Approximate Dynamic Programming for American Options

Belomestny D. [1,2], **Kaledin M.**[1], Schoenmakers J.[3]

[1]HDI Lab, Higher School of Economics, Moscow;[2]University of Duisburg-Essen
[3]Weierstrass Institute for Applied Analysis and Stochastics, Berlin

HDI Lab

Introduction

Optimal stopping problems are very common in statistics, economics and, naturally, in mathematical finance (option pricing and portfolio management). They are solved via dynamic programming and much research is concentrated around approximations of value functions. The most popular algorithms are (Longstaff and Schwartz 2001) and (Tsitsiklis and Van Roy 2001), they demonstrated the efficiency of least-squares regression approach, (Clément, Lamberton, and Protter 2002) and (Zanger 2013) established the convergence properties. So far, the methods were being designed for discrete time problems, however there are ones in continuous time (say, american options) and natural desire is to use discrete-time methods on dense time grids to obtain approximate solution. We claim that Weighted Grid Monte Carlo (WGMC) algorithm is in some sense computationally better than regression approaches.

Optimal Stopping Problem

Given

- Stochastic process $X_t \in \mathbb{R}^d$ and starting point x_0,
- Reward function $f(x,t)$ depending on state and time,
- Set of times $t \in \mathcal{T}$,

find the optimal $\tau \in \mathcal{T}$ such that

$$\mathbb{E}\left[f(X_\tau,\tau)|X_0 = x_0\right]$$

is maximized. In option pricing, the maximum value is usually called *fair price of an option*.

Let stochastic process be defined as a diffusion

$$dX_s = a(X_s)ds + \sigma(X_s)dW_s \tag{1}$$

with $a(X_s) \in \mathbb{R}^d$, $\sigma(X_s) \in \mathbb{R}^{d\times m}$ and W_s being a vector of uncorrelated Wiener processes.

Backward Induction

Consider a discrete set $\mathcal{T} = t_0,..,t_{L-1}$ on $[0,T]$ with $h = T/L$ and define the Bellman operator

$$\Gamma\left[V_{t+1}\right](X_t) = \max\left[f(X_t,t), \int V_{t+1}(x')p(dx'|X_t)\right].$$

Backward induction:

1. Define $V_L(X_{t_{L-1}}) = f(X_{t_{L-1}}, t_{L-1})$.
2. Recursively $V_t(X_t) = \Gamma\left[V_{t+1}\right](X_t)$ for $t = L-2,..,0$.
3. Output V_0.

Conditional expectation should be approximated: one way is to use regression for value functions.

Complexity Issues with Regression Approach

Using Corollary 3.10 in (Zanger 2013) for the basis of degree-m polynomials and assuming the error due to the time discretization is of order $L^{-\beta}$ for some $0<\beta<1$, the complexities for discrete and continuous problems ($L\to\infty$) are

$$\mathcal{C}_L(\varepsilon,d) = \frac{L\,5^{L(2+3d/\alpha)}}{\varepsilon^{2+3d/\alpha}}, \quad \mathcal{C}_\infty(\varepsilon,d) = O\left(\frac{\varepsilon^{-1/\beta}\,5^{(2+3d/\alpha)\varepsilon^{-1/\beta}}}{\varepsilon^{2+3d/\alpha}}\right).$$

Weighted Grid Monte Carlo Algorithm

Suppose the transition density $p_h(X_{t+h}|X_t)$ from t to $t+h$ of the process X_t is known. Use the following procedure (Broadie and Glasserman 2004):

- Simulate N trajectories $\tilde{X}_k^t,\ t = \overline{0,L}, \quad k = \overline{1,N} \quad$ from x_0;
- Use the following approximation:

$$\int V(x^{t+1})p(dx^{t+1}|\tilde{X}_i^t) \approx \sum_{j=1}^{N} V(\tilde{X}_j^{t+1})p_N(\tilde{X}_j^{t+1}|\tilde{X}_i^t),$$

$$\text{where} \qquad p_N(\tilde{X}_j^{t+1}|\tilde{X}_i^t) = \frac{p_h(\tilde{X}_j^{t+1}|\tilde{X}_i^t)}{\sum_{k=1}^{N} p_h(\tilde{X}_j^{t+1}|\tilde{X}_k^t)}.$$

Main Result

Definition 1 *Let us call the algorithm which has complexity $\mathcal{C}(\varepsilon,d)$ such that*

$$\lim_{\varepsilon\searrow 0}\frac{\log\mathcal{C}(\varepsilon,d)}{\log(1/\varepsilon)} = \infty$$

being subject to Curse of Time Discretization(CoTD).

