

Méthode de Longstaff-Schwartz pour l'évaluation d'options américaines

Martin Duguey, Chloé Amar

24 septembre 2023

Introduction

Dans le cadre du module *Méthodes de Monte-Carlo en finance*, nous nous intéressons à la valorisation d'options américaines par simulation. Plus particulièrement, nous proposons dans ce rapport une étude approfondie des méthodes utilisées dans les articles [3] et [2]. Afin de compléter notre démarche, nous nous sommes également appuyés sur les notes de cours du module *Méthodes probabilistes avancées* [1]. La première partie de ce rapport sera consacrée à la présentation du cadre général et au développement des aspects théoriques évoqués dans les articles, permettant de répondre à la problématique. La seconde partie s'attachera à l'étude de la convergence des algorithmes. De la même façon que dans les articles, elle sera démontrée en trois niveaux. Enfin, une dernière partie présentera les résultats donnés par les algorithmes dans des cas concrets, et les comparera notamment à ceux de l'article [3].

1 Méthode de Longstaff-Schwartz

Dans cette partie on explicite les différentes méthodes utilisées dans les articles [3] et [2] et leur cohérence dans une approche de simulation par méthode de Monte-Carlo.

Dans la suite, on considère un problème d'option américaine sur un sous-jacent S , supposé markovien, dont la diffusion n'est pas spécifiée, et versant un payoff $g(S_t)$, où g est une fonction supposée régulière. La durée d'exercice est notée T , c'est aussi la maturité puisqu'on souhaite valoriser l'option américaine au temps 0. On considère qu'il n'existe pas d'opportunité d'arbitrage et que le marché est complet. Il y a en particulier existence et unicité de la probabilité risque-neutre. Dans toute la suite, on se placera sous cette probabilité, notée \mathbb{Q} .

1.1 Cadre général

Alors qu'une option européenne ne donne à son détenteur le droit d'exercer qu'à un instant fixé T et d'obtenir le payoff, l'option américaine correspondante lui donne ce droit à tout instant $t \in [0, T]$. Afin de maximiser l'espérance de profit, le problème se réduit à déterminer un temps d'arrêt optimal pour excercer l'option. Soit $\mathcal{T}_{s,T}$ l'ensemble des temps d'arrêt dans $[s, T]$. On cherche le temps d'arrêt optimal, c'est-à-dire le temps où excercer permet de maximiser son espérance de profit. Le problème se réduit donc à déterminer :

$$\sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E} \left[e^{-\int_0^\tau r_s ds} g(S_\tau) \right]$$

avec $(r_s)_{s \geq 0}$ le taux sans risque (possiblement stochastique). Dans la suite, on le supposera positif et on notera $\mathcal{T} = \mathcal{T}_{0,T}$.

Proposition 1.1. Notons $u(t, S_t) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} \left[e^{-\int_t^\tau r_s ds} g(S_\tau) \mid \mathcal{F}_t \right] = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} \left[e^{-\int_t^\tau r_s ds} g(S_\tau) \mid S_t \right]$. Le supremum est alors obtenu pour le temps d'arrêt

$$\tau_t^* = \inf \{s \geq t : u(s, S_s) = g(S_s)\} \wedge T$$

Le prix de l'option américaine à la date t est donc donné par $\mathbb{E}[e^{-\int_t^{\tau_t^*} r_s ds} g(S_{\tau_t^*}) \mid \mathcal{F}_t]$.

En particulier, le prix à la date 0 est $\mathbb{E}[e^{-\int_0^{\tau_0^*} r_s ds} g(S_{\tau_0^*})]$.

Le problème qui se pose pour les options américaines est l'absence de solution analytique car le prix dépend de la trajectoire suivie par le sous-jacent entre l'instant 0 et la maturité T . Les options américaines sont donc *path-dependent*, ce qu'il est possible de contourner en ayant recours à des méthodes de Monte-Carlo.

L'idée principale d'une méthode de Monte Carlo pour la valorisation d'option est de simuler, si la loi est explicite, ou alors d'approcher par un schéma numérique, un grand nombre de trajectoires pour le processus sous-jacent et de calculer empiriquement une moyenne des payoffs obtenus pour chaque trajectoire. Cette méthode repose sur la loi forte des grands nombres, qui assure la convergence de la moyenne empirique vers le vrai prix (non actualisé) sous des hypothèses relativement simples. C'est une adaptation de cette méthode à notre problème que l'on cherche à appliquer et qui est proposée dans les articles [2] et [3].

1.1.1 Problème de programmation dynamique

Notons $(V_t)_{0 \leq t \leq T}$ le processus de valorisation de l'option américaine. Par définition, on a donc que

$$V_t = \mathbb{E}[e^{-\int_t^{\tau_t^*} r_s ds} g(S_{\tau_t^*}) | \mathcal{F}_t]$$

Proposition 1.2. *Le processus $(V_t)_{0 \leq t \leq T}$ est solution du problème de programmation dynamique suivant*

$$\begin{cases} V_T = g(S_T) \\ V_t = \max(g(S_t), \mathbb{E}[e^{-\int_t^{t+\delta t} r_s ds} V_{t+\delta t} | \mathcal{F}_t]) \quad \forall t \in [0, T[\end{cases} \quad (1)$$

Démonstration. On propose une démonstration intuitive essentiellement basée sur le caractère markovien du sous-jacent S .

À l'instant $t = T$: Remarquons que si l'option n'a pas été exercé avant la date finale T , elle vaudra en T le payoff $g(S_T)$. Par absence d'opportunité d'arbitrage, il est naturel d'avoir $V_T = g(S_T)$.

À l'instant $t < T$: L'acheteur de l'option peut choisir entre exercer l'option, et recevoir le payoff $g(S_t)$, ou alors ne pas l'exercer, dans ce cas le faire à une date postérieure, ce qui revient à recevoir à la date $t + \delta t$, le montant $\mathbb{E}[e^{-\int_t^{t+\delta t} r_s ds} V_{t+\delta t} | \mathcal{F}_t]$, qui représente les flux de cash générés par l'option à cette date.

On a donc que $V_t = \max(g(S_t), \mathbb{E}[e^{-\int_t^{t+\delta t} r_s ds} V_{t+\delta t} | \mathcal{F}_t])$.

□

En pratique, il n'est pas possible de mettre en place le problème 1 en temps continu. Une approche usuelle de la littérature est de considérer une grille fixée de temps d'exercice pour l'option. On ne parle plus dans ce cas d'option américaine mais d'option bermudéenne.

1.1.2 Approximation par option bermudéenne

Une option bermudéenne donne à son détenteur le droit d'exercer à chaque instant t compris dans un ensemble \mathcal{R} , préalablement fixé, de temps d'exercice dans l'intervalle $[0, T]$. On suppose que $\mathcal{R} := \{0 = t_0 < t_1 < \dots < t_K = T\}$. Ce type d'option est donc à mi-chemin entre les option européennes, dont l'exercice est uniquement possible en T , et les options américaines, dont l'exercice est possible à chaque instant t de l'intervalle $[0, T]$. On peut alors définir le prix de l'option bermudéenne par :

$$\sup_{\tau \in \mathcal{R}} \mathbb{E} \left[e^{-\int_0^\tau r_s ds} g(S_\tau) \right]$$

À travers cette définition, nous pouvons remarquer que deux cas se présentent. En effet, soit le temps d'arrêt optimal pour exercer l'option américaine fait partie de l'ensemble \mathcal{R} des temps d'exercices de l'option bermudéenne associée, soit il n'en fait pas partie. Dans tous les cas, le temps d'exercice de l'option bermudéenne associée est sous-optimal pour l'option américaine, c'est-à-dire que :

$$\sup_{\tau \in \mathcal{R}} \mathbb{E} \left[e^{-\int_0^\tau r_s ds} g(S_\tau) \right] \leq \sup_{\tau \in \mathcal{T}} \mathbb{E} \left[e^{-\int_0^\tau r_s ds} g(S_\tau) \right]$$

Proposition 1.3. *Notons $(V_t^b)_{0 \leq t \leq T}$ le processus de valorisation de l'option bermudéenne.*

Soit $\mathcal{R} := \{0 = t_0 < t_1 < \dots < t_K = T\}$, l'ensemble des temps d'exercice de l'option. Le processus $(V_t^b)_{0 \leq t \leq T}$ est solution du problème de programmation dynamique suivant

$$\begin{cases} V_T^b = g(S_T) \\ V_{t_k}^b = \max(g(S_{t_k}), \mathbb{E}[e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b | \mathcal{F}_{t_k}]) \quad \forall k \in \{0, \dots, K-1\} \end{cases} \quad (2)$$

En pratique, on peut estimer le prix d'une option américaine en enrichissant l'ensemble \mathcal{R} . Intuitivement, plus on se donne de temps d'exercice possibles pour l'option bermudéenne, plus celle-ci se rapproche d'une option américaine. Ainsi, on tend vers un problème continu à mesure que l'on raffine la grille \mathcal{R} .

L'option américaine donnant plus de droits que l'option bermudéenne correspondante, le gain espéré par le détenteur de l'option est plus grand dans le cas américain que bermudéen, donc le prix de l'option américaine est plus élevé. Par conséquent, le prix de l'option bermudéenne tend par valeurs inférieures vers le prix de l'option américaine, ce qui sera prouvé rigoureusement dans la seconde partie. Notons que l'on peut également obtenir des majorations du prix des options américaines, en passant par une approche dite duale, qui ne sera pas développée dans ce rapport. Ainsi, dans toute la suite, on développera des méthodes d'évaluation d'options bermudéennes, ce qui permettra de résoudre le problème en temps discret.

