Unbiased Optimal Stopping via the MUSE

Guanyang Wang (Rutgers Stats)

joint work with

Zhengqing Zhou Stanford Math/SIG

Jose Blanchet Stanford MS&E

Peter Glynn Stanford MS&E

Department Seminar
University of Minnesota
Department of Statistics
October 7, 2021

Outline

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes of maximizing the utility/minimizing the cost.

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.

Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.

Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.
Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?
  - If the first roll ends up with 6. Then ...

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.

Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?
  - If the first roll ends up with 6. Then ... you want to stop.
  - If the first roll ends up with 2. Then ...

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.

Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?
  - If the first roll ends up with 6. Then ... you want to stop.
  - If the first roll ends up with 2. Then ... you want to continue.

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.

Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?
  - If the first roll ends up with 6. Then ... you want to stop.
  - If the first roll ends up with 2. Then ... you want to continue.
- Quick calculation:
  - Expectation of tossing a dice once equals $\sum_{k=1}^{6} k/6 = 3.5$
  - The optimal stopping rule will be: stop when you get $4, 5, 6$, continue when you get $1, 2, 3$.
  - The expected utility of the optimal strategy is $U_2 = 4.25$.

What is an optimal stopping problem?

Informal: Choose a time to take an action, in the hopes to maximizing the utility/minimizing the cost.
Example (Dice Game):

- You are playing a dice gambling game. You can roll the dice up to twice and the reward will be the last roll.
- Question: What is the optimal strategy?
  - If the first roll ends up with 6. Then ... you want to stop.
  - If the first roll ends up with 2. Then ... you want to continue.
- Quick calculation:
  - Expectation of tossing a dice once equals $\sum_{k=1}^{6} k/6 = 3.5$
  - The optimal stopping rule will be: stop when you get $4, 5, 6$, continue when you get $1, 2, 3$.
  - The expected utility of the optimal strategy is $U_2 = 4.25$.
- Follow up question: What if you can roll T times?
  - Dynamic programming solution: $U_T = \mathbb{E}[\max\{X_1, U_{T-1}\}]$.

What is an optimal stopping problem?

Example (American put option):

- Suppose there is a stock which price X_t follows a Geometric Brownion motion, i.e., $\mathrm{d}X_t = \mu X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}W_t, t \in [0, T]$.
- Buying a put option at strike price K means you earn money when the stock price is lower than K. The utility function is $f(x) := \max\{K - x, 0\}$

What is an optimal stopping problem?

Example (American put option): Suppose there is a stock which price X_t follows a Geometric Brownion motion, i.e.,
$\mathrm{d}X_t = \mu X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}W_t, t \in [0, T]$. Buying a put option at strike price K means you earn money when the stock price is lower than K. The utility function is $f(x) := \max\{K - x, 0\}$

- American option allows 'early exercise'.

What is an optimal stopping problem?

Example (American put option): Suppose there is a stock which price X_t follows a Geometric Brownion motion, i.e., $\mathrm{d}X_t = \mu X_t \,\mathrm{d}t + \sigma X_t \,\mathrm{d}W_t, t \in [0, T]$. Buying a put option at strike price K means you earn money when the stock price is lower than K. The utility function is $f(x) := \max\{K - x, 0\}$.

- American option allows 'early exercise'.
- The expected utility is: $U_T = \sup_\tau \mathbb{E}[f(X_\tau)]$.
- Can still be solved using the dynamic programming idea numerically.

Finite-horizon optimal stopping problem

Given an underlying process $(X_1, \cdots, X_T) \sim \pi$ and a reward function f. We are interested in computing the expected utility of the optimal strategy:

$$U_T := \sup_{\tau \in \mathcal{T}_T} \mathbb{E}\left[f\left(X_\tau\right)\right], \tag{1}$$

where $\mathcal{T}_T$ is the set of all stopping times taking values in $\{1, 2, \cdots, T\}$. A stopping time is a random variable τ with $\{\tau \leq k\} \in \sigma(X_1, \cdots, X_k)$ for every k.

