

Estimation des sensibilités de la CVA par Adjoint Différentiation

Adrien Misko, Ismail Bninou

Nexialog Consulting, Paris, France



Résumé Cet article aborde deux grandes familles de méthodes qui permettent l'estimation des sensibilités, à savoir, les méthodes par différence finie et les méthodes Pathwise. On s'intéresse particulièrement à l'application de l'Adjoint Différentiation, une des méthodes Pathwise, dans le domaine de la finance avec la finalité de calculer les sensibilités des prix des instruments financiers par rapport à différents facteurs. Dans cet article, on utilise la méthode d'Adjoint Différentiation dans le but de calculer les sensibilités de la CVA pour certains produits de taux exotiques. On montre que cette méthode est aussi précise que la méthode par différence finie centrée tout en demandant un temps de calcul moindre.

Keywords: Différence finie · Adjoint Différentiation · Monte Carlo · produits financiers.

Table des matières

Estimation des sensibilités de la CVA par Adjoint Différentiation	1
<i>Adrien Misko, Ismail Bninou</i>	
1 Approximation par différence finie	4
1.1 Biais de l'estimation par différence finie	4
1.2 Variance de l'estimation par différence finie	5
1.3 Équilibre biais variance	6
Dérivées de second ordre	7
2 Les méthodes Pathwise	8
2.1 Introduction à l'Adjoint Différentiation	9
2.2 L'Adjoint Différentiation pour les instruments financiers à terme	9
2.3 Optimisation de l'Adjoint Différentiation	11
2.4 Adjoint Différentiation pour les instruments exotiques	12
3 Résultats numériques sur l'AD	14
4 Credit Value Adjustment CVA	15
4.1 Définition de la CVA	15
4.2 Vision générale sur le calcul de la CVA	16
4.3 Méthode de calcul des sensibilités de la CVA par AD	17
Méthode pour un Swap	17
Méthode pour une Bermuda Swaption	19
5 Résultats numériques sur la CVA	21
5.1 Calcul de la CVA	21
5.2 Calcul des sensibilités de la CVA pour une Bermuda Swaption	23
A Least Square Monte Carlo	27
B La méthode MC^2	28
C Modèle de diffusion des taux courts : Vasicek	28
C.1 Prix de l'obligation zéro coupon	29
C.2 Discount factor	30
C.3 Taux Libor	30
D Procédure d'AD pour un Swap	30

Table des figures

1	Cash Flow d'un Swap.	11
2	Courbes de delta et vega de la Bermuda Swaption en fonction de r_0 par AD et par différence finie	14
3	Intervalle de confiance sur l'estimation du delta et du vega de la Bermuda Swaption par AD.	15
4	Comparaison de l'exposition $V_c(t_{\phi(i)})^+$ par MC ² et LSMC.	22
5	Valeurs des probabilités de défaut historiques.	22
6	Courbes de delta et vega de la CVA de la Bermuda Swaption en fonction de r_0 par AD et par différence finie	23
7	Intervalle de confiance sur l'estimation du delta du vega de la CVA de la Bermuda Swaption par AD.	24
8	Illustration du principe de la méthode MC ²	28



Introduction

Le calcul des grecs est d'une très grande importance dans le domaine de la finance, par exemple dans le cadre de la FRTB ou de la couverture $(\delta, \text{vega}, \gamma)$.

Dans cet article, on présente la méthodologie de calcul des grecs par AD (Adjoint Différentiation), une méthode appartenant à la famille Pathwise, puisqu'elle présente certains avantages par rapport aux méthodes de la famille des méthodes par différence finie.

Dans un deuxième temps, on décrit l'utilisation de l'AD dans le calcul optimal des grecs du Swap et de la Bermuda Swaption et on fournit des résultats numériques permettant à la fois de confirmer la méthode et de montrer son optimalité.

Dans un troisième temps, on définit la CVA (Credit Value Adjustment) et on présente une adaptation de l'AD permettant le calcul des grecs de celle-ci. On finit par exposer des résultats numériques relatifs à la CVA des deux produits de taux précédents.

1 Approximation par différence finie

Soit une variable aléatoire $P(\theta)$ qui dépend du paramètre θ . On suppose que l'on dispose d'un mécanisme qui nous permet de simuler $P(\theta)$ et on note $P(\theta)^{[p]}$ la p -ième simulation de $P(\theta)$.

