0$$
- Take an arbitrary control $u \in \mathcal{A}$ and define the stopping time
$$\tau = \inf\{s > t_0 \ : \ X_s^u \notin \mathcal{N}_r\}$$
- Since X is continuous
$$X_\tau^u \in \partial N_r$$
and therefore
$$\varphi(\tau, X_\tau^u) \geq \eta + H(\tau, X_\tau^u) \tag{3}$$

Stochastic Optimal Control

- Apply Itô's lemma to φ to find

$$\varphi(\tau, X^u_\tau) = \varphi(t_0, x_0) + \int_{t_0}^{\tau} (\partial_t + \mathscr{L}^u_t)\varphi(s, X^u_s)\, ds + \int_{t_0}^{\tau} \partial_x \varphi(s, X^u_s)\, \sigma(s, X^u_s, u)\, dW_s$$

- Therefore,

$$\begin{aligned} V(t_0, x_0) &= \varphi(t_0, x_0) \\ &= \mathbb{E}_{t_0, x_0}\left[\varphi(\tau, X^u_\tau) - \int_{t_0}^{\tau} (\partial_t + \mathscr{L}^u_t)\varphi(s, X^u_s)\, ds\right] \end{aligned}$$

- From (2), for $(t, x) \in \mathcal{N}_r$

$$\psi(t, x) = \partial_t \varphi(t, x) + \sup_{\upsilon} \left\{\mathscr{L}^{\upsilon}_t \varphi(t, x) + F(t,\, x, \upsilon)\right\} < 0$$

so that

$$\partial_t \varphi(t, x) + \mathscr{L}^u_t \varphi(t, x) + F(t,\, x, u_t) \leq \psi(t, x) < 0$$

Stochastic Optimal Control

- Therefore, we have

$$
\begin{aligned}
V(t_0, x_0) &= \varphi(t_0, x_0) \\
&= \mathbb{E}_{t_0,x_0}\left[\varphi(\tau, X_\tau^u) - \int_{t_0}^{\tau} (\partial_t + \mathscr{L}_t^u)\varphi(s, X_s^u)\, ds\right] \\
&\geq \mathbb{E}_{t_0,x_0}\left[\varphi(\tau, X_\tau^u) + \int_{t_0}^{\tau} \left(F(s, X_s^u, u_s) - \psi(s, X_s^u)\right)\, ds\right]
\end{aligned}
$$

- On boundary $\mathcal{N}_r$, φ dominates H, i.e. from (3), we have

$$
\begin{aligned}
V(t_0, x_0) &\geq \mathbb{E}_{t_0,x_0}\left[\eta + H(\tau, X_\tau^u) + \int_{t_0}^{\tau} \left(F(s, X_s^u, u_s) - \psi(s, X_s^u)\right)\, ds\right] \quad (\text{since } \psi < 0) \\
&\geq \eta + \mathbb{E}_{t_0,x_0}\left[H(\tau, X_\tau^u) + \int_{t_0}^{\tau} F(s, X_s^u, u_s)\, ds\right] \\
&> \mathbb{E}_{t_0,x_0}\left[H(\tau, X_\tau^u) + \int_{t_0}^{\tau} F(s, X_s^u, u_s)\, ds\right]
\end{aligned}
$$

This **violates the DPP**!

Stochastic Optimal Control

Hence we obtain the **Dynamic Programming Equation** (DPE), aka **Hamilton-Jacobi-Bellman** (HJB) equation

$$\boxed{\partial_t H(t,x) + \sup_{v} \{\mathscr{L}_t^v H(t,x) + F(t,\, x, v)\} = 0}$$

- This is a **non-linear PDE** that the **value function must satisfy**
- It is not clear that if we solve this PDE, the solution is the value function!
- This requires a **verification theorem**