Finite-horizon optimal stopping problem

Given an underlying process $(X_1, \cdots, X_T) \sim \pi$ and a reward function f. We are interested in computing the expected utility of the optimal strategy:

$$U_T := \sup_{\tau \in \mathcal{T}_T} \mathbb{E}\left[f\left(X_\tau\right)\right], \tag{2}$$

where $\mathcal{T}_T$ is the set of all stopping times taking values in $\{1, 2, \cdots, T\}$. A stopping time is a random variable τ with $\{\tau \leq k\} \in \sigma(X_1, \cdots, X_k)$ for every k.

- Each $X_i \in \mathbb{R}^d$ is assumed to be a multivariate random variable.
- The utility function can be generalized to $f(x, t)$ such as $e^{-\gamma t} x$ to include penalties on late decisions.
- Can be viewed as a special bandit problem.
- We are interested in estimating U_T, assuming we can simulate the underlying process.

Outline

Solving the optimal stopping problem

Definition (One more notation)

- $U_{T-k}(x_1, \cdots, x_k)$: The utility if one starts the $(T-k)$-stage optimal stopping problem after observing the first k outcomes.

Then the dynamical programming can be written as:

$$\begin{cases} U_1 = \mathbb{E}\left[f\left(X_T\right) \mid \{x_i\}_{i=1}^{T-1}\right], \\ U_{T-k} = \mathbb{E}\left[\max\left\{f(X_{k+1}), U_{T-(k+1)}\right\} \mid \{x_i\}_{i=1}^{k}\right], \qquad k = 0, \cdots, T-1. \end{cases}$$

Solving the optimal stopping problem

Definition (One more notation: conditional utility)

- $U_{T-k}(x_1, \cdots, x_k)$: The utility if one starts the $(T-k)$-stage optimal stopping problem after observing the first k outcomes.

Then the dynamical programming can be written as:

$$\begin{cases} U_1(x_1, \cdots, x_{T-1}) = \mathbb{E}\left[f\left(X_T\right) \mid \{x_i\}_{i=1}^{T-1}\right], \\ U_{T-k}(x_1, \cdots, x_k) = \mathbb{E}\left[\max\left\{f(X_{k+1}), U_{T-(k+1)}\right\} \mid \{x_i\}_{i=1}^{k}\right], \quad k \leq T-1. \end{cases}$$

Main idea: Solve $U_1, U_2, \cdots$ in a backward recursively way, and finally solve U_T.

Systematic bias for the plug-in estimators

Consider the two stage optimal stopping problem:

$$U_2 = \mathbb{E}[\max\{U_1(X_1), f(X_1)\}] = \int \max\{U_1(x_1), f(x_1)\}\pi_1(dx_1).$$

The plug-in estimator works as follows:

Systematic bias for the plug-in estimators

Consider the two stage optimal stopping problem:

$$U_2 = \mathbb{E}[\max\{U_1(X_1), f(X_1)\}] = \int \max\{U_1(x_1), f(x_1)\}\pi_1(dx_1).$$

The plug-in estimator works as follows:

- Simulate $x_1 \sim \pi_1$.
- Estimate $U_1(x_1)$ by an unbiased estimator $\hat{U}_1(x_1)$.
- Estimate U_2 by $\hat{U}_2 := \max\{\hat{U}_1(x_1), f(x_1)\}$.

Systematic bias for the plug-in estimators

Consider the two stage optimal stopping problem:

$$U_2 = \mathbb{E}[\max\{U_1(X_1), f(X_1)\}] = \int \max\{U_1(x_1), f(x_1)\}\pi_1(dx_1).$$

The plug-in estimator works as follows:

- Simulate $x_1 \sim \pi_1$.
- Estimate $U_1(x_1)$ by an unbiased estimator $\hat{U}_1(x_1)$.
- Estimate U_2 by $\max\{\hat{U}_1(x_1), f(x_1)\}$.