On définit

$$\alpha(\theta) = \mathbb{E}[P(\theta)].$$

On note $\alpha'(\theta)$ la dérivée de la fonction α par rapport au paramètre θ . L'objectif est d'estimer $\alpha'(\theta)$. Dans la pratique, $\alpha(\theta)$ pourrait être le prix d'un produit dérivé et θ la volatilité du modèle associé.

Ci-dessous, on explique comment estimer $\alpha'(\theta)$ par deux méthodes par différence finie : la différence finie à droite et la différence finie centrée. Les estimations par différence finie évoquent un dilemme biais variance qui est aussi abordé.

1.1 Biais de l'estimation par différence finie

Soit $h > 0$. Soit N_p un entier naturel. On réalise les N_p simulations $P^{[p]}(\theta)$ et les N_p simulations $P^{[p]}(\theta + h)$

L'estimation par différence finie à droite de $\alpha'(\theta)$ est

$$\hat{\Delta}_F(N_p, h) = \frac{1}{N_p} \sum_{p=1}^{N_p} \frac{P^{[p]}(\theta + h) - P^{[p]}(\theta)}{h}.$$

Si α est deux fois dérivable, la formule de Taylor permet d'écrire

$$\alpha(\theta + h) = \alpha(\theta) + \alpha'(\theta)h + \frac{1}{2}\alpha''(\theta)h^2 + o(h^2).$$

On en déduit le biais de l'estimation $\hat{\Delta}_F$

$$Biais(\hat{\Delta}_F)(N_p, h) = \mathbb{E}[\hat{\Delta}_F(N_p, h) - \alpha'(\theta)] = \frac{1}{2}\alpha''(\theta)h + o(h).$$

Par ailleurs, on définit l'estimation par différence finie centrée par

$$\hat{\Delta}_C(N_p, h) = \frac{1}{N_p} \sum_{p=1}^{N_p} \frac{P^{[p]}(\theta + h) - P^{[p]}(\theta - h)}{h}.$$

Si α est trois fois dérivable, on peut montrer que $Biais(\hat{\Delta}_C(N_p, h)) = \frac{1}{6}\alpha'''(\theta)h^2 + o(h^2)$.

L'estimation centrée est plus précise que l'estimation à droite. Le biais de l'estimation par différence finie dépend à la fois de h et de N_p .

Afin de s'assurer de la qualité de l'estimation, il faut calculer la variance de l'estimateur.

1.2 Variance de l'estimation par différence finie

Considérons l'estimateur par différence finie à droite dont la variance est

$$\mathbb{V}ar[\hat{\Delta}_F(N_p, h)] = h^{-2} \mathbb{V}ar\left[\frac{1}{N_p} \sum_{p=1}^{N_p} (P^{[p]}(\theta + h) - P^{[p]}(\theta))\right]. \quad (1)$$

Par la suite, on se met dans le cas simple où les $(P^{[p]}(\theta), P^{[p]}(\theta + h))$ sont iid. (se simplifie en

$$\mathbb{V}ar[\hat{\Delta}_F(N_p, h)] = \frac{1}{h^{-2}N_p} \mathbb{V}ar[(P(\theta + h) - P(\theta))]. \quad (2)$$

La valeur de la variance est donc déterminée à partir de l'ordre de grandeur de $\mathbb{V}ar[P(\theta + h) - P(\theta)]$ qui est une fonction de h .

Dans la pratique, on se retrouve généralement dans l'un des cas suivants

$$\mathbb{V}ar[P(\theta + h) - P(\theta)] = \begin{cases} O(1), & \text{cas } i \\ O(h), & \text{cas } ii \\ O(h^2), & \text{cas } iii. \end{cases} \quad (3)$$

De manière générale, on se retrouve dans le cas i si on simule $P(\theta)$ et $P(\theta + h)$ de manière indépendante. Le cas échéant et sous la condition peu exigeante de continuité de $\mathbb{V}ar[P(\theta)]$ en θ , on a

$$\mathbb{V}ar[P(\theta + h) - P(\theta)] = \mathbb{V}ar[P(\theta + h)] + \mathbb{V}ar[P(\theta)] \xrightarrow{h \rightarrow 0} 2\mathbb{V}ar[P(\theta)].$$

Par conséquent, la variance de l'estimateur par différence finie à droite diverge quand h tend vers 0.

