Multilevel primal and dual approaches for pricing American options

Denis Belomestny\(^1\)  Marcel Ladkau\(^2\)  John Schoenmakers\(^2\)

\(^1\)Duisburg-Essen University, Duisburg, Germany
\(^2\)Weierstrass Institute, Berlin, Germany

1. Let \((Z_j)_{j \geq 0}\) be a nonnegative adapted process on a filtered probability space \((\Omega, \mathcal{F} = (\mathcal{F}_j)_{j \geq 0}, \mathbb{P})\) representing the discounted payoff of an American option, so that the holder of the option receives \(Z_j\) if the option is exercised at time \(j \in \{0, \ldots, T\}\) with \(T \in \mathbb{N}_+\). The pricing of American options can be formulated as a primal-dual problem. The primal representation corresponds to the following optimal stopping problems

\[ Y_j^* := \max_{\tau \in \mathcal{T}[j, \ldots, T]} \mathbb{E}_{\mathcal{F}_j}[Z_\tau], \quad j = 0, \ldots, T, \]

where \(\mathcal{T}[j, \ldots, T]\) is the set of \(\mathcal{F}\)-stopping times taking values in \(\{j, \ldots, T\}\). The process \(\left(Y_j^*\right)_{j \geq 0}\) is called the Snell envelope. \(Y^*\) is a supermartingale satisfying the Bellman principle

\[ Y_j^* = \max_{s \in \{j, \ldots, T\}} \left(Z_s - \pi_s^* + \pi_j^*\right) \quad \text{a.s.,} \]

where

\[ Y_j^* = Y_0^* + \pi_j^* - A_j^* \]  \hspace{1cm} (1)

is the (unique) Doob decomposition of the supermartingale \(Y_j^*\). That is, \(\pi^*\) is a martingale and \(A^*\) is an increasing process with \(\pi_0 = A_0 = 0\) such that (1) holds.

2. Assume that we are given a stopping family \((\tau_j)\) that is consistent, i.e.

\[ \tau_j > j \Rightarrow \tau_j = \tau_{j+1}, \quad j = 0, \ldots, T - 1. \]
The stopping policy defines a lower bound for \( Y^* \) via
\[
Y_j = \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_j}], \quad j = 0, \ldots, T.
\]
Consider now a new family \((\hat{\tau}_j)_{j=0,\ldots,T}\) defined by
\[
\hat{\tau}_j := \inf \{ k : j \leq k < T, Z_k \geq \mathbb{E}_{\mathcal{F}_k}[Z_{\tau_k+1}] \} \wedge T.
\]  
(2)
The basic idea behind (2) goes back to [6] in fact. For more general versions of policy iteration and their analysis, see [7]. Next, we introduce the \((\mathcal{F}_j)\)-martingale
\[
\pi_j = \sum_{k=1}^{j} \left( \mathbb{E}_{\mathcal{F}_k}[Z_{\tau_k}] - \mathbb{E}_{\mathcal{F}_{k-1}}[Z_{\tau_k}] \right), \quad j = 0, \ldots, T,
\]  
(3)
and then consider,
\[
\tilde{Y}_j := \mathbb{E}_{\mathcal{F}_j} \left[ \max_{k=j,\ldots,T} (Z_k - \pi_k + \pi_j) \right],
\]
onlyx{along with}
\[
\hat{Y}_j := \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_j}], \quad j = 0, \ldots, T.
\]

The following theorem states that \( \hat{Y} \) is an improvement of \( Y \) and that the Snell envelope process \( Y^*_j \) lies between \( \hat{Y}_j \) and \( \tilde{Y}_j \) with probability 1.

**Theorem 2.** It holds
\[
Y_j \leq \hat{Y}_j \leq Y^*_j \leq \tilde{Y}_j, \quad j = 0, \ldots, T \quad a.s.
\]

**3.** The main issue in the Monte Carlo construction of \( \hat{Y} \) and \( \tilde{Y} \) in a Markovian environment is the estimation of the conditional expectations in (2) and (3). We thus assume that the cash-flow \( Z_j \) is of the form \( Z_j = Z_j(X_j) \) for some underlying (possibly high-dimensional) Markovian process \( X^\prime \). A canonical approach is the use of sub simulations. In this respect we consider an enlarged probability space \((\Omega, \mathbb{F}', \mathbb{P})\), where \( \mathbb{F}' = (\mathcal{F}'_j)_{j=0,\ldots,T} \) and \( \mathcal{F}_j \subset \mathcal{F}'_j \) for each \( j \). On the enlarged space we consider \( \mathcal{F}'_j \) measurable estimations \( C_{j,M} \) of \( C_j = \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_{j+1}}] \) as being standard Monte Carlo estimates based on \( M \) sub simulations. More precisely
\[
C_{j,M} = \frac{1}{M} \sum_{m=1}^{M} Z_{\tau^{(m)}_{j+1}}(X^j_{\tau^{(m)}_{j+1}})
\]
where the \( \tau^{(m)}_{j+1} \) are evaluated on \( M \) sub trajectories all starting at time \( j \) in \( X_j \). Obviously, \( C_{j,M} \) is an unbiased estimator for \( C_j \) with respect to \( \mathbb{E}_{\mathcal{F}_j} [\cdot] \). We thus end up with a simulation based versions of (2) and (3) respectively,
\[
\hat{\tau}_{j,M} := \inf \{ k : j \leq k < T, Z_k > C_{k,M} \} \wedge T, \quad j = 0, \ldots, T,
\]
\[
\pi_{j,M} := \sum_{k=1}^{j} (Z_k - C_{k-1,M}) 1_{\{\tau_k = k\}} + \sum_{k=1}^{j} (C_{k,M} - C_{k-1,M}) 1_{\{\tau_k > k\}}.
\]