However... this estimator has systematic bias ☹

Systematic bias for the plug-in estimators

Consider the two stage optimal stopping problem:

$$U_2 = \mathbb{E}[\max\{U_1(X_1), f(X_1)\}] = \int \max\{U_1(x_1), f(x_1)\}\pi_1(dx_1).$$

The plug-in estimator works as follows:

- Simulate $x_1 \sim \pi_1$.
- Estimate $U_1(x_1)$ by an unbiased estimator $\hat{U}_1(x_1, x_2)$.
- Estimate U_2 by $\hat{U}_2 := \max\{\hat{U}_1(x_1), f(x_1)\}$.

However... this estimator has systematic bias We can calculate:

$$\begin{aligned}
\mathbb{E}[\hat{U}_2] &= \int\int \max\{\hat{U}_1(x_1, x_2), f(x_1)\}\pi_1(dx_1)\pi(dx_2|x_1) \\
&\geq \int \pi_1(dx_1)\max\{\mathbb{E}_{x_2|x_1}[\hat{U}_1], f(x_1)\} = \mathbb{E}[\max\{U_1(X_1), f(X_1)\}] = U_2.
\end{aligned}$$

Bias analysis, and debiasing techniques

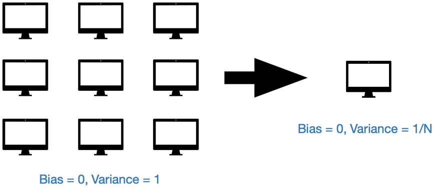
- Essentially, the bias comes from the following fact: If X is unbiased for μ, then $g(X)$ is biased for $g(\mu)$ in general.
- Examples include:
  - $g(x) = x^2$, then $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$,
  - $g(x) = \max\{x, c\}$,
  - $g(x)$ convex/concave.
- Why we need unbiasedness?

Bias analysis, and debiasing techniques

- Essentially, the bias comes from the following fact: If X is unbiased for μ, then $g(X)$ is biased for $g(\mu)$ in general.
- Examples include:
  - $g(x) = x^2$, then $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$,
  - $g(x) = \max\{x, c\}$,
  - $g(x)$ convex/concave.
- Why we need unbiasedness?
  - Easily implemented on parallel processors!
  - Uncertainty quantification.

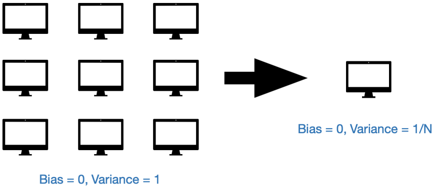
Why unbiasedness?

- Unbiased estimators

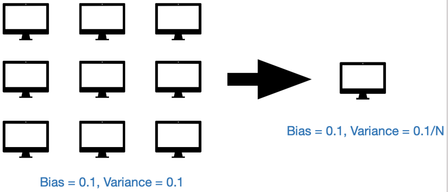


Why unbiasedness?

- Unbiased estimators



- Biased estimators



Short history of debiasing techniques

- (Applied probability) Generating new coins from old:
  - von Neumann gave an algorithm for generating fair coins from biased coins in 1951.
  - The general 'Bernoulli factory' problem is mathematically solved by Keane and O'Brien [10] in 1994.
  - Recent works connect 'Bernoulli factory' with MCMC. [11].

Short history of debiasing techniques

- (Applied probability) Generating new coins from old:
  - von Neumann gave an algorithm for generating fair coins from biased coins in 1951.
  - The general 'Bernoulli factory' problem is mathematically solved by Keane and O'Brien [10] in 1994.
  - Recent works connect 'Bernoulli factory' with MCMC. [11].
- (Stats) The unbiased MCMC by Jacob, O'Leary, and Atchadé [8].
  - Remove the bias in MCMC estimators. Applications include Bayesian inference, convergence diagnosis [3], asymptotic variance estimation.