Le cas *ii* se présente typiquement lorsqu'on simule $P(\theta)$ et $P(\theta + h)$ de manière dépendante, c'est à dire lorsque les réalisations aléatoires utilisées pour simuler $P^{[p]}(\theta)$ sont sauvegardées et réutilisées pour simuler $P^{[p]}(\theta + h)$

Pour retrouver le cas *iii*, on a généralement besoin, outre l'utilisation des mêmes réalisations pour les simulations de $Y(\theta)$ et de $Y(\theta + h)$, de la continuité presque sûre de $Y(\cdot)$. Le prix de l'option digitale appartient au cas *ii* sans appartenir au cas *iii*, tandis que le prix de l'option européenne est un exemple du cas *iii*.

Excepté le cas *iii*, la variance de l'estimateur (2) diverge pour les petites valeurs de h . Au contraire, diminuer la valeur de h permet de diminuer le biais. C'est le dilemme biais variance.

1.3 Équilibre biais variance

On se place dans un cadre général et on note $\hat{\Delta} = \hat{\Delta}(n, h)$ un estimateur générique par différence finie de $\alpha'(\theta)$. Notre objectif est de minimiser l'erreur quadratique moyenne suivante :

$$MSE(\hat{\Delta}(n, h)) = \mathbb{E}[\hat{\Delta} - \alpha'(\theta)]^2 + \mathbb{V}ar[\hat{\Delta}]. \quad (4)$$

Soient

$$\mathbb{E}[\hat{\Delta} - \alpha'(\theta)] = bh^\beta + o(h^\beta) \quad \text{et} \quad \mathbb{V}ar[\hat{\Delta}] = \frac{\sigma^2}{nh^\eta} + o(h^{-\eta}), \quad (5)$$

respectivement le biais et la variance de $\hat{\Delta}$ avec β, η, σ des réels strictement positifs et b un réel non nul.

Si $\hat{\Delta}$ est l'estimateur par différence finie à droite, $\beta = 1$ et si $\hat{\Delta}$ est l'estimateur par différence finie centrée, $\beta = 2$.

On remarque aussi que $\eta = 2$ correspond au cas *i* de (3) et que $\eta = 1$ correspond au cas *ii*.

On considère la séquence d'estimateurs $\hat{\Delta}(n, h_n)$ telles que

$$h_n = h_* n^{-\gamma}, \quad (6)$$

où h_* et γ sont positifs.

En injectant les deux expressions (6) dans (5), on obtient

$$MSE(\hat{\Delta}) = b^2 h_n^{2\beta} + \frac{\sigma^2}{n h_n^2}.$$

Ainsi, Asymptotiquement, la valeur de γ maximisant la vitesse de décroissance de la MSE est $\gamma = \frac{1}{2\beta+\eta}$. On voit que l'on. Pour ce γ optimal, on a

$$MSE(\hat{\Delta}) = O(n^{-\frac{2\beta}{2\beta+\eta}}). \quad (7)$$

Plus précisément, on a

$$n^{\frac{2\beta}{2\beta+\eta}} MSE(\hat{\Delta}) \xrightarrow{n \rightarrow +\infty} b^2 h_*^{2\beta} + \sigma^2 h_*^{-\eta}.$$

On aboutit donc à la valeur optimale de h :

$$h_* = \left(\frac{\eta \sigma^2}{2\beta b^2} \right)^{\frac{1}{2\beta+\eta}} \quad (8)$$

La table ci-dessous de [1] récapitule les résultats dans le cas d'une différence finie à droite et d'une différence finie centrée.

Estimateur	variance	biais	h_n optimal	ordre de convergence	h_*
de l' MSE					
$\Delta_{F,i}$	$\frac{\sigma_{F,i}^2}{h_n^2}$	$\frac{1}{2} \alpha''(\theta) h$	$O(n^{-\frac{1}{4}})$	$O(n^{-\frac{1}{2}})$	$\frac{4\sigma_{F,i}^2}{\alpha''(\theta)^2}$
$\Delta_{F,ii}$	$\frac{\sigma_{F,ii}^2}{h_n^2}$	$\frac{1}{6} \alpha'''(\theta) h$	$O(n^{-\frac{1}{6}})$	$O(n^{-\frac{2}{3}})$	$\frac{18\sigma_{F,ii}^2}{\alpha'''(\theta)^2}$
$\Delta_{C,i}$	$\frac{\sigma_{C,i}^2}{h_n}$	$\frac{1}{2} \alpha''(\theta) h$	$O(n^{-\frac{1}{3}})$	$O(n^{-\frac{2}{3}})$	$\frac{2\sigma_{C,i}^2}{\alpha''(\theta)^2}$
$\Delta_{C,ii}$	$\frac{\sigma_{C,ii}^2}{h_n}$	$\frac{1}{6} \alpha'''(\theta) h$	$O(n^{-\frac{1}{5}})$	$O(n^{-\frac{4}{5}})$	$\frac{9\sigma_{C,ii}^2}{\alpha'''(\theta)^2}$

Vitesse de convergence pour les estimateurs à incrément optimal h_n . L'indice i (resp. ii) dans l'estimateur Δ est relatif au cas i (resp. ii).