Denote
\[
\tilde{Y}_{j,M} := \mathbb{E}_{\mathcal{F}_j}[Z_{\tau_j,M}], \quad j = 0, \ldots, T
\]
and
\[
\tilde{Y}_{j,M} := \mathbb{E}_{\mathcal{F}_j} \left[ \max_{k=j,...,T} (Z_k - \pi_{k,M} + \pi_{j,M}) \right].
\]

**Theorem 3.** Let us assume that there exist constants \( B_{0,j} > 0, j = 0, \ldots, T - 1, \) and \( \alpha > 0, \) such that for any \( \delta > 0 \) and \( j = 0, \ldots, T - 1, \)
\[
\mathbb{P}(|\mathbb{E}_{\mathcal{F}_j}[Z_{\tau_{j+1}}] - Z_j| \leq \delta) \leq B_{0,j} \delta^\alpha.
\]

Further suppose that there are constants \( B_1 \) and \( B_2, \) such that \( |Z_j| < B_1 \) and
\[
\text{Var}_{\mathcal{F}_j}[Z_{\tau_{j+1}}] := \mathbb{E}_{\mathcal{F}_j}[(Z_{\tau_{j+1}} - C_j)^2] < B_2, \text{ a.s.} \quad (4)
\]
for \( j = 0, \ldots, T - 1. \) It then holds,
\[
|\tilde{Y}_0 - \tilde{Y}_{0,M}| \leq M^{-\frac{\delta + \alpha}{2}} B \sum_{k=0}^{T-1} B_{0,k},
\]
with some constant \( B \) depending only on \( \alpha, B_1 \) and \( B_2. \) Moreover, if for any \( \delta > 0 \)
\[
\mathbb{P}(|\mathbb{E}_{\mathcal{F}_j}[Z_{\tau_{j+1}}] - Z_j| \leq \delta) \leq \overline{B}_{0,j} \delta^\overline{\alpha}
\]
with some positive constants \( \overline{\alpha} \) and \( \overline{B}_{0,j}, \) \( j = 0, \ldots, T - 1, \) then
\[
\mathbb{E}[(Z_{\tau_0,M} - Z_{\bar{\tau}_0})^2] \leq M^{-\pi/2} 2 B_1^2 \overline{B} \sum_{j=0}^{T-1} \overline{B}_{0,j}.
\]

**Theorem 4.** Introduce for \( Z := \max_{j=0,...,T} (Z_j - \pi_j), \) the random set
\[
Q = \{ j : Z_j - \pi_j = Z \},
\]
and the \( \mathcal{F}_T \) measurable random variable
\[
\Lambda := \min_{j \notin Q} (Z - Z_j + \pi_j),
\]
with \( \Lambda := +\infty \) if \( Q = \{0, \ldots, T\}. \) Obviously \( \Lambda > 0 \text{ a.s.} \) Further suppose that
\[
\mathbb{E}[\Lambda^{-\xi}] < \infty \text{ for some } 0 < \xi \leq 1, \quad \text{and} \quad \#Q = 1.
\]
It then holds,
\[
|\tilde{Y}_0 - \tilde{Y}_{0,M}| \leq C M^{-\frac{\xi + 1}{2}}
\]
for some constant \( C. \)
For a fixed natural number $L$ and a geometric sequence $m_l = m_0 \kappa^l$, for some $m_0, \kappa \in \mathbb{N}, \kappa \geq 2$, we consider in the spirit of [4] the telescoping sum

$$\hat{Y}_{m_L} = \hat{Y}_{m_0} + \sum_{l=1}^{L} (\hat{Y}_{m_l} - \hat{Y}_{m_{l-1}}),$$

where $\hat{Y}_m := \hat{Y}_{0,m}$. Next we take a set of natural numbers $n : = (n_0, \ldots, n_L)$ satisfying $n_0 > \ldots > n_L \geq 1$, and simulate an initial set of cash-flows

$$\left\{ Z_{\hat{\tau}_{n_0}}^{(j)}, \quad j = 1, \ldots, n_0 \right\},$$

due to an initial set of trajectories $\{X_{0,x}^{(j)}, \quad j = 1, \ldots, n_0\}$, where

$$Z_{\hat{\tau}_{n_0}}^{(j)} := Z_{\hat{\tau}_{0,n_0}}^{(j)} \left( X_{0,x}^{(j)} \right).$$