Short history of debiasing techniques

- (Applied probability) Generating new coins from old:
  - von Neumann gave an algorithm for generating fair coins from biased coins in 1951.
  - The general ‘Bernoulli factory’ problem is mathematically solved by Keane and O’Brien [10] in 1994.
  - Recent works connect ‘Bernoulli factory’ with MCMC. [11].
- (Stats) The unbiased MCMC by Jacob, O’Leary, and Atchadé [8].
  - Remove the bias in MCMC estimators. Applications include Bayesian inference, convergence diagnosis [3], asymptotic variance estimation.
- (Operation research) The Randomized Multilevel Monte Carlo (MLMC) framework proposed by Glynn, Rhee, Blanchet et al.
  - The MLMC framework is developed by Heinrich and Giles [7, 5, 6] for SDE related problems.
  - Remove the bias in estimating $g(\mu)$ where μ is an expectation.
  - Applications include gradient estimation, robust optimization, PDEs.

The main idea: telescoping sum expression

- Target: Unbiased estimators of $g(\mu) = g(\mathbb{E}_\pi(X))$ given a simulator of $X_1, X_2, \cdots \sim \pi$.
- Idea: Write the expectation as telescoping sums. Let k_n be a sequence going to infinity with $n \to \infty$.

The main idea: telescoping sum expression

- Target: Unbiased estimators of $g(\mu) = g(\mathbb{E}_\pi(X))$ given a simulator of $X_1, X_2, \dots \sim \pi$.
- Idea: Write the expectation as telescoping sums. Let k_n be a sequence going to infinity with $n \to \infty$.

$$
\begin{aligned}
g(\mu) &= \lim_{n\to\infty} \mathbb{E}\left[g\left(\frac{S_{k_n}}{k_n}\right)\right] \qquad \text{(by LLN)} \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[g\left(\frac{S_{k_n}}{k_n}\right) - g\left(\frac{S_{k_{n-1}}}{k_{n-1}}\right)\right] \quad (S_0/0 := 0) \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[\Delta_n\right] \quad (\Delta_n := g\left(\frac{S_{k_n}}{k_n}\right) - g\left(\frac{S_{k_{n-1}}}{k_{n-1}}\right))
\end{aligned}
$$

The main idea: telescoping sum expression

- Now we have $g(\mu) = \sum_{n=1}^{\infty} \mathbb{E}[\Delta_n]$. The unbiased MLMC algorithm samples N with $\mathbb{P}(N = n) = p_n$, and samples $i.i.d.$ random variables $X_1, ..., X_{k_N}$. The 'final' estimator is:

$$\widehat{g(\mu)} := \frac{\Delta_N}{p_N}.$$

The MLMC estimator: Δ_N/p_N

- Now we have $g(\mu) = \mathbb{E}[\sum_{n=1}^{\infty} \Delta_n]$. The unbiased MLMC algorithm samples N with $\mathbb{P}(N = n) = p_n$, and samples $i.i.d.$ random variables $X_1, ..., X_{k_N}$. The 'final' estimator is:

$$\widehat{g(\mu)} := \frac{\Delta_N}{p_N}.$$

- Quick validity proof:

$$\begin{aligned}\mathbb{E}[\Delta_N/p_N] &= \mathbb{E}[\mathbb{E}[\Delta_N/p_N \mid N]] \quad \text{(law of iterated expectation)}\\ &= \sum_{n=1}^{\infty} \frac{\mathbb{E}[\Delta_n]}{p_n} \cdot p_n = \sum_{n=1}^{\infty} \mathbb{E}[\Delta_n] = g(\mu).\end{aligned}$$

The MLMC estimator: Δ_N / p_N

- There are still technical details such as the variance of the estimator, the computational cost that need to be addressed. The final estimator by Blanchet and Glynn [4] uses: $k_n = 2^n, N \sim \mathsf{Geo}(p)$, and

$$\Delta_n := g(S_{2^n}/2^n) - \frac{g(S^{\mathsf{E}}_{2^n}/2^{n-1}) + g(S^{\mathsf{O}}_{2^n}/2^{n-1})}{2}$$

  where $S^{\mathsf{E}}_{2^n}, S^{\mathsf{O}}_{2^n}$ sums up the even and odd terms in $X_1, \cdots, X_{2^n}$ respectively.

