Sorbonne Université
Période 2

Master M2 PF - PMA
Machine Learning pour les produits dérivés

Examen du 25 mars 2025

Exercise 1 (Deep Optimal Stopping):
We consider a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and a discrete-time filtration $(\mathcal{F}_n)_{n=0,\dots,N}$. On this filtered space, we consider an inhomogeneous Markov chain $(X_n)_{n=0,\dots,N}$ taking values in $\mathbf{R}_+^d$ (with large d) with a deterministic initial condition $X_0 = x_0 \in \mathbf{R}_+^d$, and a transition kernel defined at time $n = 1, \dots, N$ by $P_n(x, A) = \mathbf{P}[X_n \in A | X_{n-1} = x]$ for $A \in \mathcal{B}(\mathbf{R}^d)$. Let $(\mathcal{F}_n)_{n=0,\dots,N}$ the natural filtration of $(X_n)_{n=0,\dots,N}$.

We assume that $(X_n)_{n=0,\dots,N}$ is a martingale under $\mathbf{P}$ and consider an American derivative $(Z_n = \varphi_n(X_n))_{n=0,\dots,N}$ that can be exercised at any date n. We assume that $\varphi_n : \mathbf{R}_+^d \to [0, K]$ (K is a deterministic constant), meaning $Z_n \in [0, K]$.

Let $\mathcal{T}_0$ be the set of stopping times taking values in $\{0, \dots, N\}$. We aim to solve $V_0 = \sup_{\tau \in \mathcal{T}_0} \mathbf{E}[Z_\tau]$ and determine τ^* that achieves this supremum.

1. Provide the backward recurrence relation for the sequence of value functions $(V_n)_{n=0,\dots,N}$ that theoretically allows computing $V_0(x_0) = V_0$. This relation involves the transition kernel of the chain $(X_n)_{n=0,\dots,N}$.

2. Define the continuation function at $n = N$ by $C_N = 0$, and at $n = 0, \dots, N-1$ by $C_n(x) = P_{n+1} V_{n+1}(x)$. Introduce $\tau_n = \min\{k \geq n, V_k(X_k) = \varphi_k(X_k)\}$. Provide the argument justifying the representation:

$$C_n(x) = \mathbf{E}\big[Z_{\tau_{n+1}} | X_n = x\big].$$

$S_n = \operatorname{esssup} \mathbf{E}[Z_\tau | \mathcal{F}_n]$

3. Deduce that the random variables $(\tau_n)_{n=0,\dots,N}$ satisfy the recurrence relation:

$$\begin{cases} \tau_N = N, \\ \tau_n = n B_n + \tau_{n+1}(1 - B_n), \quad n = N-1, \dots, 0 \end{cases} \quad \text{with} \quad B_n = \mathbf{1}_{Z_n \geq \mathbf{E}[Z_{\tau_{n+1}} | X_n]},$$

and that $V_0(x_0) = \mathbf{E}[Z_{\tau_0}] = \max\big(\varphi_0(x_0), \mathbf{E}[Z_{\tau_1}]\big)$.

Using the notations from the course, we introduce a Feedforward Neural Network (multi-layer perceptron) of fixed depth $L \geq 2$ with m neurons per hidden layer, i.e. the parametric function $x \in \mathbf{R}^d \mapsto \Phi_m(x; \theta) \in \mathbf{R}$ where $\theta \in \Theta_m$. Additionally, we introduce the parametric function f_m taking values in $\{0, 1\}$ defined by adding an indicator activation function at the output of the network Φ_m

$$\forall \theta \in \Theta_m, \quad f_m(x; \theta) = \mathbf{1}_{]0,+\infty[} \circ \Phi_m(x; \theta)$$

