Optimal Stopping with Neural Networks

Sebastian Becker
ZENAI

Patrick Cheridito
RiskLab, ETH Zurich

ETH Zurich, November 8, 2019

based on the the paper *Deep Optimal Stopping*, JMLR, 2019

The Problem

- Let $X_0, X_1, \dots, X_N$ be a d-dimensional Markov process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, *i.e.* $X_0, X_1, \dots, X_N \colon \Omega \to \mathbb{R}^d$ *are random vectors such that*

$$\mathbb{P}[X_{n+1} \in B \mid X_n] = \mathbb{P}[X_{n+1} \in B \mid X_0, \dots, X_n]$$

every random sequence can be made Markov by adding enough past information to the current state

- $g \colon \{0, 1, \dots, N\} \times \mathbb{R}^d \to \mathbb{R}$ a measurable function such that $\mathbb{E}|g(n, X_n)| < \infty$ for all n

- **Optimal Stopping Problem**

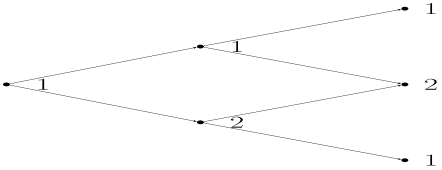
$$\sup_{\tau \in \mathcal{T}} \mathbb{E}\, g(\tau, X_\tau)$$

where $\mathcal{T}$ is the set of all X-stopping times $\tau \colon \Omega \to \{0, 1, \dots, N\}$

that is, $1_{\{\tau = n\}} = h_n(X_0, \dots, X_n)$ *for all* n

Example I

$$X_0 = 0, \quad \mathbb{P}[X_1 = \pm 1] = \frac{1}{2}, \quad \mathbb{P}[X_2 = X_1 \pm 1 \mid X_1] = \frac{1}{2}$$



$$\tau^* = \begin{cases} 2 & \text{if } X_1 = 1 \\ 1 & \text{if } X_1 = -1 \end{cases}$$

$$\mathbb{E}\, g(\tau^*, X_{\tau^*}) = \frac{1}{2} \times 2 + \frac{1}{4} \times 1 + \frac{1}{4} \times 2 = 1.75$$

General Solution: Snell Envelope

- The Snell envelope is recursively given by

$$\begin{aligned} H_N &:= g(N, X_N) \\ H_n &:= g(n, X_n) \vee \mathbb{E}[H_{n+1} \mid X_n] \end{aligned}$$

- It is the smallest $(\mathcal{F}_n^X)$-super-martingale such that $H_n \geq g(n, X_n)$ for all n,
  where $\mathcal{F}_n^X = \sigma(X_0, X_1, \ldots, X_n)$

  Proof:

  1) By definition, $H_n \geq \mathbb{E}[H_{n+1} \mid X_n] = \mathbb{E}[H_{n+1} \mid \mathcal{F}_n^X]$ and $H_n \geq g(n, X_n)$

  2) Every $(\mathcal{F}_n^X)$-super-martingale (Y_n) that dominates $g(n, X_n)$ satisfies
     $Y_n \geq \mathbb{E}[Y_{n+1} \mid \mathcal{F}_n^X]$ and $Y_n \geq g(n, X_n)$. So $Y_n \geq H_n$. □

- For every, X-stopping time τ,

$$\mathbb{E}\, g(\tau, X_\tau) \leq \mathbb{E}\, H_\tau \leq \mathbb{E}\, H_0$$

Optimal Solution

- Define

$$\begin{aligned}\tau_N^* &:= N\\ \tau_n^* &:= n1_{\{g(n,X_n)=H_n\}} + \tau_{n+1}^* 1_{\{g(n,X_n)<H_n\}}, \quad n \le N-1\end{aligned}$$

- Then, one verifies recursively for all $n \le N-1$,

$$\begin{aligned}\mathbb{E}\left[g\left(\tau_n^*, X_{\tau_n^*}\right) \mid X_n\right] &= \mathbb{E}\left[g\left(n, X_n\right) 1_{\{g(n,X_n)=H_n\}} + g\left(\tau_{n+1}^*, X_{\tau_{n+1}^*}\right) 1_{\{g(n,X_n)<H_n\}} \mid X_n\right]\\ &= g\left(n, X_n\right) 1_{\{g(n,X_n)=H_n\}} + \mathbb{E}\left[g\left(\tau_{n+1}^*, X_{\tau_{n+1}^*}\right) \mid X_n\right] 1_{\{g(n,X_n)<H_n\}}\\ &= g(n, X_n) \vee \mathbb{E}\left[H_{n+1} \mid X_n\right] = H_n\end{aligned}$$