En reprenant la grille de temps \mathcal{R} définie précédemment, on a donc que :

$$V_{t_k}^b = \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right]$$

avec

$$\tau_k := \inf\{j \geq k \mid t_j \in \mathcal{R}, V_{t_j}^b = g(S_{t_j})\}$$

1.2 Algorithmes de résolution numérique

Dans cette partie, nous présenterons deux méthodes de résolution du problème de valorisation d'option bermudéenne après avoir présenté une méthodologie permettant d'approximer l'espérance conditionnelle, nécessaire à la mise en place de ces deux algorithmes de résolution numérique.

Les deux méthodes proposées sont des méthodes de Monte Carlo, donc le processus de prix du sous-jacent, sera simulé un grand nombre de fois. Notons N ce nombre, on notera alors $S^i = (S_{t_k}^i)_{t_k \in \mathcal{R}}$ la i -ème trajectoire simulée pour le spot. On notera $S = (S^i)_{i \in \llbracket 1, N \rrbracket}$ de dimension $N \times K$. De la même manière, le processus de valorisation sera perçu comme la matrice $V^b = (V_{t_k}^{i,b})_{t_k \in \mathcal{R}, i \in \llbracket 1, N \rrbracket}$ également de dimension $N \times K$. Enfin, on considérera les temps d'arrêt $(\tau^i)_{i \in \llbracket 1, N \rrbracket}$, vérifiant $\forall t_k \in \mathcal{R}, \tau_k^i = \inf\{j \geq k \mid t_j \in \mathcal{R}, V_{t_j}^{i,b} = g(S_{t_j}^i)\}$.

La stratégie optimale d'exercice ne dépendant que du choix entre l'exercice du payoff ou pas, le détenteur de l'option compare le payoff reçu instantanément avec le montant moyen qui sera reçu dans le futur si l'option continue de vivre, caractérisé par l'espérance conditionnelle dans le problème 2. Le calcul de ce dernier terme est le point clé des méthodes que nous allons présenter.

1.2.1 Calcul de l'espérance conditionnelle

Soit $F^i(t_j, t_k)$ le flux généré à la date $t_j > t_k$ par l'option sur le sous-jacent S^i si l'option n'a pas été exercée avant t_k . Ce flux peut correspondre, par définition du prix de l'option bermudéenne sous forme de solution au problème 2, au payoff en t_j ou à la valeur de continuation à ce même instant. On peut alors écrire la valeur de continuation pour le i -ème sous-jacent simulé comme suit :

$$C_{t_k}^i = \mathbb{E} \left[\sum_{j=k+1}^K e^{-\int_{t_k}^{t_j} r_s ds} F^i(t_j, t_k) | \mathcal{F}_{t_k} \right]$$

En effet, la valeur de continuation représente, à la date t_k , la valeur moyenne des flux futurs actualisés perçus par le détenteur de l'option s'il décide de ne pas l'exercer à cette date. Par conséquent, si l'option est exercée à une date $t_m \in [t_{k+1}, t_K]$, on considérera que ces flux sont nuls après t_m .

L'idée de la méthode est de considérer que $\forall t_k \in \mathcal{R}, C_{t_k}^i$ peut être représenté comme une combinaison linéaire des informations disponibles sur le sous-jacent à la date t_k . Cette hypothèse est justifiée par le fait qu'une espérance conditionnelle est la projection orthogonale sur un espace L^2 . En effet, rappelons que pour $Y \in L^2(\Omega, \mathcal{F}, \mathbb{Q})$, si $X \in L^2(\Omega, \mathcal{G}, \mathbb{Q})$ pour $\mathcal{G} \subset \mathcal{F}$, alors $\mathbb{E}(Y|X)$ est la projection orthogonale de Y sur $L^2(\Omega, \mathcal{G}, \mathbb{Q})$. Par définition de la projection orthogonale, on a donc que $\|Y - \mathbb{E}(Y|X)\|_{L^2}^2 \leq \|Y - U\|_{L^2}^2 \quad \forall U \in L^2(\Omega, \mathcal{G}, \mathbb{Q})$. Dans notre cas, on a donc que $\forall t_k \in \mathcal{R}$

$$C_{t_k}^i = \operatorname{argmin}_{U \in L^2(\mathcal{F}_{t_k})} \mathbb{E} \left[\left\| \sum_{j=k+1}^K e^{-\int_{t_k}^{t_j} r_s ds} F^i(t_j, t_k) - U \right\|^2 \right]$$

Le prix du sous-jacent étant par hypothèse un processus markovien, on a que $\forall t_k \in \mathcal{R}, \mathcal{F}_{t_k} = \sigma(S_{t_k})$, donc

$$C_{t_k}^i = \operatorname{argmin}_{f \in L^2(\mathcal{F}_{t_k})} \mathbb{E} \left[\left| \sum_{j=k+1}^K e^{-\int_{t_k}^{t_j} r_s ds} F^i(t_j, t_k) - f(S_{t_k}^i) \right|^2 \right]$$

Autrement dit, en regroupant les flux de cash futurs, on a que

$$C_{t_k}^i = \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} \mid \mathcal{F}_{t_k} \right]$$

et ainsi

$$C_{t_k}^i = \operatorname{argmin}_{f \in L^2(\mathcal{F}_{t_k})} \mathbb{E} \left[\left| e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} - f(S_{t_k}^i) \right|^2 \right]$$

Comme $L^2(\mathcal{F}_{t_k})$ est un espace de Hilbert, il possède une base orthonormée, ce qui permet de représenter l'espérance conditionnelle comme une fonction linéaire des éléments de cette base. On peut donc écrire

$$C_{t_k}^i = \sum_{j=0}^{\infty} a_{j,i}^{(k)} L_j(S_{t_k}^i)$$

avec $(L_j)_{j \geq 0}$ une base de polynôme et $(a_{j,i}^{(k)})_{j \geq 0}$ une suite de coefficients constants. Elle dépend du temps t_k et de la i -ème trajectoire simulée pour le spot. Dans la suite, on notera $\forall t_k \in \mathcal{R}, C_{t_k} = (C_{t_k}^i)_{i \in \llbracket 1, N \rrbracket}, L_j(S_{t_k}) = (L_j(S_{t_k}^i))_{i \in \llbracket 1, N \rrbracket}$ et $a_j^{(k)} = (a_{j,i}^{(k)})_{i \in \llbracket 1, N \rrbracket}$, les vecteurs de taille N correspondant à ces quantités pour chaque trajectoire $i \in \llbracket 1, N \rrbracket$. On a ainsi que

$$C_{t_k} = \begin{pmatrix} \sum_{j=0}^{\infty} a_{j,1}^{(k)} L_j(S_{t_k}^1) \\ \vdots \\ \sum_{j=0}^{\infty} a_{j,N}^{(k)} L_j(S_{t_k}^N) \end{pmatrix} = \sum_{j=0}^{\infty} a_j^{(k)} L_j(S_{t_k})$$

L'impact du choix de la base de polynôme sera discuté dans la dernière partie. Il est en effet possible, d'après l'article [3] d'utiliser une grande variété de polynômes, comme par exemple les polynômes de Laguerre, Legendre, Hermite, Chebyshev ou encore Jacobi. Comme nous le montrerons, de simples puissances du spot $L_j(x) = x^j|_{x=S_{t_k}^i}$ permettent d'obtenir de bons résultats.

En pratique, considérer un tel ensemble de fonction est irréaliste, car trop grand ; on se restreint donc à une base de taille fixée, notée J , et on estime l'espérance conditionnelle en régressant sur les polynômes de degré au plus J de la base. On peut ainsi écrire la valeur de continuation obtenue à partir des polynômes de degré au plus J comme suit quelque soit $t_k \in \mathcal{R}$:

$$C_{t_k}^{a,[J]} = \sum_{j=0}^J a_j^{(k)} L_j(S_{t_k}) := a^{[J],[k]} \cdot L^{[J]}(S_{t_k})$$

Ainsi, $\forall t_k \in \mathcal{R}$, la suite $(a_j^{(k)})_{0 \leq j \leq J}$ est déterminée de façon à minimiser sur les $(\alpha_j^{(k)})_{0 \leq j \leq J}$ l'expression

$$\mathbb{E} \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b - \sum_{j=0}^J \alpha_j^{(k)} L_j(S_{t_k}) \right)^2 \right]$$

En pratique, on détermine, $\forall t_k \in \mathcal{R}$, la suite $(\hat{a}_j^{(k)})_{0 \leq j \leq J}$ minimisant ¹

1. Comme $V_{t_{k+1}}^b = \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mid \mathcal{F}_{t_{k+1}} \right]$, on a que $C_{t_k} = \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mid \mathcal{F}_{t_k} \right]$, donc cela revient à minimiser

$$\frac{1}{N} \sum_{i=1}^N \left[\left(e^{-\int_{t_k}^{\tau_{k+1}^i} r_s ds} g(S_{\tau_{k+1}^i}^i) - \sum_{j=0}^J \alpha_{j,i}^{(k)} L_j(S_{t_k}^i) \right)^2 \right]$$

$$\frac{1}{N} \sum_{i=1}^N \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} - \sum_{j=0}^J \alpha_{j,i}^{(k)} L_j(S_{t_k}^i) \right)^2 \right]$$

D'après les auteurs, cette régression peut se faire en considérant uniquement les trajectoires in the money, ce qui permet de limiter la région sur laquelle on calcule l'espérance conditionnelle et par conséquent réduire le temps de calcul sans pour autant réduire la qualité de l'estimation. Nous en discuterons dans la dernière partie.