Next we simulate independently for each level $l = 1, \ldots, L$, a set of pairs

$$\left\{ (Z_{\hat{\tau}_{n_l}}^{(j)}, Z_{\hat{\tau}_{n_l-1}}^{(j)}), \quad j = 1, \ldots, n_l \right\}$$

due to a set of trajectories $X_{0,x}^{(j)}, \quad j = 1, \ldots, n_l$, to obtain the multilevel estimator

$$\hat{\gamma}_{n,m} := \frac{1}{n_0} \sum_{i=1}^{n_0} Z_{\hat{\tau}_{n_0}}^{(i)} + \sum_{l=1}^{L} \frac{1}{n_l} \sum_{j=1}^{n_l} (Z_{\hat{\tau}_{n_l}}^{(j)} - Z_{\hat{\tau}_{n_l-1}}^{(j)})$$

for estimating $\hat{Y}$. (5)

4. With the notations of the previous section we define

$$\tilde{Y}_{m_L} = \tilde{Y}_{m_0} + \sum_{l=1}^{L} [\tilde{Y}_{m_l} - \tilde{Y}_{m_{l-1}}],$$

where $\tilde{Y}_m := \tilde{Y}_{0,m}$. Given a sequence $n : = (n_0, \ldots, n_L)$ with $1 \leq n_0 < \ldots < n_L$, we then simulate for $l = 0$ an initial set of trajectories

$$\left\{ (Z_{i,j}^{(i)}, \pi_{j,m_0}^{(i)}), \quad i = 1, \ldots, n_0, \quad j = 0, \ldots, T, \right\}$$

of the two-dimensional vector process $(Z_i, \pi_{j,m_0})$, and then for each level $l = 1, \ldots, L$, independently, a set of trajectories

$$\left\{ (Z_{i,j}^{(i)}, \pi_{j,m_l}^{(i)}, \pi_{j,m_l}^{(i)}), \quad i = 1, \ldots, n_l, \quad j = 0, \ldots, T \right\}$$

of the vector process $(Z_i, \pi_{j,m_l-1}, \pi_{j,m_l})$. Based on this simulation we consider the following multilevel estimator:

$$\tilde{\gamma}_{n,m} := \frac{1}{n_0} \sum_{i=1}^{n_0} Z_{m_0}^{(i)} + \sum_{l=1}^{L} \frac{1}{n_l} \sum_{i=1}^{n_l} [Z_{m_l}^{(i)} - Z_{m_l-1}^{(i)}]$$

(6)
with $Z_{ml}^{(i)} := \max_{j=0,...,T} \left( Z_j^{(i)} - \pi_{j,ml}^{(i)} \right)$, $i = 1,...,n_l$, $l = 0,...,L$.

5. We now consider the numerical complexity of the multilevel estimators (5) and (6), for convenience generically denoted by $X_{n,m}$. Assume that there are some positive constants $\gamma, \beta, \mu_\infty, \sigma_\infty$ and $\mathcal{V}_\infty$ such that $\text{Var}[X_m] \leq \sigma_\infty^2$, $|X - \mathbb{E}[X_m]| \leq \mu_\infty m^{-\gamma}$, $m \in \mathbb{N}$ and

$$\mathbb{E}[X_{ml} - X_{m_{l-1}}]^2 \leq \mathcal{V}_\infty m_l^{-\beta}, \quad l = 1,...,L.$$ 

**Theorem 5.** Let us assume that $0 < \beta \leq 1$, $\gamma \geq \frac{1}{2}$ and $m_l = m_0 \kappa^l$ for some fixed $\kappa$ and $m_0 \in \mathbb{N}$. Fix some $0 < \epsilon < 1$. Let $L = L(\epsilon)$ be the integer part of

$$\gamma^{-1} \ln^{-1} \kappa \ln \left[ \frac{\sqrt{2\mu_\infty}}{m_0^\gamma} \right], \quad \text{and} \quad n_l = n_0 \kappa^{-(1+\beta)/2} \text{ with } n_0 = n_0(\epsilon) = \frac{2\sigma_\infty^2}{\epsilon^2} + \frac{2\mathcal{V}_\infty}{\epsilon^2 m_0^\beta} \kappa^{(1-\beta)/2} - 1 \kappa^{(1-\beta)/2}.$$

Then the complexity needed to achieve the accuracy $\epsilon := \sqrt{\mathbb{E}[(X - X_{n,m})^2]} < \epsilon$ is

$$C_{n,m}^{ML}(\epsilon) = O(\epsilon^{-2-\frac{1-\beta}{\gamma}}), \quad \epsilon \searrow 0, \quad \text{for } \beta < 1,$$

$$C_{n,m}^{ML}(\epsilon) = O(\epsilon^{-2} \ln^2 \epsilon), \quad \epsilon \searrow 0, \quad \text{for } \beta = 1.$$

**References**