Dérivées de second ordre Soit l'estimateur de la dérivée de second ordre suivant

$$\frac{\bar{Y}_p(\theta + h) - 2\bar{Y}_p(\theta) + \bar{Y}_p(\theta - h)}{h^2}.$$

Même en utilisant des variables aléatoires communes pour les trois simulations $\bar{Y}^{[p]}(\theta - h)$, $\bar{Y}^{[p]}(\theta)$ et $\bar{Y}^{[p]}(\theta + h)$ pour une fonction $\alpha(\theta)$ dérivable d'ordre quatre, on peut au mieux atteindre une vitesse de convergence de 4 d'ordre $O(h^{-4/7})$. Ainsi, Si on les compare avec les dérivées premières, les dérivées secondes sont plus difficiles à approcher.

Conclusion

Les méthodes par différence finie fournissent une bonne estimation de la sensibilité certes, mais à défaut d'un bon choix de l'incrément, on obtient des résultats instables et imprécis. D'un autre côté, pour une fonction à plusieurs variables, le nombre d'évaluations de la fonction nécessaires afin de déterminer toutes ses sensibilités avec une méthode par différence finie augmente linéairement en fonction du nombre de variables. Ainsi, le coût de la méthode devient onéreux pour les fonctions à grand coût d'évaluation. La famille des méthodes présentée ci-dessous ne présente pas cet inconvénient.

2 Les méthodes Pathwise

Au contraire de la méthode par différence finie, les méthodes Pathwise ne nécessitent qu'un seul jeu de simulation, celui qui est relatif au point θ .

De manière plus formelle, soit $(t_i)_{0 \leq i \leq N_t}$ la suite strictement croissante d'instantes et considérons un instrument financier général dont le prix P dépend du paramètre θ et des taux courts $(r_{t_i})_{0 \leq i \leq N_t}$ et soit $V(\theta) = \mathbb{E}[P((r_{t_i})_{0 \leq i \leq N_t}, \theta)]$ sa valeur. Étant donné les conditions de régularité suivantes [1] :

- (C1) Pour tout θ et pour tout $0 \leq i \leq N_t$, $r'_{t_i}(\theta)$ existe avec une probabilité 1
Soit \mathcal{D}_P l'ensemble de points de différentiabilité de P
- (C2) Pour tout θ , $\mathbb{P}[(r_{t_i})_{0 \leq i \leq N_t}, \theta] \in \mathcal{D}_P = 1$
- (C3) Il existe une constante k_P telle que pour tout $x, y \in \mathbb{R}^{N_t+1} \times \mathbb{R}$,
 $|P(x) - P(y)| \leq k_P \|x - y\|$
- (C4) Pour tous θ_1, θ_2 et pour tout $1 \leq i \leq N_t$, il existe une variable aléatoire κ_i telle que $|r_{t_i}(\theta_2) - r_{t_i}(\theta_1)| \leq \kappa_i |\theta_2 - \theta_1|$ et $\mathbb{E}[\kappa_i] < \infty$,

l'intervention des signes de dérivation et d'intégration est possible et on peut écrire

$$\frac{\partial V(\theta)}{\partial \theta} = \mathbb{E} \left[\frac{\partial P((r_{t_i})_{0 \leq i \leq N_t}, \theta)}{\partial \theta} \right].$$

La dérivation de la variable aléatoire à l'intérieure de l'espérance dans (2) se fait pour chaque trajectoire de la diffusion de $(r_{t_i})_{0 \leq i \leq N_t}$. En ce qui suit, on explique comment calculer cette dérivée en utilisant l'Adjoint Différentiation, une des méthodes Pathwise.

2.1 Introduction à l'Adjoint Différentiation

Le travail ci-dessous se base sur les travaux d'Antonov [2]. Dans son article, [2] explique comment calculer les sensibilités par AD pour des produits dérivés *path-dependant* : Il présente d'abord l'AD pour les instruments financiers à terme. Ensuite il explique comment appliquer l'AD pour les produits exotiques à travers l'exemple des Bermuda Swaption et il finit par décrire la méthode *Backward differentiation* qui permet d'alléger les coûts de stockage en appliquant l'AD classique sur les valeurs de continuation du produit dérivé.