1.2.2 Algorithme Least Square Monte Carlo

Nous explicitions dans cette partie l'algorithme LSM complet, reposant sur le calcul de l'espérance conditionnelle détaillé précédemment. L'approche utilisée, présentée ci-dessous, repose sur une méthode de Monte Carlo.

Algorithm 1 Bermudean Option Pricing Algorithm

$S \leftarrow (S_0^i, \dots, S_T^i)_{0 \leq i \leq N}$

$V \leftarrow (0_{\mathbb{R}^N}, \dots, 0_{\mathbb{R}^N}, g(S_T^i)_{0 \leq i \leq N})$

for $k = K - 1 : 0$ **do**

$P \leftarrow 0_{\mathbb{R}^N}$

$C \leftarrow 0_{\mathbb{R}^N}$

Compute $(\hat{\alpha}_j^{(k)})_{0 \leq j \leq J} \in \underset{(\alpha^{(k)})_{0 \leq j \leq J}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} - \sum_{j=0}^J \alpha_{j,i}^{(k)} L_j(S_{t_k}^i) \right)^2 \right]$

for $i = 1 : N$ **do**

$C^{i,[J]} \leftarrow \sum_{j=0}^J \hat{\alpha}_{j,i}^{(k)} L_j(S_{t_k}^i)$

$P^i \leftarrow g(S_{t_k}^i)$

if $P^i > C^{i,[J]}$ **then**

$V_{t_k}^i \leftarrow P^i$

else

$V_{t_k}^i \leftarrow e^{-\int_{t_k}^{t_{k+1}} r_s ds} C^{i,[J]}$

end if

end for

end for

$P_{am} \leftarrow \frac{1}{N} \sum_{i=1}^N e^{-\int_{t_0}^{t_1} r_s ds} V_{t_0}^i$

return P_{am}

Cet algorithme est donné de manière à présenter la méthode de façon la plus intuitive possible, mais on peut remarquer sa sous-optimalité. On l'implémentera donc en pratique de façon plus optimale.

1.2.3 Approche par temps d'arrêts

Sur la base de la démarche de Lonstaff et Schwartz, l'article [2] détaille une autre méthodologie pour l'évaluation d'options bermudéennes. Le point de départ des auteurs, Clément, Lamberton et Protter, est de réécrire le problème de programmation dynamique sur le processus de valorisation de l'option comme un problème de programmation dynamique

sur les temps d'arrêt.

Rappelons que $V_{t_k}^b = \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right]$ avec $\tau_k := \inf\{j \geq k \mid t_j \in \mathcal{R}, V_{t_j}^b = g(S_{t_j})\}$.

On avait également défini $C_{t_k} = \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b | \mathcal{F}_{t_k} \right] = \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right]$ par définition de $V_{t_{k+1}}^b$.

On explicite alors une suite discrète de temps d'arrêt qui est solution du problème de programmation dynamique.

Proposition 1.4. *Le problème 2 est équivalent au problème*

$$\begin{cases} \tau_K = t_K \\ \tau_k = t_k \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + \tau_{k+1} \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}}, \quad \forall k \in \{0, \dots, K-1\} \end{cases} \quad \text{avec } C_{t_k} = \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b | \mathcal{F}_{t_k} \right] \quad (3)$$

et le prix de l'option à la date t_k est donné par $V_{t_k}^b = \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right] \quad \forall k \in \{0, \dots, K\}$.

Démonstration. On raisonne par récurrence rétrograde.

Initialisation : Pour $k = K$, on a que $\tau_K = t_K = T$. Par définition du prix, on a $V_T^b = g(S_T) = \mathbb{E} [g(S_T) | \mathcal{F}_T]$, et puisque $\tau_K = T$, on a bien $V_T^b = \mathbb{E} [g(S_{\tau_K}) | \mathcal{F}_{\tau_K}]$.

Hérédité : Soit $k < K$ et on suppose l'hypothèse vraie au rang $k+1$: $V_{t_{k+1}}^b = \mathbb{E} \left[e^{-\int_{t_{k+1}}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_{k+1}} \right]$. D'après le problème 2, on a :

$$\begin{aligned} V_{t_k}^b &= \max \left(g(S_{t_k}), \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b | \mathcal{F}_{t_k} \right] \right) \\ &= \max \left(g(S_{t_k}), \mathbb{E} \left[e^{-\int_{t_k}^{t_{k+1}} r_s ds} \mathbb{E} \left[e^{-\int_{t_{k+1}}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_{k+1}} \right] | \mathcal{F}_{t_k} \right] \right) \quad \text{par hypothèse de récurrence} \\ &= \max \left(g(S_{t_k}), \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \right) \quad \text{car } \mathcal{F}_{t_k} \subset \mathcal{F}_{t_{k+1}} \\ &= g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} \\ &= \mathbb{E} \left[g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right] \\ &= \mathbb{E} \left[g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right] \quad \text{car sur } \{g(S_{t_k}) < C_{t_k}\}, \tau_{k+1} = \tau_k \\ &= \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right] \quad \text{car sur } \{g(S_{t_k}) \geq C_{t_k}\}, \tau_k = t_k \\ &= \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right] \quad \text{par définition de } \tau_k \end{aligned}$$

L'hypothèse de récurrence est donc vraie au rang k et la propriété est vérifiée. \square

La problématique est donc la même que dans l'article [3], il faut calculer $\left(\mathbb{E}\left[V_{t_{k+1}}^b|\mathcal{F}_{t_k}\right]\right)_{0\leq k\leq K}$ pour pouvoir construire notre suite $(\tau_k)_{0\leq k\leq K}$ et déterminer, grâce au problème 4 :

$$V_0^b := \sup_{0\leq k\leq K} \mathbb{E}\left[e^{-\int_0^{\tau_k} r_s ds} g(S_{\tau_k})\right] = \mathbb{E}\left[e^{-\int_0^{\tau_0} r_s ds} g(S_{\tau_0})\right]$$

C'est une approximation équivalente à celle de l'article [2], puisque nous choisissons d'effectuer notre récurrence sur les temps d'arrêt jusqu'au temps initial de valorisation $t_0 = 0$, mais on aurait très bien pu obtenir la formulation de l'article en étudiant les temps d'exercice possibles juste après le temps $t_0 = 0$, donc au temps t_1 . Dans ce cas on aurait alors obtenu, à un coefficient d'actualisation près, le résultat de l'article, c'est-à-dire le maximum entre l'exercice au temps initial et l'espérance de gain futur si l'exercice se fait au temps τ_1 .

En pratique, on calcule à chaque instant l'espérance conditionnelle, par les méthodes présentées dans 1.2.1 et on compare cette valeur à $g(S_{t_k})$. Dans ce contexte, on introduit la suite $(\hat{\tau}_k^{[J]})_{0\leq k\leq K}$ définie comme suit :

$$\begin{cases} \hat{\tau}_K^{[J]} = t_K \\ \hat{\tau}_k^{[J]} = t_k \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} + \hat{\tau}_{k+1}^{[J]} \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}}, \quad \forall k \in \{0, \dots, K-1\} \end{cases} \quad (4)$$

où $C_{t_k}^{[J]}$ est la valeur de continuation estimée à partir des J premiers polynômes de la base choisie. On peut alors évaluer une option bermudéenne en suivant l'algorithme présenté ci-dessous, encore une fois de manière sous-optimale.

Algorithm 2 Bermudean Option Pricing Algorithm - Stopping Times

$S \leftarrow (S_0^i, \dots, S_T^i)_{0\leq i\leq N}$

$V \leftarrow (0_{\mathbb{R}^N}, \dots, 0_{\mathbb{R}^N}, g(S_T^i))_{0\leq i\leq N}$

for $k = K - 1 : 0$ **do**

$\tau \leftarrow (0_{\mathbb{R}^N}, \dots, 0_{\mathbb{R}^N}, (t_k)_{0\leq i\leq N})$

$P \leftarrow 0_{\mathbb{R}^N}$

$C \leftarrow 0_{\mathbb{R}^N}$

Compute $(\hat{\alpha}_j^{(k)})_{0\leq j\leq J} \in \underset{(\alpha^{(k)})_{0\leq j\leq J}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} - \sum_{j=0}^J \alpha_{j,i}^{(k)} L_j(S_{t_k}^i) \right)^2 \right]$

for $i = 1 : N$ **do**

$C^{i,[J]} \leftarrow \sum_{j=0}^J \hat{\alpha}_{j,i}^{(k)} L_j(S_{t_k}^i)$

$P^i \leftarrow g(S_{t_k}^i)$

if $P^i \geq C^{i,[J]}$ **then**

$\hat{\tau}_k^i \leftarrow t_k$

$V_{t_k}^i \leftarrow P^i$

else

$\hat{\tau}_k^i \leftarrow \hat{\tau}_{k+1}^i$

end if

end for

end for

$P_{am} \leftarrow \frac{1}{N} \sum_{i=1}^N e^{-\int_{t_0}^{\hat{\tau}_0^i} r_s ds} V_{\hat{\tau}_0^i}^i$

return P_{am}

2 Résultats de convergence

Dans la première partie, nous avons présenté les méthodes d'évaluation d'option bermudéennes comme étudiées dans les articles [2] et [3], en détaillant leur mise en place numérique et notamment le caractère bien défini des algorithmes 2 et 1. Dans cette partie, on souhaite vérifier la convergence des algorithmes. Dans un premier temps, nous présenterons un résultat permettant de justifier l'approximation de l'espérance conditionnelle faite précédemment. Ensuite, nous aborderons le fait que par la simulation, on peut obtenir des résultats de convergence vers le prix de l'option bermudéenne. Enfin dans une dernière section, et toujours dans un contexte de simulation, nous justifierons la convergence du prix d'une option bermudéenne vers celui d'une option américaine lorsque le pas de discrétisation des temps d'exercice de l'option bermudéenne converge vers 0.