Clearly, the Longstaff-Schwartz method generally possesses CoTD.

Definition 2 *Let there is an algorithm $\mathcal{A}$ which is able to solve the problem with complexity $\mathcal{C}(d,\epsilon)$, then the number*

$$\Gamma^{\mathcal{A}} := \lim_{d\nearrow\infty}\lim_{\varepsilon\searrow 0}\frac{\log\mathcal{C}(\varepsilon,d)}{d\log(1/\varepsilon)}$$

is called the semi-tractability index of algorithm $\mathcal{A}$.

If $\Gamma^{\mathcal{A}} = 0$, we call such algorithm semi-tractable. The key assumptions for the results are

(AG) Suppose that $c_g > 0$ is such that

$$g(t,x) \le c_g(1+|x|) \quad \text{for all } 0\le t\le T,\ x\in\mathbb{R}^d. \tag{2}$$

(AX) Assume that there exists some $c_{\bar{X}} > 0$ such that for all $0\le l\le L$, $x\in\mathbb{R}^d$,

$$\mathrm{E}_{\mathcal{F}_t}\left[\sup_{l\le l'\le L}|X_{l'h}|\,\middle|\,X_{lh} = x\right] \le c_X(1+|x|), \tag{3}$$

uniformly in the choice of h. This is satisfied under Lipschitz conditions on the coefficients of the SDE (1), and can be proved by using the Burkholder-Davis-Gundy inequality and the Gronwall lemma.

(AP) $(X_{lh},\ l = 0,\dots,L)$ is time homogeneous with transition densities $p_h(y|x)$ that satisfy the Aronson type inequality: there exist positive constants $\overline{\varkappa}$ and $\overline{\alpha}$ such that for any $x,y\in\mathbb{R}^d$ and any $h>0$ it holds that

$$p_h(y|x) \le \frac{\overline{\varkappa}}{(2\pi\overline{\alpha}h)^{d/2}}e^{-\frac{|x-y|^2}{2\overline{\alpha}h}}.$$

This assumption holds if the coefficients in (1) are bounded and σ is uniformly elliptic.

Theorem 3 *(Belomestny,Kaledin,Schoenmakers) Under the assumptions WGMC is not subject to CoTD and one obtains the following semi-tractability indices:*

	WGMC	*Longstaff-Schwartz*
Discrete Time	*0*	$3/\alpha$
Continious Time	*2*	∞

Table 1: *Comparison of semi-tractability indices for considered algorithms*

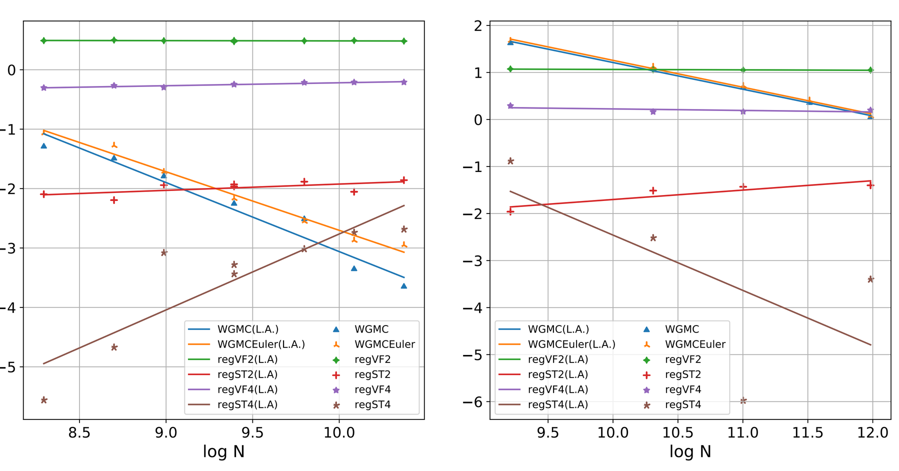
If the transition density p_h is not known one still can prove for suitable approximations of density $\overline{p_h}$ and corresponding Markov process $\overline{X}_t$ the same type of results using similar assumptions.

Theorem 4 *(Belomestny, Kaledin, Schoenmakers) Suppose that there exists an approximating sequence $\{\overline{p}_n\}$ such that for some $\varkappa > 1$ and natural $m > 0$*

$$\left|\frac{\overline{p}_n(y|x_0) - p_h(y|x_0)}{\overline{p}_n(y|x_0)}\right| \lesssim \frac{h^{\varkappa}(1+|y-x_0|^m)}{n!}.$$

Then WGMC with approximated density is correct and is not subject to CoTD.