Using this parametric function f_m, we define the following pseudo-algorithm:

$$\begin{cases} \bar{\tau}_N^{(m)} = N, \\ \bar{\tau}_n^{(m)} = n f_m(X_n; \bar{\theta}_n^{(m)}) + \bar{\tau}_{n+1}^{(m)}\big(1 - f_m(X_n; \bar{\theta}_n^{(m)})\big), \quad n = N-1, \dots, 0 \end{cases} \tag{1}$$
$$\text{with} \quad \bar{\theta}_n^{(m)} = \operatorname{argmax}_{\theta \in \Theta_m} \mathbf{E}\Big[Z_n f_m(X_n; \theta) + Z_{\bar{\tau}_{n+1}^{(m)}}\big(1 - f_m(X_n; \theta)\big)\Big].$$

The goal of the following questions is to prove that for all $n = 0, \dots, N$, we have convergence in $\mathbf{L}^2$ of $E_n^{(m)} = \mathbf{E}\big[Z_{\bar{\tau}_n^{(m)}} - Z_{\tau_n} | \mathcal{F}_n\big]$ to 0 as m tends to infinity.

4. Show that for all $n = 0, \ldots, N-1$, $E_n^{(m)} = A_n^{(m)} + \mathbf{E}\big[E_{n+1}^{(m)}|X_n\big](1 - f_m(X_n; \bar{\theta}_n^{(m)}))$ where

$$A_n^{(m)} = \big(Z_n - \mathbf{E}[Z_{\tau_{n+1}}|\mathcal{F}_n]\big)\big(f_m(X_n; \bar{\theta}_n^{(m)}) - B_n\big)$$

and deduce that the convergence in $\mathbf{L}^2$ to 0 of $E_n^{(m)}$ follows from that of $A_n^{(m)}$ (for fixed n, as m tends to infinity).

$Z_n B_n - \mathbb{E}[Z_\tau]B_m - Z_n f_m$

5. Show that $A_n^{(m)} \leq 0$.

ce qui ne bouge pas

6. Show that

$$\begin{aligned}|A_n^{(m)}| \leq & Z_n B_n + \mathbf{E}\big[Z_{\bar{\tau}_{n+1}^{(m)}}|X_n\big](1 - B_n)\\ & - \Big(Z_n f_m(X_n; \bar{\theta}_n^{(m)}) + \mathbf{E}\big[Z_{\bar{\tau}_{n+1}^{(m)}}|X_n\big](1 - f_m(X_n; \bar{\theta}_n^{(m)}))\Big)\\ & + \mathbf{E}\big[E_{n+1}^{(m)}|\mathcal{F}_n\big](B_n - f_m(X_n; \bar{\theta}_n^{(m)})).\end{aligned}$$

We now introduce $\tilde{\theta}_n^{(m)}$ which realizes

$$\tilde{\theta}_n^{(m)} = \operatorname{argmin}_{\theta \in \Theta_m} \mathbf{E}\Big[\big(B_n - f_m(X_n; \theta)\big)^2\Big].$$

7. Recall the universal approximation result in $\mathbf{L}^2$ seen in class and deduce that $(f_m(X_n; \tilde{\theta}_n^{(m)}))_{m \geq 2}$ converges to B_n in $\mathbf{L}^2$.

8. Show that

$$\mathbf{E}\big[|A_n^{(m)}|\big] \leq \mathbf{E}\Big[\big(Z_n - \mathbf{E}[Z_{\bar{\tau}_{n+1}^{(m)}}|X_n]\big)\big(B_n - f_m(X_n; \tilde{\theta}_n^{(m)})\big)\Big] + 2\mathbf{E}\big[|E_{n+1}^{(m)}|\big].$$

9. Deduce the convergence in $\mathbf{L}^1$ and then in $\mathbf{L}^2$ of $A_n^{(m)}$.

10. From the pseudo-algorithm (1), write an implementable algorithm based on a set of M independent trajectories $(X_n^{(j)})_{n=0,\ldots,N}$ for $1 \leq j \leq M$ to approximate $V_0(x_0)$. What errors are committed?

$Z_n B_n - Z_n f_m - \mathbb{E}[Z_\tau]B_m + \mathbb{E}[Z_\tau]f_m + 2\mathbb{E}[|E_{n+1}^{(m)}|]$