2.1 Convergence de l'estimateur de l'espérance conditionnelle

On s'intéresse dans cette partie à justifier la convergence de l'estimateur de l'espérance conditionnelle vers la valeur théorique lorsque le nombre de fonctions de la base tend vers l'infini. Comme nous avons pu l'expliquer précédemment, le calcul numérique à travers les algorithmes 1 et 2 nous permet d'obtenir la valeur :

$$V_0^{b,[J]} := \mathbb{E}[e^{-\int_0^{\hat{\tau}_0^{[J]}} r_s ds} g(S_{\hat{\tau}_0^{[J]}})]$$

Ainsi, prouver la convergence de l'estimateur de l'espérance conditionnelle revient à montrer que :

$$V_0^{b,[J]} \xrightarrow{J \rightarrow +\infty} V_0^b$$

Dans cette partie, on observe le processus² $(S_t)_{0 \leq t \leq T}$ et on considère qu'il vérifie les hypothèses de la première partie. On peut introduire le résultat suivant issu de l'article [2].

Théorème 2.1. *Quelque soit t_k dans \mathcal{R} , $(L_j)_{j \geq 0}$ une base orthonormée de l'espace de Hilbert $L^2(\mathcal{F}_{t_k})$ on a que*

$$\mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_k^{[J]}} r_s ds} g(S_{\hat{\tau}_k^{[J]}}) | \mathcal{F}_{t_k} \right] \xrightarrow{J \rightarrow +\infty} \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right] \text{ dans } L^2$$

Démonstration. On raisonne par récurrence rétrograde.

Initialisation : Pour $k = K$, on a bien que $\mathbb{E} \left[e^{-\int_{t_K}^{\hat{\tau}_K^{[J]}} r_s ds} g(S_{\hat{\tau}_K^{[J]}}) | \mathcal{F}_{t_K} \right] \xrightarrow{J \rightarrow +\infty} \mathbb{E} \left[e^{-\int_{t_K}^{\tau_K} r_s ds} g(S_{\tau_K}) | \mathcal{F}_{t_K} \right]$ dans L^2 car $\hat{\tau}_K^{[J]} = \tau_K$ par définition.

Hérédité : Soit $k < K$ et supposons l'hypothèse vraie au rang $k+1$:

$$\mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) | \mathcal{F}_{t_{k+1}} \right] \xrightarrow{J \rightarrow +\infty} \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_{k+1}} \right] \text{ dans } L^2$$

Notons dans toute la suite $A_k := \mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_k^{[J]}} r_s ds} g(S_{\hat{\tau}_k^{[J]}}) - e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right]$.

Par définition de $\hat{\tau}_k^{[J]}$, on a que³

$$e^{-\int_{t_k}^{\hat{\tau}_k^{[J]}} r_s ds} g(S_{\hat{\tau}_k^{[J]}}) = g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} + e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}}$$

2. Il n'y a donc à ce stade pas de simulation, le processus S est unidimensionnel

3. Comme dans la démonstration de la proposition 1.4, le résultat provient du fait que sur $\{g(S_{t_k}) < C_{t_k}\}$, $\hat{\tau}_{k+1}^{[J]} = \hat{\tau}_k^{[J]}$, et sur $\{g(S_{t_k}) \geq C_{t_k}\}$, $\hat{\tau}_k^{[J]} = t_k$

D'où :

$$A_k = \mathbb{E} \left[g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} + e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} | \mathcal{F}_{t_k} \right] - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right]$$

De la même façon, pour $C_{t_k} = \mathbb{E}[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k}]$ (cf proposition 1.4), on a que :

$$\mathbb{E} \left[e^{-\int_{t_k}^{\tau_k} r_s ds} g(S_{\tau_k}) | \mathcal{F}_{t_k} \right] = \mathbb{E} \left[g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} + e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right]$$

Donc :

$$\begin{aligned} A_k &= g(S_{t_k}) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) \\ &\quad + \mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} | \mathcal{F}_{t_k} \right] \\ &\quad - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right] \\ &= g(S_{t_k}) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) + \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} | \mathcal{F}_{t_k} \right] \\ &\quad - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} | \mathcal{F}_{t_k} \right] + \mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} | \mathcal{F}_{t_k} \right] \\ &\quad - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} | \mathcal{F}_{t_k} \right] \end{aligned}$$

En regroupant les termes, on obtient :

$$\begin{aligned} A_k &= g(S_{t_k}) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) + \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \left(\mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} \right) \\ &\quad + \mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) - e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} \end{aligned}$$

Or on remarque que :

$$\left(\mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}\}} \right) = \left(1 - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - (1 - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}}) \right) = - \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right)$$

Donc en notant $\delta_k := \mathbb{E} \left[e^{-\int_{t_k}^{\hat{\tau}_{k+1}^{[J]}} r_s ds} g(S_{\hat{\tau}_{k+1}^{[J]}}) - e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{[J]}\}}$, on a que

$$A_k = \left(g(S_{t_k}) - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \right) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) + \delta_k$$

On sait que δ_k converge vers 0 dans L^2 par hypothèse de récurrence. On définit dans la suite B_k comme :

$$\begin{aligned} B_k &:= \left(g(S_{t_k}) - \mathbb{E} \left[e^{-\int_{t_k}^{\tau_{k+1}} r_s ds} g(S_{\tau_{k+1}}) | \mathcal{F}_{t_k} \right] \right) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) \\ &= (g(S_{t_k}) - C_{t_k}) \left(\mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}\}} \right) \end{aligned}$$

Ainsi :

$$\begin{aligned}
|B_k| &= |g(S_{t_k}) - C_{t_k}| \mathbb{1}_{\{C_{t_k} > g(S_{t_k}) \geq C_{t_k}^{[J]}\}} - \mathbb{1}_{\{C_{t_k}^{[J]} > g(S_{t_k}) \geq C_{t_k}\}} \\
&\leq |g(S_{t_k}) - C_{t_k}| \mathbb{1}_{\{|g(S_{t_k}) - C_{t_k}| \leq |C_{t_k}^{[J]} - C_{t_k}|\}} \\
&\leq |C_{t_k}^{[J]} - C_{t_k}|
\end{aligned}$$

Donc $\|B_k\|_2 \leq \|A_{k+1}\|_2$ qui converge vers 0 par hypothèse de récurrence. Donc l'hypothèse de récurrence est vérifiée au rang k . \square

Le résultat précédent est valable quelque soit t_k dans \mathcal{R} . En prenant en particulier $t_0 = 0$, on a exactement

$$\mathbb{E} \left[e^{-\int_0^{\hat{\tau}_0^{[J]}} r_s ds} g(S_{\hat{\tau}_0^{[J]}}) \right] \xrightarrow{J \rightarrow +\infty} \mathbb{E} \left[e^{-\int_0^{\tau_0} r_s ds} g(S_{\tau_0}) \right]$$

On a donc montré que $V_0^{b,[J]}$ converge vers V_0^b dès lors que J tend vers l'infini. De cette façon et en dehors de tout contexte de simulation, on a bien justifié l'approximation faite pour la projection de l'espérance conditionnelle. C'est un résultat essentiel pour le calcul numérique puisqu'il nous permet de travailler avec une base orthonormée finie. La suite de la démarche consiste à vérifier que cette convergence peut être mise à profit dans un contexte de simulation.

2.2 Convergence de l'estimateur du prix de l'option bermudéenne vers le prix de l'option bermudéenne

On vient de montrer que l'on peut effectivement approcher l'espérance conditionnelle par une combinaison linéaire d'une base polynomiale. Jusque là, on ne considérait que des espérance théoriques. On souhaite désormais se placer dans un contexte de simulation, et simuler le processus de spot un nombre N de fois. On peut alors mentionner le résultat de la proposition 2 de l'article [3].

Celui-ci permet d'affirmer (sous certaines hypothèses) qu'un nombre fini d'élément de la base $(L_j)_{j \geq 0}$ permet d'obtenir la convergence en probabilité de l'estimateur LSM vers le prix de l'option bermudéenne lorsque N , le nombre de trajectoires simulées du processus S , tend vers l'infini.

Dans cette partie, on souhaite prouver concrètement la convergence des algorithmes présentés précédemment. On observe désormais une multitude de trajectoires du processus S , et on cherche à montrer que :

$$\frac{1}{N} \sum_{i=0}^N e^{-\int_0^{\hat{\tau}_0^{i,[J]}} r_s ds} g(S_{\hat{\tau}_0^{i,[J]}}) \xrightarrow{N \rightarrow +\infty} \mathbb{E} \left[e^{-\int_0^{\hat{\tau}_0^{[J]}} r_s ds} g(S_{\hat{\tau}_0^{[J]}}) \right]$$

Un tel résultat assurerait la convergence des algorithmes puisqu'en pratique le membre de gauche est exactement l'estimateur empirique de $V_0^{b,[J]}$ construit par méthode de Monte-Carlo.

Théorème 2.2. *Si $\forall t_k \in \mathcal{R} \setminus \{t_K\}$, $\mathbb{P}(C_{t_k} = g(S_{t_k})) = 0$, alors on a que*

$$\frac{1}{N} \sum_{i=0}^N e^{-\int_0^{\hat{\tau}_k^{i,[J]}} r_s ds} g(S_{\hat{\tau}_k^{i,[J]}}) \xrightarrow{N \rightarrow +\infty} \mathbb{E} \left[e^{-\int_0^{\hat{\tau}_k^{[J]}} r_s ds} g(S_{\hat{\tau}_k^{[J]}}) \right]$$

presque sûrement.

