Machine Learning in Finance

Exercise sheet 12

Exercise 12.1 (American Option) We consider discrete price process $S = (S^0, S^1)$ consisting of a risk-free bond $S_t^0 = (1+r)^t$ and a risky asset $(S_t^1)_{t=1}^T$. Assume the market is complete and no arbitrage, and let Q be the equivalent martingale measure. Consider the American option with payoff $f_t = (K - S_t^1)_+$ and V^* be the fair price process of the American option. Recall from Exercise Sheet 11 that

$$V_{t-1}^* = \max\{f_{t-1}, (1+r)^{-1}\mathbb{E}_Q[V_t^*|\mathcal{F}_{t-1}]\}, \quad t = 1, \ldots, T. \tag{1}$$

(a) Prove that if $\mathcal{F} = \sigma(S)$, then there exist $(g_t)_{t=0}^{T-1}$ such that

$$g_{t-1}(S_{t-1}^1) = \mathbb{E}_Q[(1+r)^{-1}V_t^*|\mathcal{F}_{t-1}], \quad t = 1, \ldots, T. \tag{2}$$

and

$$g_{t-1} = \arg\min_{h\in\mathcal{L}^0} \mathbb{E}_Q\left[\left(h(S_{t-1}^1) - (1+r)^{-1}V_t^*\right)^2\right], \quad t = 1, \ldots, T. \tag{3}$$

(b) Prove the following backward algorithm pricing American option:
Step 1: Initiate $p_T = f_T$
Step 2: Compute conditional expectation by regression:

$$g_{t-1} = \arg\min_{h\in\mathcal{L}_0} \mathbb{E}_Q\left[\left(h(S_{t-1}^1) - (1+r)^{-1}p_t\right)^2\right] \tag{4}$$

Step 4: Compute Snell Envelope

$$\begin{aligned} p_{t-1} &= \max\{f_{t-1}, g_{t-1}(S_{t-1}^1)\} \\ &= f_{t-1} \cdot \mathbb{1}_{\left\{f_{t-1} \le g_{t-1}(S_{t-1}^1)\right\}} + g_{t-1}(S_{t-1}^1) \cdot \mathbb{1}_{\left\{f_{t-1} \ge g_{t-1}(S_{t-1}^1)\right\}} \\ &= f_{t-1} \cdot \mathbb{1}_{\left\{f_{t-1} \le g_{t-1}(S_{t-1}^1)\right\}} + \mathbb{E}_Q[(1+r)^{-1}p_t|S_{t-1}^1] \cdot \mathbb{1}_{\left\{f_{t-1} \ge g_{t-1}(S_{t-1}^1)\right\}} \end{aligned}$$

(c) Prove that if we replace Q with path samples $\mathbb{S}$, $\mathcal{H}$ with shallow neural networks, $\mathbb{E}_Q[(1+r)^{-1}p_t|S_{t-1}^1]$ with unbiased estimator $(1+r)^{-1}p_t$, and keep the optimal stopping formula at the final step

$$p_0 = \max\{g(x_0), \frac{1}{m}\sum_{i=m+1}^{2m}(1+r)^{-1}p_1^i\}. \tag{5}$$

Then we obtain Algorithm 1 in [3] after a special train-test split.

Remark: We only learn the optimal stopping decision but not the stopped prices. Also we do regression on training data set ($i = 1, \ldots, m$) and calculate price on testing data set ($i = m+1, \ldots, 2m$), see [3].

Exercise 12.2 Understand the code of [3] on Github.

References

[1] Dimitri Bertsekas. *Reinforcement learning and optimal control.* Athena Scientific, 2019.

[2] Robert J. Elliott and P. Ekkehard Kopp. *Mathematics of Financial Markets.* Springer, 2005.

[3] Calypso Herrera, Florian Krach, Pierre Ruyssen, and Josef Teichmann. Optimal stopping via randomized neural networks. *arXiv preprint arXiv:2104.13669*, 2021.

[4] Francis A Longstaff and Eduardo S Schwartz. Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, 14(1):113–147, 2001.

[5] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

[6] John N Tsitsiklis and Benjamin Van Roy. Regression methods for pricing complex american-style options. *IEEE Transactions on Neural Networks*, 12(4):694–703, 2001.