- In particular, $\mathbb{E}\, g\left(\tau_0^*, X_{\tau_0^*}\right) = \mathbb{E}\, H_0 = \sup_{\tau \in \mathcal{T}} \mathbb{E}\, g\left(\tau, X_\tau\right)$ that is, τ_0^* is optimal
- In fact, $\mathbb{E}\, g\left(\tau_n^*, X_{\tau_n^*}\right) = \mathbb{E}\, H_n = V_n := \sup_{n \le \tau \le N} \mathbb{E}\, g\left(\tau, X_\tau\right)$ for all $n = 0, 1, \ldots, N$.

Dual Problem

Let $H_n = H_0 + M_n^H - A_n^H$ be the *Doob decomposition* with respect to $(\mathcal{F}_n^X)$, that is, (M_n^H) is an $(\mathcal{F}_n^X)$-martingale and (A_n^H) a non-decreasing $(\mathcal{F}_n^X)$-predictable process such that $M_0^H = A_0^H = 0$

Theorem (Rogers (2002), Haugh–Kogan (2004))

For every $(\mathcal{F}_n^X)$-martingale (M_n) with $M_0 = 0$, one has

$$\mathbb{E}\left[\max_{0\le n\le N}\{g(n,X_n) - M_n\}\right] \ge V_0 = \mathbb{E}\left[\max_{0\le n\le N}\{g(n,X_n) - M_n^H\}\right]$$

Proof: 1) For every X-stopping time τ,

$$\mathbb{E}\, g(\tau, X_\tau) = \mathbb{E}\left[g(\tau, X_\tau) - M_\tau\right] \le \mathbb{E}\left[\max_{0\le n\le N}\{g(n,X_n) - M_n\}\right].$$

$$\Rightarrow V_0 \le \mathbb{E}\left[\max_{0\le n\le N}\{g(n,X_n) - M_n\}\right].$$

2) $$\mathbb{E}\left[\max_{0\le n\le N}\{g(n,X_n) - M_n^H\}\right] \le \mathbb{E}\left[\max_{0\le n\le N}\{H_n - M_n^H\}\right] = \mathbb{E}\left[\max_{0\le n\le N}\{H_0 - A_n^H\}\right] = \mathbb{E}\, H_0 = V_0$$

□

Stopping Times and Stopping Decisions

- Denote by $\mathcal{T}_n$ the set of all X-stopping times τ such that $n \leq \tau \leq N$
- Let $f_n, f_{n+1}, \ldots, f_N : \mathbb{R}^d \to \{0,1\}$ be measurable functions such that $f_N \equiv 1$. Then

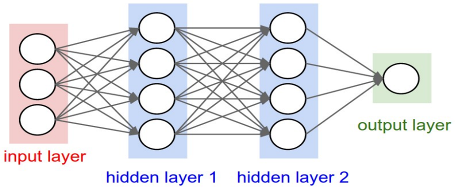
$$\tau_n = \sum_{m=n}^{N} m f_m(X_m) \prod_{j=n}^{m-1} (1 - f_j(X_j)) \qquad \text{with} \quad \prod_{j=n}^{n-1} (1 - f_j(X_j)) := 1$$

  is a stopping time in $\mathcal{T}_n$

-

$$\tau_n = n f_n(X_n) + \tau_{n+1}(1 - f_n(X_n)), \quad \text{where} \quad \tau_{n+1} = \sum_{m=n+1}^{N} m f_m(X_m) \prod_{j=n+1}^{m-1} (1 - f_j(X_j)).$$

Neural Network Approximation



Idea Recursively approximate f_n by a neural network $f^\theta \colon \mathbb{R}^d \to \{0, 1\}$ of the form

$$f^\theta = 1_{[0,\infty)} \circ a_3^\theta \circ \varphi_{q_2} \circ a_2^\theta \circ \varphi_{q_1} \circ a_1^\theta,$$

where

- q_1 and q_2 are positive integers specifying the number of nodes in the two hidden layers,
- $a_1^\theta \colon \mathbb{R}^d \to \mathbb{R}^{q_1}$, $a_2^\theta \colon \mathbb{R}^{q_1} \to \mathbb{R}^{q_2}$ and $a_3^\theta \colon \mathbb{R}^{q_2} \to \mathbb{R}$ are affine functions given by
$$a_i^\theta(x) = A_i x + b_i, \; i = 1, 2, 3,$$
- for $j \in \mathbb{N}$, $\varphi_j \colon \mathbb{R}^j \to \mathbb{R}^j$ is the component-wise ReLU activation function given by $\varphi_j(x_1, \ldots, x_j) = (x_1^+, \ldots, x_j^+)$