La démonstration repose sur deux lemmes.

Lemme 2.3. *Si quelque soit $t_k \in \mathcal{R} \setminus \{t_K\}$, on a que $\mathbb{P}(C_{t_k} = g(S_{t_k})) = 0$, alors lorsque $N \rightarrow +\infty$,*

$$(\hat{\alpha}_j^{(k)})_{0 \leq j \leq J} \in \underset{(\alpha_j^{(k)})_{0 \leq j \leq J}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^{i,b} - \sum_{j=0}^J \alpha_{j,i}^{(k)} L_j(S_{t_k}^i) \right)^2 \right]$$

converge presque sûrement vers

$$(a_j^{(k)})_{0 \leq j \leq J} \in \underset{(\alpha_j^{(k)})_{0 \leq j \leq J}}{\operatorname{argmin}} \mathbb{E} \left[\left(e^{-\int_{t_k}^{t_{k+1}} r_s ds} V_{t_{k+1}}^b - \sum_{j=0}^J \alpha_j^{(k)} L_j(S_{t_k}) \right)^2 \right]$$

La démonstration du lemme 2.3 repose essentiellement sur la loi des grands nombres.

Lemme 2.4. Soit $t_k \in \mathcal{R} \setminus \{t_K\}$, posons $C_{t_k}^{a, [J]} = \sum_{j=0}^J a_j^{(k)} L_j(S_{t_k}) := a^{[J], (k)} \cdot L^{[J]}(S_{t_k})$ et

$$D_{t_k} \left(a^{[J]}, S \right) = g(S_{t_k}) \mathbb{1}_{\{g(S_{t_k}) \geq C_{t_k}^{a, [J]}\}} + D_{t_{k+1}} \left(a^{[J]}, S \right) \mathbb{1}_{\{g(S_{t_k}) < C_{t_k}^{a, [J]}\}}$$

Alors :

$$|D_{t_k}(a^{[J]}, S) - D_{t_k}(b^{[J]}, S)| \leq \left(\sum_{t_j=t_k}^{t_K} |g(S_{t_j})| \right) \left(\sum_{t_j=t_k}^{t_{K-1}} \mathbb{1}_{\{|g(S_{t_k}) - C_{t_k}^{b, [J]}| \geq |a^{[J], (k)} - b^{[J], (k)}| |L^{[J]}(S_{t_k})|\}} \right)$$

Démonstration. En annexe. □

Le résultat de convergence du théorème 2.2 est directement issue du résultat de convergence des lemmes 2.3 et 2.4. On s'intéresse alors dans la suite à vérifier que le prix d'option bermudéenne converge bien vers le prix de l'option américaine correspondante.

2.3 Convergence du prix d'option bermudéenne vers le prix d'option américaine

Les sections précédentes ont permis de montrer la convergence des algorithmes d'évaluation d'options bermudéennes. Il reste désormais à montrer que le prix d'une option bermudéenne converge vers celui d'une option américaine à mesure que la grille \mathcal{R} se raffine, idée que nous avons présentée intuitivement dans la première partie.

Proposition 2.5. Soit $|\mathcal{R}| = \max_{1 \leq k \leq K} (t_k - t_{k-1})$. Si la fonction de payoff g est Lipschitz (resp. C_b^2), alors il existe $C > 0$ telle que :

$$\sup_{t \in [0, T]} \mathbb{E} [|V_0 - V_0^b|^2] \leq C |\mathcal{R}| \quad (\text{resp } C |\mathcal{R}|^2)$$

Démonstration. En annexe □

Nous avons ainsi démontré que le prix d'une option américaine peut être approximé par celui d'une option dont les dates d'exercice sont fixées, à condition que l'écart entre ces dates tende vers 0.

Finalement, nous avons montré que l'on pouvait approcher le prix d'une option américaine en simulant un grand nombre de trajectoires pour le spot, et en calculant le prix d'une option bermudéenne pour un grand nombre de temps d'exercice, le tout en utilisant l'approximation de l'espérance conditionnelle pour un degré J fini.

3 Résultats numériques

Dans toute cette partie, nous supposons que le processus de spot suit un modèle de Black-Scholes avec $r, \sigma > 0$ fixés. De cette façon nous pourrions comparer nos résultats avec l'article [3] et nous n'aurons pas besoin de mettre en place de schéma numérique pour simuler la loi du processus.

3.1 Méthode de régression en pratique et choix de la base polynomiale

Les algorithmes étant explicités dans la première partie, on s'intéresse dans cette partie à discuter de leur implémentation. En effet, on souhaite détailler les choix laissés libre par les auteurs et les perspectives de travail qui s'ensuivent. En particulier, on détaillera notre calcul des coefficients de la régression en pratique ainsi que les bases polynomiales étudiées pour celle-ci.

Dans une perspective d'optimiser le temps de calcul et de façon à pouvoir comparer nos résultats avec ceux obtenus par les auteurs, nous faisons le choix d'utiliser la librairie Python *scipy* et en particulier *linalg* pour effectuer la régression linéaire.

Plusieurs bases polynomiales sont proposées dans les articles. Nous avons fait le choix d'étudier l'influence de la base canonique, polynomiale de Laguerre, polynomiale de Jacobi et de la base polynomiale de Legendre. Comme souligné dans [3], il convient d'adapter le degré maximal de la décomposition en fonction de l'exemple étudié. On peut pour cela mentionner le résultat suivant tiré de [1], qui stipule que la qualité de l'approximation de l'espérance conditionnelle, utilisant une base de polynôme de taille J , est majorée par la somme d'un terme croissant en fonction de J et un terme décroissant en fonction de J . Ainsi, il y a un équilibre à trouver, ce que nous tentons de faire en visualisant la qualité de l'estimation de l'espérance conditionnelle par rapport à la taille de la base.

Prenons le cas d'un put. On présente ci-dessous les approximations de l'espérance conditionnelle en fonction de la taille de la base de polynôme utilisée. Ici, on utilise de simples puissances du spot.

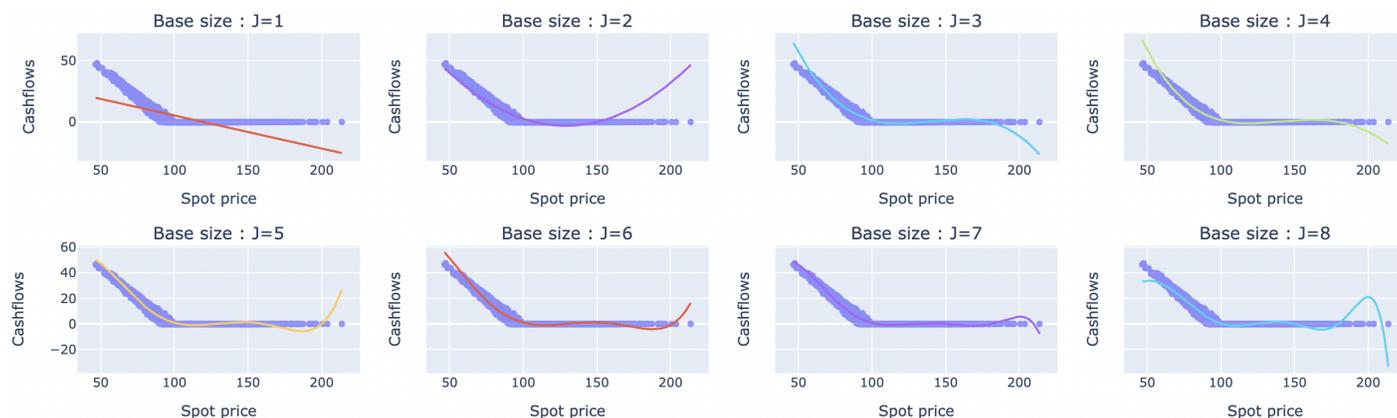


FIGURE 1 – Comparaison des flux réels et approximatés pour un put en fonction du degré maximal de la base de polynômes au temps t_{K-1} .

Les graphiques ci-dessous permettent de mieux visualiser l'impact du degré maximal sur l'estimateur de l'espérance conditionnelle. On remarque que à mesure que le degré augmente, l'estimation s'améliore, et ce jusqu'à un certain degré à partir duquel la qualité de l'estimation se détériore. L'observation faite précédemment, à savoir qu'augmenter la taille de la base pouvait mener à la détérioration de la qualité de l'estimation de l'espérance conditionnelle s'observe donc bien ici. Néanmoins, la taille de la base n'a que très peu d'impact sur le prix obtenu par les algorithmes précédents.

Une question qui subsiste est de se demander si choisir une famille de polynôme donnée peut améliorer la qualité de l'estimation. Pour y répondre, nous traçons sur la figure 2 les mêmes approximations que précédemment, mais pour une taille fixée $J=4$, et en utilisant les familles de polynôme suivantes : polynômes puissances, polynômes de Laguerre, polynômes de Jacobi et polynômes de Legendre. Les différences sont minimes.

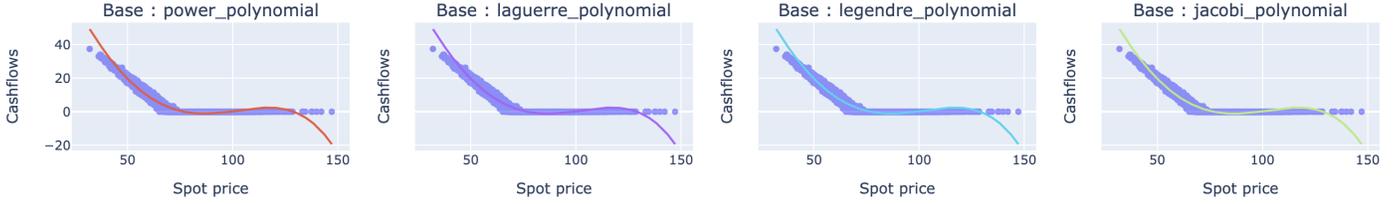


FIGURE 2 – Comparaison des flux réels et approximés pour un put en fonction de la base de polynômes utilisée au temps t_{K-1} .