2.2 L'Adjoint Différentiation pour les instruments financiers à terme

L'adjoint différentiation permet de calculer une valeur de la dérivée pour chaque trajectoire simulée, et le passage aux dérivées de $V(\theta)$ se fait par Monte Carlo (MC) i.e :

$$\frac{\partial V(\theta)}{\partial \theta} = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{\partial P((r_{t_i}^{[p]})_{0 \leq i \leq N_t}, \theta)}{\partial \theta},$$

où $r_{t_i}^{[p]}$ est la p -ème simulation du taux court à l'instant t_i .

Ci-dessous, la procédure rétrograde permettant de calculer $\frac{\partial P((r_{t_i}^{[p]})_{0 \leq i \leq N_t}, \theta)}{\partial \theta}$ est détaillée. Pour alléger la notation, on omet l'indice p .

Supposons d'abord que le processus $(r_t)_t$ est Markovien et qu'il permet une relation de récurrence de la forme

$$r_{t_{i+1}} = G(r_{t_i}, Z_i, \theta) \quad 0 \leq i < N_t.$$

où les Z_i sont des variables aléatoires iid de loi normale centrée réduite et G une fonction réelle.

À noter que ces conditions ne sont pas contraignantes et qu'elles sont atteignables par une discrétisation d'Euler par une grande partie des modèles de taux telles que le modèle de Black et Scholes et le modèle de Vasicek utilisé pour les applications numériques.

On a alors

$$P((r_{t_i})_{0 \leq i \leq N_t}, \theta) = P((r_{t_0}, \dots, r_{N_t-1}, G(r_{N_t-1}, Z_{N_t-1}, \theta)), \theta).$$

r_{N_t} n'apparaît donc plus dans la formule, en revanche, la variable aléatoire Z_{N_t-1} est apparue. Puisque Z_{N_t-1} ne dépend d'aucune autre quantité (notamment θ), on peut la considérer comme paramètre de la fonction P_{N_t-1} définie par

$$P_{N_t-1}((r_{t_i})_{0 \leq i \leq N_t-1}, \theta) = P((r_{t_i})_{0 \leq i \leq N_t}, \theta).$$

Ceci permet d'écrire

$$\frac{\partial P_{N_t-1}}{\partial \theta} = \frac{\partial P_{N_t}}{\partial \theta} + \frac{\partial P_{N_t}}{\partial r_{N_t}} \frac{\partial G(r_{N_t-1}, Z_{N_t-1}, \theta)}{\partial \theta}.$$

Ensuite, on calcule la quantité utile

$$\frac{\partial P_{N_t-1}}{\partial r_{N_t-1}} = \frac{\partial P}{\partial r_{N_t}} + \frac{\partial P_{N_t}}{\partial r_{N_t}} \frac{\partial G(r_{N_t-1}, Z_{N_t-1}, \theta)}{\partial r_{N_t-1}}.$$

De manière générale, on passe de $\frac{\partial P_i}{\partial \theta}$ à $\frac{\partial P_{i-1}}{\partial \theta}$ en utilisant d'abord

$$\frac{\partial P_{i-1}}{\partial \theta} = \frac{\partial P_i}{\partial \theta} + \frac{\partial P_i}{\partial r_{t_i}} \frac{\partial G(r_{t_{i-1}}, Z_{i-1}, \theta)}{\partial \theta}, \quad (9)$$

la quantité $\frac{\partial P_i}{\partial r_{t_i}}$ étant calculée à l'étape précédente. Ensuite, pour l'étape suivante, on calcule

$$\frac{\partial P_{i-1}}{\partial r_{t_{i-1}}} = \frac{\partial P}{\partial r_{t_{i-1}}} + \frac{\partial P_i}{\partial r_{t_i}} \frac{\partial G(r_{t_{i-1}}, Z_{i-1}, \theta)}{\partial r_{t_{i-1}}}. \quad (10)$$

où P_i est définie telle que

$$P_i((r_{t_i})_{0 \leq i \leq i-1}, \theta) = P((r_{t_i})_{0 \leq i \leq N_t}, \theta).$$

À l'étape finale on déduit les valeurs

$$\frac{\partial P_0}{\partial \theta}, \quad \frac{\partial P_0}{\partial r_{t_0}}.$$

Puisque P_0 ne fait plus apparaître que r_{t_0} et θ , $\frac{\partial P_0}{\partial \theta}$ est la dérivée totale de P par rapport à θ et $\frac{\partial P_0}{\partial r_{t_0}}$ est la dérivée totale de P par rapport à r_0 .