The MLMC estimator: Δ_N / p_N

- There are still technical details such as the variance of the estimator, the computational cost that need to be addressed. The final estimator by Blanchet and Glynn [4] uses: $k_n = 2^n, N \sim \text{Geo}(p)$, and

$$\Delta_n := g(S_{2^n}/2^n) - \frac{g(S_{2^n}^{\mathsf{E}}/2^{n-1}) + g(S_{2^n}^{\mathsf{O}}/2^{n-1})}{2}$$

  where $S_{2^n}^{\mathsf{E}}, S_{2^n}^{\mathsf{O}}$ sums up the even and odd terms in $X_1, \cdots, X_{2^n}$ respectively.
- Intuition behind this construction: Cancel out both the constant and linear terms when doing Taylor expansion around $\mu = \mathbb{E}[X]$.
- The resulting estimator is unbiased, has finite variance and finite computational complexity when certain technical assumptions are satisfied.

Outline

Go back to optimal stopping

However, the unbiased MLMC framework cannot be directly applied even for the two-stage optimal stopping problem. Since

- All the previous results require g to be continuously differentiable.
- The two-stage problem is a 'nested Monte Carlo' problem, i.e.,

$$U_2 = \mathbb{E}_{x_1 \sim \pi_1}[g_{x_1}(\mathbb{E}_{x_2 \sim \pi(\cdot|x_1)}[U_1(x_1)])]$$

  where $g_{x_1}(s) = \max\{f(x_1), s\}$ with x_1 fixed.
- The general multi-stage problem is more complicated, which requires a backward recursion.

Our contribution

- We construct an unbiased estimator MUSE (multilevel unbiased stopping estimator) for the utility U_T of the multi-stage optimal stopping problem.
- We relax the technical assumptions in the MLMC framework, and show MUSE has both finite variance and finite computational complexity.
- The computational cost to achieve ε-accuracy is $O(1/\varepsilon^2)$, which matches the optimal rate from the Central Limit Theorem (CLT).
- As by-products, we construct CLT and bootstrap-based confidence intervals for U_T. We also design an algorithm for finding the optimal stopping time.
- We test the behavior of MUSE on option pricing problems.

The two-stage MUSE

Algorithm 1 Two-stage MUSE

1: **Input:** A simulator $\mathcal{S}$ of the two-stage process (X_1, X_2), parameter $r \in (1/2, 1)$.
2: **Output:** An unbiased estimator of U_2.
3: **Step 1**. Sample $N \sim \text{Geo}(r)$.
4: **Step 2**. Sample $X_1(1)$, and $X_2(1), \cdots X_2(2^N)$ given $X_1(1)$.
5: **Step 3**. Calculate $S_{2^N} = f(X_2(1)) + \cdots f(X_2(2^N))$, and $S^{\mathsf{E}}_{2^N}, S^{\mathsf{O}}_{2^N}$ respectively.
6: **Step 4**. Calculate

$$\Delta_N = \max\left\{f\left(X_1(1)\right), \frac{S_{2^N}}{2^N}\right\} - \frac{1}{2}\left[\max\left\{f\left(X_1(1)\right), \frac{S^{\mathsf{O}}_{2^{N-1}}}{2^{N-1}}\right\} + \max\left\{f\left(X_1(1)\right), \frac{S^{\mathsf{E}}_{2^{N-1}}}{2^{N-1}}\right\}\right].$$

7: **Step 5**. **Return**: $Y := \Delta_N / p_r(N)$.