The components of θ consist of the entries of A_i and b_i, $i = 1, 2, 3$

More precisely,

- assume parameter values $\theta_{n+1}, \theta_{n+2}, \ldots, \theta_N \in \mathbb{R}^q$ have been found such that $f^{\theta_N} \equiv 1$ and the stopping time

$$\tau_{n+1} = \sum_{m=n+1}^{N} m f^{\theta_m}(X_m) \prod_{j=n+1}^{m-1} (1 - f^{\theta_j}(X_j))$$

  produces an expectation $\mathbb{E}\, g(\tau_{n+1}, X_{\tau_{n+1}})$ close to the optimal value V_{n+1}

- now try to find a maximizer $\theta_n \in \mathbb{R}^q$ of

$$\theta \mapsto \mathbb{E}\left[g(n, X_n) f^{\theta}(X_n) + g(\tau_{n+1}, X_{n+1})(1 - f^{\theta}(X_n))\right]$$

- **Goal** find an (approximately) optimal $\theta_n \in \mathbb{R}^q$ with a stochastic gradient ascent method
- **Problem** for $x \in \mathbb{R}^d$, the θ-gradient of

$$f^\theta(x) = 1_{[0,\infty)} \circ a_3^\theta \circ \varphi_{q_2} \circ a_2^\theta \circ \varphi_{q_1} \circ a_1^\theta(x)$$

  is 0 or does not exist

- As an **intermediate step** consider a neural network $F^\theta \colon \mathbb{R}^d \to (0,1)$ of the form

$$F^\theta = \psi \circ a_3^\theta \circ \varphi_{q_2} \circ a_2^\theta \circ \varphi_{q_1} \circ a_1^\theta \quad \text{for} \quad \psi(x) = \frac{e^x}{1+e^x}$$

- Use **stochastic gradient ascent** to find an approximate optimizer $\theta_n \in \mathbb{R}^q$ of

$$\theta \mapsto \mathbb{E}\left[g(n, X_n) F^\theta(X_n) + g(\tau_{n+1}, X_{\tau_{n+1}})(1 - F^\theta(X_n))\right]$$

- **Approximate** $f_n \approx f^{\theta_n} = 1_{[0,\infty)} \circ a_3^{\theta_n} \circ \varphi_{q_2} \circ a_2^{\theta_n} \circ \varphi_{q_1} \circ a_1^{\theta_n}$
- **Repeat the same steps** at times $n-1, n-2, \ldots, 0$

Training the Networks

- Let $(x_n^k)_{n=0}^N$, $k = 1, 2, \ldots$ be independent simulations of $(X_n)_{n=0}^N$
- Let $\theta_{n+1}, \ldots, \theta_N \in \mathbb{R}^q$ be given, and consider the corresponding stopping time

$$\tau_{n+1} = \sum_{m=n+1}^{N} m f^{\theta_m}(X_m) \prod_{j=n+1}^{m-1} (1 - f^{\theta_j}(X_j))$$

- τ_{n+1} is of the form $\tau_{n+1} = l_{n+1}(X_{n+1}, \ldots, X_{N-1})$ for a measurable function

$$l_{n+1} : \mathbb{R}^{d(N-n-1)} \to \{n+1, n+2, \ldots, N\}$$

- Denote

$$l_{n+1}^k = \begin{cases} N & \text{if } n = N-1 \\ l_{n+1}(x_{n+1}^k, \ldots, x_{N-1}^k) & \text{if } n \le N-2 \end{cases}$$

- The realized reward

$$r_n^k(\theta) = g(n, x_n^k) F^\theta(x_n^k) + g(l_{n+1}^k, x_{l_{n+1}^k}^k)(1 - F^\theta(x_n^k))$$

is continuous and almost everywhere differentiable in θ

Stochastic Gradient Ascent

- **Initialize** $\theta_{n,0}$ typically random; e.g. Xavier initialization
- **Standard updating** $\theta_{n,k+1} = \theta_{n,k} + \eta \nabla r_n^k(\theta_{n,k})$
- **Variants**
  - Mini-batches
  - Batch normalization
  - Momentum
  - Adagrad
  - RMSProp
  - AdaDelta
  - ADAM
  - Decoupling weight decay
  - Warm restarts
  - ...