La procédure AD ci-dessus fournit deux dérivées en utilisant les données d'une seule simulation de prix.

Afin de réaliser l'AD, il faut au préalable calculer analytiquement les dérivées :

$$\frac{\frac{\partial G(r_{t_i}, Z_i, \theta)}{\partial r_{t_i}}}{\frac{\partial G(r_{t_i}, Z_i, \theta)}{\partial \theta}} \frac{\partial P}{\partial r_{t_i}} \quad (11)$$

La complexité numérique de l'AD est $O(N_t^2 \times N_p)$.

2.3 Optimisation de l'Adjoint Différentiation

On propose une méthode pour optimiser le coût de la procédure AD à travers l'exemple d'un Swap de taux qui est un instrument financier à terme.

Swap de taux Le Swap de taux est un échange périodique de taux entre deux contreparties, le taux Libor contre le taux d'exercice K . Le graphe 1 résume les Cash Flow d'un Swap.

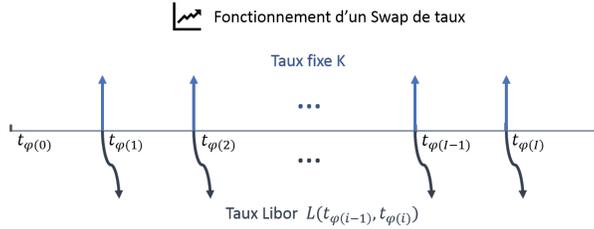


FIGURE 1: Cash Flow d'un Swap.

Soit $I > 0$ et ϕ une application de $[0, I] \cap \mathbb{N}$ dans $[0, N_t] \cap \mathbb{N}$ telle que la suite $\phi(i)_{0 \leq i \leq I}$ est strictement croissante avec $\phi(0) = 0$ et $\phi(I) = N_t$.

On note $L(t_{\phi(i-1)}, t_{\phi(i)})$ le taux Libor entre les instants $t_{\phi(i-1)}$ et $t_{\phi(i)}$ et $D(0, t_{\phi(i)})$ le discount factor par au temps $t_{\phi(i)}$.

Soit un Swap payeur qui consiste à recevoir en chaque instant $t_{\phi(i)}$ ($i > 0$) un Cash Flow de $L(t_{\phi(i-1)}, t_{\phi(i)}) - K$ (ce Cash Flow peut être négatif).

La valeur de ce Swap est

$$\mathbb{E}[P((r_{t_i})_{0 \leq i \leq N_t}, \theta)] = \mathbb{E}\left[\sum_{i=1}^I D(0, t_{\phi(i)}) (L(t_{\phi(i-1)}, t_{\phi(i)}) - K)\right].$$

L'AD classique consiste à dériver l'expression de $P((r_{t_i})_{0 \leq i \leq N_t}, \theta)$ par rapport à tous les taux courts $(r_{t_i})_{0 \leq i \leq N_t}$.

Cependant, il y a deux choses à noter :

- P ne dépend que des taux $r_{t_{\phi(i)}}$.
- $F_{\phi(i)} = L(t_{\phi(i-1)}, t_{\phi(i)}) - K$ ne dépend que de $r_{t_{\phi(i-1)}}$.

On a donc finalement besoin de calculer uniquement la dérivée de $L(t_{\phi(i-1)}, t_{\phi(i)}) - K$ par rapport à $r_{t_{\phi(i)}}$ ce qui permet de réduire considérablement le temps de calcul.

Avant mettre en place la nouvelle procédure AD qui en découle on définit d'abord la notion de dérivée conditionnelle : soit une fonction dont les taux

courts font partie des variables. Soit x un paramètre de cette fonction. On définit $\frac{\partial f}{\partial x} |_{\mathcal{F}_{t_i}}$ par la dérivée total de f par rapport à x en considérant les taux courts r_t tels que $t \leq t_i$ comme des variables propres et les taux courts r_t tels que $t > t_i$ comme des fonctions. Par exemple, on a $\frac{\partial P}{\partial r_{t_i}} |_{\mathcal{F}_{t_i}} = \frac{\partial P_i}{\partial r_{t_i}}$.