Numerical Experiments: Convergence Rates and CoTD Effect



$(d=2)$

WGMCRate≈ -1.16,

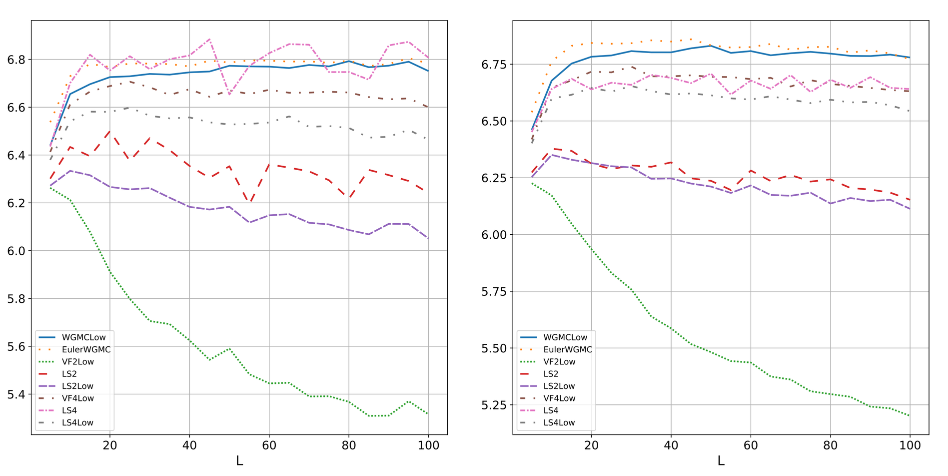
WGMCEulerRate≈ -0.98

$(d=5)$

WGMCRate≈ -0.57,

WGMCEulerRate≈ -0.57

Figure 1: Deviations of the estimates from the true price (log-log axes), 'L.A' means '*linear approximation*' . The curves 'VF' and 'ST' correspond to Tsitsiklis-VanRoy and Longstaff-Schwarz algorithm, the number in the name means the degree of polynomials in the basis. The problem is Bermudan Max-Call option pricing (example from Glasserman 2003, Ch.8, true price is 13.90). State x_t is given as a vector of independent Geometric Brownian Motions with initial state $x_0^i = 100$ for all $i = 1,..,d$. The payoff is $g(t,x) = (\max_{i=1,..,d} x_t^i - K)_+$ with strike price $K = 100$. Other model parameters include $\mu = -0.05, \sigma = 0.2,\ T = 3, h = T/L = 1/3$ .Euler-Maruyama method was used as approximate density for WGMCEuler.



$(N_{train} = 500)$

$(N_{train} = 2000)$

Figure 2: One-dimensional max-put american option, model parameters: $d = 1,\ T = 3,\ \mu = 0.08,\ \sigma = 0.2,\ g(t,x) = (K - \max_{i=1,..,d} x_t^i)_+$. True American option price is 6.9320 (Jin Kim, Ma, and Choe 2013, Table 2, inst. 23), the charts present the lower estimates for 500 and 1000 train trajectories, $N_{test} = 20000$. Lines 'VF' and 'LS' correspond to Tsitsiklis-VanRoy and Longstaff-Schwarz estimators(same legend conventions as on the left). One can immediately note that WGMC estimator is very stable and becomes more accurate as N_{train} grows while regression methods stabilize. Note also that regression estimators diverge as L grows and to overcome it one needs to increase the degree of polynomial in the basis.

References

Broadie, M. and P. Glasserman (2004). "A stochastic mesh method for pricing high-dimensional American options". In: *Journal of Computational Finance* 7.4, pp. 35–72.

Clément, E., D. Lamberton, and P. Protter (2002). "An analysis of a least squares regression method for American option pricing". In: *Finance and Stochastics* 6.4, pp. 449–471.

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Vol. 53. Springer Science & Business Media.

Jin Kim, B., Y.-K. Ma, and H. Choe (Feb. 2013). "A Simple Numerical Method for Pricing an American Put Option". In: *Journal of Applied Mathematics* 2013. DOI: 10.1155/2013/128025.

Longstaff, F. and E. Schwartz (2001). "Valuing American options by simulation: a simple least-squares approach.". In: *Review of Financial Studies* 14.1, pp. 113–147.

Tsitsiklis, J. and B. Van Roy (2001). "Regression methods for pricing complex American style options.". In: *IEEE Trans. Neural. Net.* 12.14, pp. 694–703.

Zanger, D. Z. (2013). "Quantitative error estimates for a least-squares Monte Carlo algorithm for American option pricing". In: *Finance and Stochastics* 17.3, pp. 503–534.