The two-stage MUSE, theoretical guarantees

Theorem (Theoretical properties of the two-stage MUSE)

Consider a two-stage process (X_1, X_2). *Suppose* $\mathbb{E}[\|X_i\|^{2+\delta}] < \infty$ *for* $i = 1, 2$ *and* $\delta > 0$. *Suppose* f *is at most linear growth, i.e.,* $|f(x)| \leq L\,(1 + \|x\|)$ *for some* $L > 0$. *Moreover,*

$$\mathbb{P}\left(|\mathbb{E}[f(X_2) \mid X_1] - f(X_1)| \leq \varepsilon\right) < C\varepsilon \tag{3}$$

for all $\varepsilon > 0$. *Let* $r = 1 - 2^{-\frac{2+9\delta/(80+40\delta)}{2+\delta/10}} \in (1/2, 1)$ *in Algorithm 1. Then, the resulting estimator* Y *in Algorithm 1 has the following properties:*

(1) $\mathbb{E}[Y] = U_2$.

(2) *The expected computational complexity of* Y *is finite.*

(3) $\mathbb{E}\left[|Y|^{2+\frac{\delta}{10}}\right] \leq \widetilde{C} \cdot L^{2+\delta}\left[1 + \mathbb{E}\left[\|X_2\|^{2+\delta}\right]\right]$, *where* $\widetilde{C}$ *is a constant independent of* (X_1, X_2).

Going to the general case

- The last property guarantees Y has $(2+\frac{\delta}{10})$-th moment, provided that the original random variables has $(2+\delta)$-th moment. This is stronger than Y has finite variance.
- Moreover, this property is crucial for algorithm design for the general optimal stopping problem.

Going to the general case

- The last property guarantees Y has $(2+\frac{\delta}{10})$-th moment, provided that the original random variables has $(2+\delta)$-th moment. This is stronger than Y has finite variance.
- Moreover, this property is crucial for algorithm design for the general optimal stopping problem.
- When designing the general MUSE, we will use Algorithm 1 recursively. In the end, the estimator will have $(2+\delta/10^{T-1})$-th moment.
- Main idea: construct unbiased estimators for $U_1, U_2, \cdots$ until U_T. The problem for U_1 is straightforward, for U_2 it is solved by Algorithm 1. For U_3 we need to call Algorithm 1 for 2^N times and then construct the unbiased estimator by the MLMC approach ...

The Multi-stage MUSE

Algorithm 2 Multi-stage MUSE

Input: Time index k. Trajectory history $H = \{x_1, \cdots, x_k\}$ or $\varnothing$. A simulator $\mathcal{S}$, parameters $r_{k+1}, \cdots, r_{T-1} \in (1/2, 1)$.
Output: An unbiased estimator of $U_{T-k}(H)$.
if $k = T - 1$ **then**
    Sample one x_T. **Return**: $Y := f(x_T)$.
else
    Sample x_{k+1}, add x_{k+1} to H. Sample $N_{k+1} \sim \text{Geo}(r_{k+1})$.
    Call Algorithm 2 for $2^{N_{k+1}}$ times with $(k+1; H; \mathcal{S}; r_{k+2} \cdots, r_{T-1})$, label the outputs by $Y_{k+1}(1), \cdots, Y_{k+1}(2^{N_{k+1}})$.
    Calculate $S_{2^{N_{k+1}}}, S^{\mathsf{E}}_{2^{N_{k+1}}}, S^{\mathsf{O}}_{2^{N_{k+1}}}$ respectively.
    Calculate $\Delta_{N_{k+1}}$.
    Return: $Y := \Delta_{N_{k+1}} / p_{r_{k+1}}(N_{k+1})$.
end if

The multi-stage MUSE, theoretical guarantees

Theorem (Theoretical properties of the MUSE, informal)

With the same assumption as above, consider the input

$$(0; \varnothing; \mathcal{S}, r_1, \cdots, r_{T-1})$$

in Algorithm 2 for $1 \leq i \leq T-1$. *Then, the resulting estimator* Y *in Algorithm 2 has the following properties:*

(1) $\mathbb{E}[Y] = U_T$.