Soi

On peut construire le payoff P du swap à travers les flux en exploitant la relation de Chasles du discount factor : $D(0, s) = D(0, t)D(t, s)$ pour $0 \leq t \leq s$.

La construction se fait par mise à jour de la valeur de P initialisée à 0. On commence par rajouter le flux $F_{\phi(I)}$ à P qui est ensuite actualisé à l'instant du flux précédent par le discount factor $D(t_{\phi(I-1)}, t_{\phi(I)})$. Ensuite, on rajoute le flux $F_{\phi(I-1)}$ et on actualise à travers $D(t_{\phi(I-2)}, t_{\phi(I-1)})$. Ces opérations sont répétées jusqu'à l'étape finale où on rajoute $F_{\phi(1)}$ et on actualise par $D(0, t_{\phi(1)})$. On aboutit ainsi à la valeur P du payoff du Swap.

Cette manière de construire le payoff est couplée avec l'AD afin de calculer les dérivées : Le rajout des flux et l'actualisation sont accompagnés par une modification de la valeur de la dérivée qui est initialisée à l'instar du payoff. Le rajout de flux est précédé par quelques étapes de l'AD. Suite à des simplifications mathématiques, on peut démontrer que ces étapes se résument par les deux opérations suivantes :

$$\frac{\partial P}{\partial r_{t_{\phi(i-1)}}} |_{\mathcal{F}_{t_{\phi(i-1)}}} \leftarrow \frac{\partial P}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial r_{t_{\phi(i-1)}}} |_{\mathcal{F}_{t_{\phi(i-1)}}} \quad (12)$$

$$\frac{\partial P}{\partial \sigma} |_{\mathcal{F}_{t_{\phi(i-1)}}} \leftarrow \frac{\partial P}{\partial \sigma} + \frac{\partial P}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial \sigma} |_{\mathcal{F}_{t_{\phi(i-1)}}} \quad (13)$$

Ces deux opérations permettent de simplifier la procédure classique d'AD en réduisant le nombre total d'étapes. Elle sont possibles puisque à chaque étape le prix est une fonction de $r_{t_{\phi(i-1)}}$. Les étapes du calcul sont reprises en détail en annexe.

2.4 Adjoint Différentiation pour les instruments exotiques

Le but de cette partie est d'expliquer comment utiliser l'AD afin de calculer les sensibilités des produits dérivés à caractère exotique et ce à travers l'exemple des Bermuda Swaption de taux et en utilisant la méthode optimisée ci-haut.

Bermuda Swaption On définit d'abord

$$PS(j) = \mathbb{E} \left[\sum_{i=j+1}^I D(t_{\phi(j)}, t_{\phi(i)}) (L(t_{\phi(i-1)}, t_{\phi(i)}) - K) |_{\mathcal{F}_{t_{\phi(j)}}} \right]^+.$$

$PS(j)$ est une fonction de θ et de r_{t_i} . Le calcul de $PS(j)$ et de ses dérivées se fait à travers la méthode Least Square Monte Carlo (LSMC). La méthode est détaillée en annexe.

Néanmoins, on note qu'il est possible d'avoir une formule analytique de $PS(j)$ pour certains modèles de taux.

Une Bermuda Swaption de taux est une option sur un Swap de taux, i.e. elle donne à son détenteur le droit et non pas l'obligation d'entrer dans le contrat de Swap en les instants $t_{\phi(j)}$ ($j < I$). Si l'option est exercée, la valeur de la Bermuda Swaption est

$$\mathbb{E} \left[D(0, t_{\phi(j)}) \mathbb{E} \left[\sum_{i=j+1}^I D(t_{\phi(j)}, t_{\phi(i)}) (L(t_{\phi(i-1)}, t_{\phi(i)}) - K) \middle| \mathcal{F}_{t_{\phi(j)}} \right]^+ \right] = \mathbb{E}[D(0, t_{\phi(j)}) PS(j)].$$

C'est la valeur d'une Swaption dont les Cash Flow ne commencent qu'à l'instant t_{j+1} .