3.2 Évaluation de produits dérivés américains

3.2.1 Put américain

S_0	σ	T	BS	LSM	(s.d. LSM)	DEEV (LSM)	LSM-ST	(s.d. LSM-ST)	DEEV (LSM-ST)
36	0.2	1	3.844	3.847	0.015	-0.003	4.421	0.014	-0.577
36	0.2	2	3.763	3.754	0.022	0.009	4.745	0.013	-0.982
36	0.4	1	6.711	6.670	0.025	0.041	6.947	0.030	-0.236
36	0.4	2	7.700	7.668	0.027	0.032	8.213	0.015	-0.513
38	0.2	1	2.852	2.839	0.011	0.013	3.170	0.009	-0.318
38	0.2	2	2.991	2.979	0.017	0.012	3.642	0.016	-0.651
38	0.4	1	5.834	5.807	0.030	0.027	5.987	0.033	-0.153
38	0.4	2	6.979	6.949	0.025	0.030	7.360	0.039	-0.381
40	0.2	1	2.066	2.052	0.011	0.014	2.239	0.009	-0.173
40	0.2	2	2.356	2.355	0.018	0.001	2.768	0.019	-0.412
40	0.4	1	5.060	5.022	0.024	0.038	5.128	0.025	-0.068
40	0.4	2	6.326	6.264	0.023	0.062	6.544	0.015	-0.228
42	0.2	1	1.465	1.477	0.016	-0.012	1.531	0.011	-0.066
42	0.2	2	1.841	1.831	0.013	0.010	2.081	0.013	-0.240
42	0.4	1	4.379	4.336	0.015	0.043	4.381	0.021	-0.002
42	0.4	2	5.736	5.670	0.032	0.066	5.870	0.035	-0.134
44	0.2	1	1.017	0.999	0.005	0.018	1.023	0.008	-0.006
44	0.2	2	1.429	1.422	0.018	0.007	1.554	0.016	-0.125
44	0.4	1	3.783	3.739	0.014	0.044	3.714	0.016	0.069
44	0.4	2	5.202	5.144	0.023	0.058	5.269	0.035	-0.067

TABLE 1 – Prix du Put américain obtenu avec les algorithmes LSM et $LSM - ST$ pour différentes valeurs de paramètres S_0 , σ et T . M , K et r sont respectivement fixés à 50, 40 et 0.06

Pour le problème de valorisation du put, nous avons choisi d'utiliser la base polynomiale de Laguerre pour effectuer notre régression et nous avons fixé $J = 3$. On réalise les estimations pour les mêmes paramètres de marché que dans [3], c'est-à-dire avec $K = 40$ et $r = 0.06$. L'option est exercable $M = 50$ fois. Par ailleurs, afin d'avoir une idée de la précision de chaque estimation, on répète le tirage et le calcul de valorisation 10 fois pour chaque paramètre donné. Les résultats obtenus sont visibles dans la table 1.

Pour expliciter les notations de la table 1, BS fait référence à la formule du Put européen pour le modèle de Black-Scholes, LSM et $LSM - ST$ correspondent aux deux algorithmes étudiés et implémentés, et $DEEV$ est la différence entre le prix européen BS et le prix simulé (LSM ou $LSM - ST$).

La première chose que l'on peut remarquer est que l'algorithme LSM donne des prix inférieurs dans la majorité des cas. De façon plus générale, le prix LSM est bien plus proche du prix BS , donc européen, que le prix $LSM - ST$. On pourrait alors discuter les résultats obtenus par LSM compte tenu du fait que le prix européen est toujours inférieur ou égal au prix américain de l'option associée. Pour autant, on sait aussi que l'utilisation d'option bermudéenne permet de converger vers le prix de l'option américaine par valeurs inférieures.

D'un point de vue uniquement comparatif, les résultats de l'article [3] obtenus par simulation sont bien plus proches des résultats de l'algorithme $LSM - ST$ que de l'algorithme LSM .

3.2.2 Call asiatique-américain

Toujours dans le but d'évaluer la performance de nos algorithmes, nous souhaitons dans cette section comparer les valeurs de prix obtenus par nos algorithmes et celles données dans [3] pour un call asiatique. Une option asiatique permet à son détenteur d'exercer à une date T , le payoff n'étant pas calculé à partir du niveau final du sous-jacent S_T , mais avec pour référence le cours moyen du sous-jacent entre la date 0 et la date T .

On considère ici une option asiatique américaine, que l'on approxime en utilisant une option bermudéenne qui peut être exercée à des temps t_k dans une grille \mathcal{R} . Dans ce cas, le payoff est calculé à partir de la moyenne de la valeur du sous-jacent entre la date $t_0 = 0$ et la date d'exercice t_k . On veut donc pouvoir considérer l'exercice de l'option à chaque date. Étant donné qu'en $t_0 = 0$ il n'y a pas de moyenne, l'article [3] propose de considérer la moyenne du processus S sur une période de 3 mois avant la date de valorisation. Cette valeur est notée A_0 .

Notons l le nombre de pas de temps entre t_0 et la date correspondant à 3 mois avant t_0 . Alors, dans le cas du call, le payoff à chaque instant d'exercice t_k est donné par

$$\left(\left(\frac{1}{k+l} \sum_{m=0}^k S_{t_m} + lA_0 \right) - K \right)^+$$

Le cas présenté dans l'article est plus complexe car l'option n'est exercable que 3 mois après la vente de l'option. Nous avons donc adapté les versions des algorithmes LSM et $LSM - ST$ de façon à tenir compte de ces problématiques.

De plus, de manière à reproduire le paramétrage de l'article, nous approximations à chaque temps t_k l'espérance conditionnelle en effectuant une régressions linéaire sur une constante, les deux premiers polynômes de Laguerre évalués au prix du spot à t_k , les deux premiers polynômes de Laguerre évalués à la moyenne des prix du spot entre t_0 et t_k , et le produit croisé des trois premiers polynômes de Laguerre évalués au prix du spot à t_k et ceux évalués à la moyenne des prix du spot entre t_0 et t_k . Ainsi, on utilise au total 8 polynômes pour la régression.

Les résultats obtenus, pour $T = 2$, $r = 0.06$ et $\sigma = 0.2$, sont visibles dans la table 2. De la même façon que pour le put américain classique, nous répétons le processus de simulation et de valorisation 10 fois ce qui nous permet d'estimer la précision de nos résultats. Nous avons aussi implémenté un algorithme Monte-Carlo nous donnant le prix européen⁴ de l'option associée.

On remarque que le prix donné par l'algorithme LSM est plus proche du prix européen que le prix donné par l'algorithme $LSM - ST$. De façon générale c'est l'algorithme LSM qui donne les prix simulés les plus proches de ceux observés dans l'article [3]. Les temps de calculs sont équivalents pour les deux algorithmes.

4. On entend par là prix de l'option exercable uniquement à maturité

A_0	S_0	EUR	(s.d.EUR)	LSM	(s.d.LSM)	DEEV (LSM)	LSM-ST	(s.d.LSM-ST)	DEEV (LSM-ST)
90	80	0.922	0.015	0.936	0.015	-0.014	0.777	0.018	0.145
90	90	3.168	0.047	3.217	0.047	-0.049	3.006	0.046	0.162
90	100	7.444	0.070	7.560	0.071	-0.116	7.872	0.080	-0.428
90	110	13.544	0.058	13.756	0.059	-0.212	14.868	0.053	-1.324
90	120	21.027	0.075	21.356	0.076	-0.323	23.203	0.089	-2.176
<hr/>									
100	80	1.050	0.024	1.066	0.025	-0.016	0.902	0.026	0.148
100	90	3.504	0.041	3.559	0.042	-0.055	3.454	0.044	0.050
100	100	8.002	0.068	8.127	0.069	-0.125	8.955	0.061	-0.953
100	110	14.413	0.055	14.638	0.056	-0.225	16.632	0.051	-2.219
100	120	21.999	0.152	22.343	0.155	-0.344	24.986	0.142	-2.987
<hr/>									
110	80	1.204	0.025	1.223	0.025	-0.019	1.113	0.037	0.091
110	90	3.899	0.048	3.960	0.049	-0.061	4.811	0.028	-0.912
110	100	8.672	0.053	8.808	0.054	-0.136	11.548	0.042	-2.876
110	110	15.208	0.062	15.447	0.064	-0.239	18.954	0.049	-3.746
110	120	22.881	0.079	23.241	0.080	-0.36	26.960	0.066	-4.079

TABLE 2 – Prix du put asiatique américain obtenu avec les algorithmes LSM et $LSM - ST$ pour différentes valeurs de paramètres S_0, A_0 .

3.3 Call américain et européen, convergences

Dans cette section, on souhaite observer les convergences prouvées dans la partie précédente. Ayant supposé le taux d'intérêt strictement positif, nous savons que le prix du call européen est égal à celui du call américain car il n'est pas optimal d'exercer ce dernier avant la maturité. De ce point de vue, le prix américain obtenu par méthode de Monte-Carlo et le prix européen doivent être relativement proches. On se propose donc de vérifier la convergence des algorithmes en se donnant pour référentiel la formule fermée du prix d'un call européen donnée par le modèle de Black-Scholes.