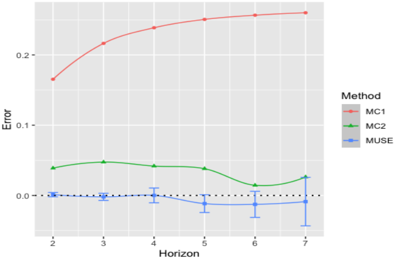
(2) The expected computational complexity is finite.

(3) $\mathrm{Var}\,(Y) < \infty$.

Outline

Optimal Stopping of Independent Random Variables

- Consider the case where X_i are $i.i.d.$ $\mathcal{N}(0,1)$ random variables with reward $f(x) = x$. The utility can be solved numerically. With each fixed time horizon $T \in \{2, \cdots, 7\}$, three estimators – MUSE and two vanilla Monte Carlo estimators MC1 and MC2 are implemented.



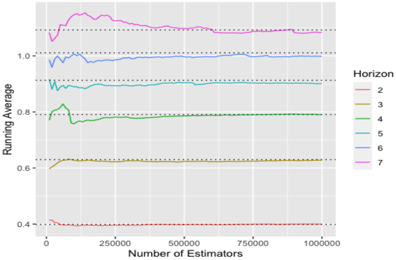


Figure: Left: Comparison between the errors of the MUSE (blue), MC1(red), and MC2(green) for estimating the utility for $i.i.d.$ standard Gaussian random variables. Blue error bars stands for the 95% confidence intervals of the MUSE. Black dotted line stands for the ground truth (error $= 0$). Right: The traceplot of the running averages of the MUSE with different horizons. Black dotted line stands for the ground truth.

Pricing the Bermudan options with high-dimensional inputs on a computer cluster

- The underlying process is a $X(t) := (X^{(1)}(t), \cdots, X^{(d)}(t))$ is a d-dimensional geometric Brownian motion with drift $r - \delta$ and volatility σ.
- Utility: $f(t, X_t) = e^{-rt} \max\{0, K - \sum_{i=1}^{d} X_t^{(i)}/d\}$ at each t, where K is the strike price and e^{-r} is often referred to as the discounting factor.
- Bermudan option is only exercisable in a discrete set of times, which transforms the pricing problem to solving the optimal stopping problem: $U_T := \sup_{\tau \in \{T_1, \cdots T_k\}} \mathbb{E}\left[f\left(\tau, X_\tau\right)\right]$, where $0 \leq T_1 \leq \cdots \leq T_k \leq T$ are all the exercisable dates.
- Existing experiments on Bermudan options often assumes $d \leq 20$, though it can be as large as 5000 in practice [1].

Pricing the Bermudan options

- In our experiment we adopt the standard parameters where $T = 3$ (years), $\sigma = 0.2, r = 0.05, \delta = 0, K = X_0^{(i)} = 100$ for every i.
- We first benchmark our result with existing results when $d = 5$, next we present our results for $d \in \{10, 20, 100, 1000\}$. For each d, we use 10^7 MUSEs generated by a 500-core CPU-based computer cluster, where the parameters r_i are set to be 0.6 for each stage. The results when $d = 5$ is presented in Table 1, the MUSE matches the results from other methods while preserving unbiasedness and having a relatively small standard error.