Stochastic Gradient Ascent

- **Initialize** $\theta_{n,0}$ typically random; e.g. Xavier initialization
- **Standard updating** $\theta_{n,m+1} = \theta_{n,m} + \eta \nabla r_n^m(\theta_{n,m})$
- **Variants**
  - Mini-batches
  - Batch normalization
  - Momentum
  - Adagrad
  - RMSProp
  - AdaDelta
  - ADAM
  - Decoupling weight decay
  - Warm restarts
  - ...

Lower bound

- The candidate optimal stopping time

$$\tau^{\Theta} = \sum_{n=1}^{N} n f^{\theta_n}(X_n) \prod_{j=0}^{n-1} (1 - f^{\theta_j}(X_j))$$

  yields a lower bound

$$L = \mathbb{E}\, g(\tau^{\Theta}, X_{\tau^{\Theta}}) \quad \text{for the optimal value} \quad V_0 = \sup_{\tau} \mathbb{E}\, g(\tau, X_{\tau})$$

- Let $(y_n^k)_{n=0}^N$, $k = 1, 2, \ldots, K$, be a new set of independent simulations of $(X_n)_{n=0}^N$
- τ^{Θ} can be written as $\tau^{\Theta} = l(X_0, \ldots, X_{N-1})$ for a measurable function $l : \mathbb{R}^{dN} \to \{0, 1, \ldots, N\}$
- Denote $l^k = l(y_0^k, \ldots, y_{N-1}^k)$
- Use the Monte Carlo approximation

$$\hat{L} = \frac{1}{K} \sum_{k=1}^{K} g(l^k, y_{l^k}^k) \quad \text{as an estimate for} \quad L$$

Upper Bound and Confidence Intervals

Analogously, one can estimate

- An upper bound $\hat{U}$ by solving the dual problem and a
- 95% confidence interval $[\hat{L} - \Delta_L \, , \, \hat{U} + \Delta_U]$

Example II

1. **Bermudan max-call options**

   Consider d assets with prices evolving according to a multi-dimensional Black–Scholes model

$$S_t^i = s_0^i \exp\left([r - \delta_i - \sigma_i^2/2]t + \sigma_i W_t^i \right), \quad i = 1, 2, \ldots, d,$$

for

- initial values $s_0^i \in (0, \infty)$
- a risk-free interest rate $r \in \mathbb{R}$
- dividend yields $\delta_i \in [0, \infty)$
- volatilities $\sigma_i \in (0, \infty)$
- and a d-dimensional Brownian motion W with constant correlation ρ_{ij} between increments of different components W^i and W^j

A Bermudan max-call option has time-t payoff $\left(\max_{1 \le i \le d} S_t^i - K\right)^+$ and can be exercised at one of finitely many times $0 = t_0 < t_1 = \frac{T}{N} < t_2 = \frac{2T}{N} < \cdots < t_N = T$

$$\text{Price:} \qquad \sup_{\tau \in \{t_0, t_1, \ldots, T\}} \mathbb{E}\left[e^{-r\tau} \left(\max_{1 \le i \le d} S_\tau^i - K \right)^+ \right] = \sup_{\tau \in \mathcal{T}} \mathbb{E}\, g(\tau, X_\tau)$$

Numerical results

for $s_0^i = 100$, $\sigma_i = 20\%$, $r = 5\%$, $\delta = 10\%$, $\rho_{ij} = 0$, $K = 100$, $T = 3$, $N = 9$:

# Assets	Point Est.	Comp. Time	95% Conf. Int.	Bin. Tree	Broadie–Cao 95% Conf. Int.
2	13.899	28.7s	[13.880, 13.910]	13.902	
3	18.690	28.9s	[18.673, 18.699]	18.69	
5	26.159	28.1s	[26.138, 26.174]		[26.115, 26.164]
10	38.337	30.5s	[38.300, 38.367]		
20	51.668	37.5s	[51.549, 51.803]		
30	59.659	45.5s	[59.476, 59.872]		
50	69.736	59.1s	[69.560, 69.945]		
100	83.584	95.9s	[83.357, 83.862]		
200	97.612	170.1s	[97.381, 97.889]		
500	116.425	493.5s	[116.210, 116.685]		