3.3.1 Convergence de l'estimateur du prix du call bermudéen vers le prix théorique du call bermudéen

Nous traçons ci-dessous l'évolution des prix d'un call calculés par les algorithmes LSM et $LSM-ST$ en fonction du nombre de trajectoires simulées pour le spot, N . On considérera $T = 1, S_0 = 100, K = 100, \sigma = 0.2$ et $r = 0.06$. M , le nombre de pas de temps utilisé est fixé à 100. L'espérance conditionnelle est approximée au moyen de puissances du spot jusqu'à l'ordre 5.

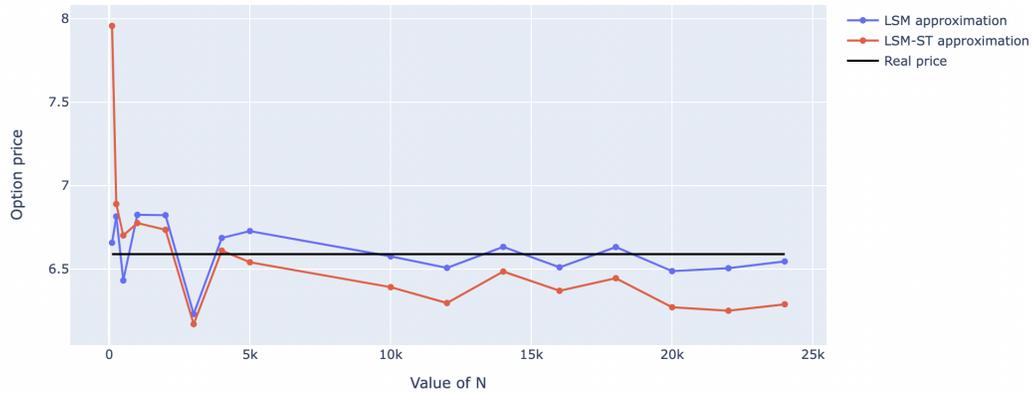


FIGURE 3 – Comparaison du prix théorique d'un call de maturité 1 an et des prix donnés par les algorithmes LSM et $LSM-ST$ pour $S_0 = 100, K = 100, \sigma = 0.2$ et $r = 0.06$ en fonction de N

A mesure que N croît, les prix simulés tendent vers le prix théorique. On observe encore que l'algorithme $LSM-ST$ tend vers le prix théorique par valeurs inférieures.

3.3.2 Convergence du prix du call bermudéen vers le prix du call américain

A présent, nous cherchons à montrer qu'à mesure que M , le nombre de pas de temps utilisés pour évaluer une option bermudéenne croît, le prix estimé tend vers le prix d'une option bermudéenne. Nous nous plaçons dans les mêmes configurations que précédemment. Cette fois, N est fixé à 10 000 et M varie.

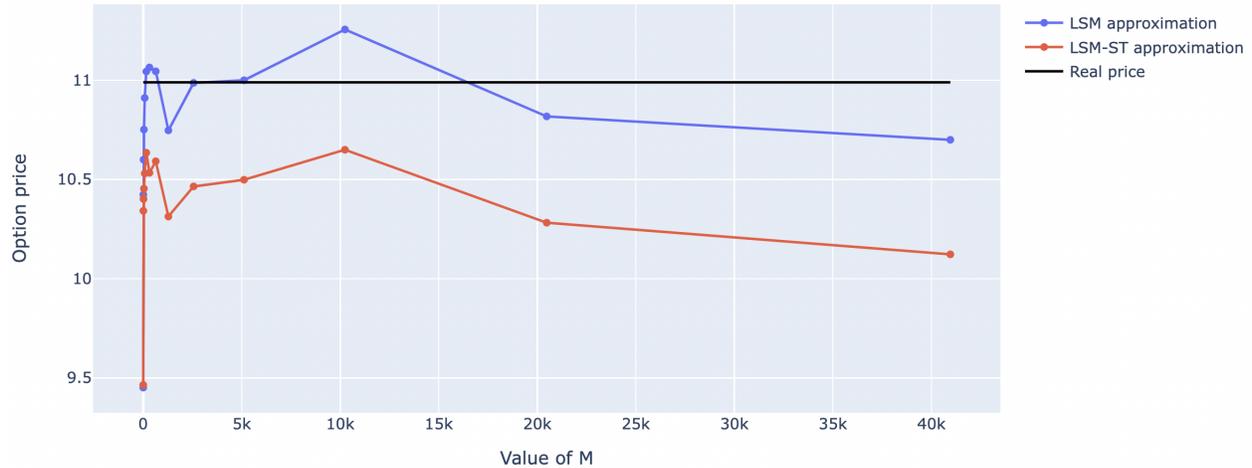


FIGURE 4 – Comparaison du prix théorique d'un call de maturité 1 an et des prix donnés par les algorithmes LSM et LSM-ST pour $S_0 = 100, K = 100, \sigma = 0.2, r = 0.06$ en fonction de M

On observe qu'à mesure que M augmente, le prix simulé tend vers le prix théorique. La convergence n'est néanmoins pas aussi claire que précédemment. Ajoutons à cela qu'augmenter considérablement M pour mieux visualiser la convergence aurait eu un impact majeur sur le temps de calcul, déjà élevé pour $M = 25000$. Cette observation permet de visualiser les limites, en pratique, de l'approche par la méthode de Longstaff-Schwartz.

Conclusion

Pour conclure, nous avons présenté dans ce rapport deux algorithmes permettant d'évaluer des options américaines, découlant des articles [2] et [3]. Après avoir détaillé la convergence et le caractère bien posé des algorithmes, nous avons comparé leur performance en évaluant des options classiques. Comme nous l'avons observé, leurs performances sont similaires, pour autant on constate une précision différente en fonction des cas d'application.

Le calcul de l'estimation du prix du put semble être meilleur, en se basant sur résultats obtenus dans [3], pour l'algorithme *LSM - ST*. Néanmoins l'estimation du prix du call asiatique est meilleure, toujours d'après les calculs de références de l'article [3], en utilisant l'algorithme *LSM*. Il est aussi important de noter que les algorithmes mis en place ont un coût de calcul équivalent, et peuvent facilement être adaptés en fonction de l'exemple étudié.

Cette étude nous a donc permis d'observer et comprendre l'intérêt numérique des méthodes décrites par les articles [2] et [3], dont la facilité d'implémentation constitue un atout majeur. Dans une perspective d'amélioration des résultats obtenus, on peut penser à améliorer la manière dont est effectuée la régression linéaire. Ayant constaté l'influence du degré maximal de la base polynomiale sur la qualité de la régression, cette observation laisser penser qu'une méthode de régression pondérée, donnant plus de poids aux points regroupés dans une zone à faible densité de flux, pourrait aider à augmenter la précision. Pour autant, nos travaux rendent bien compte du caractère limitant des méthodes étudiées dans ce rapport, pour la valorisation d'options américaines.

Annexes

Démonstration du Lemme 2.4

Notons $B_{t_k}^a = \{g(S_{t_k}) < C_{t_k}^{a,[J]}\}$ et $B_{t_k}^b = \{g(S_{t_k}) < C_{t_k}^{b,J}\}$. On a que

$$\begin{aligned}
D_{t_k}(a^{[J]}, S) - D_{t_k}(b^{[J]}, S) &= g(S_{t_k})(\mathbb{1}_{\overline{B}_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_k}^b}) + \left(D_{t_{k+1}}(a^{[J]}, S) + D_{t_{k+1}}(b^{[J]}, S)\right) (\mathbb{1}_{B_{t_k}^a} - \mathbb{1}_{B_{t_k}^b}) \\
&= g(S_{t_k})(\mathbb{1}_{\overline{B}_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_k}^b}) + \left(g(S_{t_{k+1}})(\mathbb{1}_{\overline{B}_{t_{k+1}}^a} - \mathbb{1}_{\overline{B}_{t_{k+1}}^b}) + D_{t_{k+2}}(a^{[J]}, S)\mathbb{1}_{B_{t_{k+1}}^a} + D_{t_{k+2}}(b^{[J]}, S)\mathbb{1}_{B_{t_{k+1}}^b}\right) (\mathbb{1}_{B_{t_k}^a} - \mathbb{1}_{B_{t_k}^b}) \\
&= \dots \\
&= g(S_{t_k})(\mathbb{1}_{\overline{B}_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_k}^b}) + \sum_{t_j=t_k+1}^{t_{K-1}} g(S_{t_j}) \left(\mathbb{1}_{\overline{B}_{t_j}^a, \overline{B}_{t_{j-1}}^a, \dots, B_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_j}^b, \overline{B}_{t_{j-1}}^b, \dots, B_{t_k}^b}\right) + g(S_{t_K}) \left(\mathbb{1}_{B_{t_k}^a, B_{t_{K-1}}^a, \dots, B_{t_k}^a} - \mathbb{1}_{B_{t_k}^b, B_{t_{K-1}}^b, \dots, B_{t_k}^b}\right)
\end{aligned}$$

Or on a que

$$\begin{aligned}
|g(S_{t_k})| |\mathbb{1}_{\overline{B}_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_k}^b}| &= \begin{cases} |g(S_{t_k})| & \text{si } C_{t_k}^{a,[J]} \leq g(S_{t_k}) < C_{t_k}^{b,J} \text{ ou si } C_{t_k}^{b,J} \leq g(S_{t_k}) < C_{t_k}^{a,[J]} \\ 0 & \text{sinon} \end{cases} \\
&\leq |g(S_{t_k})| \mathbb{1}_{\{C_{t_k}^{a,[J]} \leq g(S_{t_k}) < C_{t_k}^{b,J}\}} + \mathbb{1}_{\{C_{t_k}^{b,J} \leq g(S_{t_k}) < C_{t_k}^{a,[J]}\}} \\
&\leq |g(S_{t_k})| \mathbb{1}_{\{g(S_{t_k}) \leq |C_{t_k}^{b,J}| \leq |a^{[J],(k)} - b^{[J],(k)}| |L^{[J]}(S_{t_k})|\}}
\end{aligned}$$