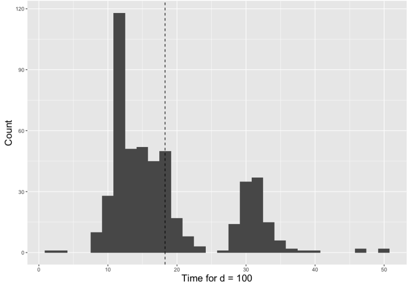
Method	LSM (s.e.)	SGM direct (s.e.)	SGM LB (s.e.)	BKS (95% CI)	**MUSE (s.e.)**
$d = 5$	$2.163(0.001)$	$2.141(0.008)$	$2.134(0.012)$	$[2.154, 2.164]$	$\mathbf{2.161}(\mathbf{0.004})$

Table: Comparison between different methods when $d = 5$. SGM and BKS stands for results reported by [9] and [2] respectively. LSM stands for Longstaff–Schwartz method, reported by [9].

Pricing the Bermudan options

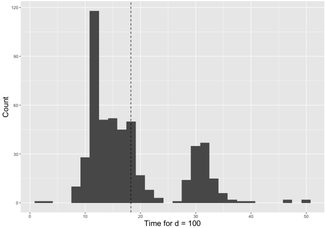
- The table below records the estimates and the standard errors of the MUSE when d is increasing. The plot shows the histogram of the computing time among 500 processors when $d = 100$.

d	MUSE (s.e.)	Average Time (s) per processor
5	2.161 (0.004)	15.922
10	0.985 (0.002)	14.787
20	0.355 (0.001)	16.004
100	0.0043 ($< 10^{-4}$)	18.271
1000	0(0)	32.191



Pricing the Bermudan options

d	MUSE (s.e.)	Average Time (s) per processor
5	2.161 (0.004)	15.922
10	0.985 (0.002)	14.787
20	0.355 (0.001)	16.004
100	0.0043 ($< 10^{-4}$)	18.271
1000	0(0)	32.191



- Comparing with other competitors, MUSE is the only unbiased estimator.
- The MUSE scales well with the dimensionality d.
- The average computing cost is OK, but variance can be high.
- It does not scale well with the number of horizon T.

Thanks!

- Check out the paper on arxiv https://arxiv.org/abs/2106.02263 "Unbiased Optimal Stopping via the MUSE"

SCAN ME

References I

Sebastian Becker, Patrick Cheridito, Arnulf Jentzen, and Timo Welti.
Solving high-dimensional optimal stopping problems using deep learning.
arXiv preprint arXiv:1908.01602, 2019.

Christian Bender, Anastasia Kolodko, and John Schoenmakers.
Policy iteration for American options: Overview.
Monte Carlo Methods and Applications, 2006.

Niloy Biswas, Pierre E Jacob, and Paul Vanetti.
Estimating convergence of markov chains with l-lag couplings.
Advances in neural information processing systems, 32, 2019.

Jose H. Blanchet and Peter W. Glynn.
Unbiased Monte Carlo for optimization and functions of expectations via multi-level randomization.
2015 Winter Simulation Conference (WSC), pages 3656–3667, 2015.

Michael B Giles.
Multilevel Monte Carlo path simulation.
Operations Research, 56(3):607–617, 2008.

Michael B Giles.
Multilevel Monte Carlo methods.
Acta Numer., 24:259–328, 2015.

Stefan Heinrich.
Multilevel Monte Carlo methods.
In *International Conference on Large-Scale Scientific Computing*, pages 58–67. Springer, 2001.

References II

Pierre E Jacob, John O'Leary, and Yves F Atchadé.
Unbiased Markov chain Monte Carlo methods with couplings.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82(3):543–600, 2020.

Shashi Jain and Cornelis W Oosterlee.
Pricing high-dimensional Bermudan options using the stochastic grid method.
International Journal of Computer Mathematics, 89(9):1186–1211, 2012.

MS Keane and George L O'Brien.
A Bernoulli factory.
ACM Transactions on Modeling and Computer Simulation (TOMACS), 4(2):213–219, 1994.

Dootika Vats, Flávio Gonçalves, Krzysztof Łatuszyński, and Gareth O Roberts.
Efficient bernoulli factory mcmc for intractable posteriors.
arXiv preprint arXiv:2004.07471, 2020.