De plus,

$$\begin{aligned}
\sum_{t_j=t_k+1}^{t_{K-1}} |g(S_{t_j})| \left| \mathbb{1}_{\overline{B}_{t_j}^a, \overline{B}_{t_{j-1}}^a, \dots, B_{t_k}^a} - \mathbb{1}_{\overline{B}_{t_j}^b, \overline{B}_{t_{j-1}}^b, \dots, B_{t_k}^b} \right| &\leq \sum_{t_j=t_k+1}^{t_{K-1}} |g(S_{t_j})| \sum_{t_l=t_j}^{t_{k-1}} |\mathbb{1}_{\overline{B}_{t_l}^a} - \mathbb{1}_{\overline{B}_{t_l}^b}| + |\mathbb{1}_{B_{t_l}^a} - \mathbb{1}_{B_{t_l}^b}| \\
&\leq \sum_{t_j=t_k+1}^{t_{K-1}} |g(S_{t_j})| \sum_{t_l=t_j}^{t_k} |\mathbb{1}_{\overline{B}_{t_l}^a} - \mathbb{1}_{\overline{B}_{t_l}^b}| \\
&\leq \sum_{t_j=t_k+1}^{t_{K-1}} \sum_{t_l=t_j}^{t_k} |g(S_{t_j})| \mathbb{1}_{\{g(S_{t_l}) \leq |C_{t_l}^{b,J}| \leq |a^{[J],(l)} - b^{[J],(l)}| |L^{[J]}(S_{t_l})|\}}
\end{aligned}$$

par précédemment. Donc

$$\begin{aligned}
\left| D_{t_k}(a^{[J]}, S) - D_{t_k}(b^{[J]}, S) \right| &\leq \sum_{t_j=t_k}^{t_{K-1}} \sum_{t_l=t_j}^{t_k} |g(S_{t_j})| \mathbb{1}_{\{g(S_{t_l}) \leq |C_{t_l}^{b,J}| \leq |a^{[J],(l)} - b^{[J],(l)}| |L^{[J]}(S_{t_l})|\}} \\
&\quad + |g(S_{t_K})| \left| \mathbb{1}_{B_{t_k}^a, B_{t_{K-1}}^a, \dots, B_{t_k}^a} - \mathbb{1}_{B_{t_k}^b, B_{t_{K-1}}^b, \dots, B_{t_k}^b} \right| \\
&\leq \left(\sum_{t_j=t_k}^{t_K} |g(S_{t_j})| \right) \left(\sum_{t_j=t_k}^{t_{K-1}} \mathbb{1}_{\{|g(S_{t_k}) - C_{t_k}^{b,J}| \geq |a^{[J],(k)} - b^{[J],(k)}| |L^{[J]}(S_{t_k})|\}} \right)
\end{aligned}$$

3.4 Démonstration de la proposition 2.5

On donne la démonstration dans le cas où g est Lipschitz de constante C_g et le processus de spot a une dynamique de type

$$dS_t = b(S_t)dt + \sigma(S_t)dW_t$$

pour b, σ bornées et $(W_t)_{t \geq 0}$ un mouvement brownien.

L'option américaine donnant plus de droits que l'option bermudéenne correspondante, on sait que $0 \leq V_0 - V_0^b$. De plus,

$$V_0 - V_0^b = \mathbb{E} \left[e^{-\int_0^{\tau_0^*} r_s ds} g(S_{\tau_0^*}) - e^{-\int_0^{\tau_0^{*b}} r_s ds} g(S_{\tau_0^{*b}}) \right]$$

avec τ_0^* et τ_0^{*b} respectivement les temps d'arrêt optimaux de l'option américaine et de l'option bermudéenne.

Le problème qui se pose est qu'il n'est pas possible de comparer les termes à l'intérieur de l'espérance car il est possible que τ_0^* n'appartienne pas à la grille \mathcal{R} sur laquelle est définie l'option bermudéenne. On définit donc sa projection sur \mathcal{R} par

$$\tau_0^{*,\mathcal{R}} = t_{k+1} \mathbb{1}_{\{t_k < \tau_0^* \leq t_{k+1}\}}$$

pour un certain $k \in \llbracket 0, K \rrbracket$. On a alors que

$$V_0 - V_0^b \leq \mathbb{E} \left[e^{-\int_0^{\tau_0^*} r_s ds} g(S_{\tau_0^*}) - e^{-\int_0^{\tau_0^{*,\mathcal{R}}} r_s ds} g(S_{\tau_0^{*,\mathcal{R}}}) \right]$$

car $\tau_0^{*,\mathcal{R}}$ est (possiblement) sous optimal pour l'exercice de l'option bermudéenne donc

$$\mathbb{E}[e^{-\int_0^{\tau_0^{*,\mathcal{R}}} r_s ds} g(S_{\tau_0^{*,\mathcal{R}}})] \geq \mathbb{E}[e^{-\int_0^{\tau_0^{*b}} r_s ds} g(S_{\tau_0^{*b}})]. \text{ Ainsi :}$$

$$\begin{aligned} |V_0 - V_0^b| &\leq \left| \mathbb{E} \left[e^{-\int_0^{\tau_0^*} r_s ds} g(S_{\tau_0^*}) - e^{-\int_0^{\tau_0^{*,\mathcal{R}}} r_s ds} g(S_{\tau_0^{*,\mathcal{R}}}) \right] \right| \\ &\leq \mathbb{E} \left[e^{-\int_0^{\tau_0^{*,\mathcal{R}}} r_s ds} |g(S_{\tau_0^*}) - g(S_{\tau_0^{*,\mathcal{R}}})| \right] \text{ car } \tau_0^* \leq \tau_0^{*,\mathcal{R}} \\ &\leq \mathbb{E} \left[|g(S_{\tau_0^*}) - g(S_{\tau_0^{*,\mathcal{R}}})| \right] \text{ car } \tau_0^* \leq \tau_0^{*,\mathcal{R}} \\ &\leq C_g \mathbb{E} \left[|S_{\tau_0^*} - S_{\tau_0^{*,\mathcal{R}}}| \right] \text{ car } g \text{ est Lipschitz} \\ &\leq C_g \left(\mathbb{E} \left[|S_{\tau_0^*} - S_{\tau_0^{*,\mathcal{R}}}|^2 \right] \right)^{\frac{1}{2}} \text{ par l'inégalité de Cauchy-Schwartz} \end{aligned}$$

Or par le lemme d'Itô,

$$\begin{aligned} S_{\tau_0^{*,\mathcal{R}}} &= S_{\tau_0^*} + \int_{\tau_0^*}^{\tau_0^{*,\mathcal{R}}} b(S_t) dt + \int_{\tau_0^*}^{\tau_0^{*,\mathcal{R}}} \sigma(S_t) dW_t \\ &= S_{\tau_0^*} + \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} b(S_t) dt + \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} \sigma(S_t) dW_t \end{aligned}$$

Donc finalement en injectant cette expression, on obtient :

$$\begin{aligned} \left(\mathbb{E} \left[|S_{\tau_0^*} - S_{\tau_0^{*,\mathcal{R}}}|^2 \right] \right)^{\frac{1}{2}} &= \left(\mathbb{E} \left[|S_{\tau_0^{*,\mathcal{R}}} - S_{\tau_0^*}|^2 \right] \right)^{\frac{1}{2}} \\ &= \left(\mathbb{E} \left[\left| \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} b(S_t) dt + \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} \sigma(S_t) dW_t \right|^2 \right] \right)^{\frac{1}{2}} \\ &= \left(\mathbb{E} \left[\left| \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} b(S_t) dt \right|^2 \right] + \mathbb{E} \left[\left| \int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} \sigma(S_t) dW_t \right|^2 \right] \right)^{\frac{1}{2}} \text{ (terme croisé nul)} \\ &\leq \left(\mathbb{E} \left[\int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} |b(S_t)|^2 dt \right] + \mathbb{E} \left[\int_0^T \mathbb{1}_{\{\tau_0^* \leq t \leq \tau_0^{*,\mathcal{R}}\}} |\sigma(S_t)|^2 dt \right] \right)^{\frac{1}{2}} \text{ (isométrie d'Itô)} \\ &\leq C \left(\mathbb{E} \left[\tau_0^{*,\mathcal{R}} - \tau_0^* \right] \right)^{\frac{1}{2}} \text{ (} b, \sigma \text{ bornés)} \\ &\leq C |\mathcal{R}|^{\frac{1}{2}} \end{aligned}$$

avec C une constante liée au fait que b et σ soient bornées. On obtient alors

$$|V_0 - V_0^b| \leq D|\mathcal{R}|^{\frac{1}{2}}$$

où D est une constante dépendante de la constante de lipschitz C_g . Par passage au sup à gauche, on obtient bien le résultat souhaité.

Références

- [1] J.-F. CHASSAGNEUX, *Notes de cours, M2MO, Méthodes probabilistes avancées*, 2023.
- [2] E. CLEMENT, D. LAMBERTON, AND P. PROTTER, *An analysis of a least squares regression method for American option pricing*, Finance and Stochastics, 2002.
- [3] F. A. LONGSTAFF AND E. S. SCHWARTZ, *Valuing American Options by Simulation : A Simple Least-Squares Approach.*, The Review of Financial Studies, 2001.