La procédure de l'AD pour une Bermuda Swaption est comme suit

- *Première étape.* On se place à l'instant finale t_{N_t} et on définit la valeur $V_{N_t} = 0$.
- *Deuxième étape* De manière générale supposons que $V_{\phi(i)}$ définie. On calcule d'abord

$$V_c(t_{\phi(i-1)}) = \mathbb{E}[D(t_{\phi(i-1)}, t_{\phi(i)}) V(t_{\phi(i)}) | \mathcal{F}_{t_{\phi(i-1)}}],$$

par LSMC. On calcule aussi ses dérivées

$$\frac{\partial V_c(t_{\phi(i-1)})}{\partial r_{t_{\phi(i-1)}}} \quad \text{et} \quad \frac{\partial V_c(t_{\phi(i-1)})}{\partial \sigma}$$

- *Troisième étape* Calcul de $PS(j-1)$ et ses dérivées

$$\frac{\partial PS(j-1)}{\partial t_{\phi(j-1)}} \quad \text{et} \quad \frac{\partial PS(j-1)}{\partial \sigma}$$

- *Quatrième étape* Définition de $V_{\phi(i-1)}$ et de ses dérivées

$$V_{\phi(i-1)} = V_c(t_{\phi(i-1)}) \mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(j-1)} + PS(j-1) \mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(j-1)}$$

$$\frac{\partial V_{\phi(i-1)}}{\partial r_{t_{\phi(i-1)}}} = \frac{\partial V_c(t_{\phi(i-1)})}{\partial r_{t_{\phi(i-1)}}} \mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(i-1)} + \frac{\partial PS(i-1)}{\partial t_{\phi(i-1)}} \mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(i-1)}$$

$$\frac{\partial V_{\phi(i-1)}}{\partial \sigma} = \frac{\partial V_c(t_{\phi(i-1)})}{\partial \sigma} \mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(i-1)} + \frac{\partial PS(i-1)}{\partial \sigma} \mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(i-1)}$$

- *Cinquième étape* À l'instar du Swap, on aboutit aux valeurs des grecs

3 Résultats numériques sur l'AD

Afin de vérifier l'exactitude de la procédure développée pour l'AD, on trace les courbes de delta et de vega de ce produit et on les compare avec les courbes générées par une méthode par différence finie centrée.

Pour la méthode par différence finie, on prend $h = 10^{-4}$ (un point de base), qui est du même ordre de h_* , pour toutes les applications.

On considère la Bermuda Swaption simple sur 5 années permettant à son détenteur d'entrer dans un contrat de Swap aux instants $t_{\phi(i)}$. Le Swap sous-jacent permet un flux financier chaque 6 mois.

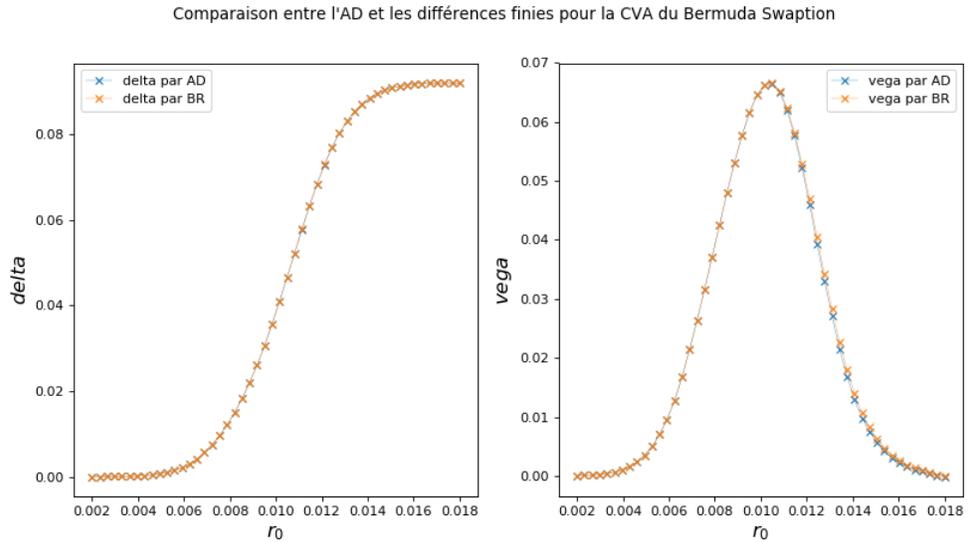


FIGURE 2: Courbes de delta et vega de la Bermuda Swaption en fonction de r_0 par AD et par différence finie

Les deux méthodes fournissent des valeurs très proches confirmant ainsi la méthode d'AD. Quantitativement, on peut majorer la différence d'estimation du delta par 7.10^{-3} et du vega par 5.10^{-2} pour chaque point à la fois pour le delta et le vega.

On peut aussi tracer l'intervalle de confiance à 95% de l'estimation de chaque point. Les deux méthodes donnent des résultats très semblables. Ci-dessous les résultats de l'AD :

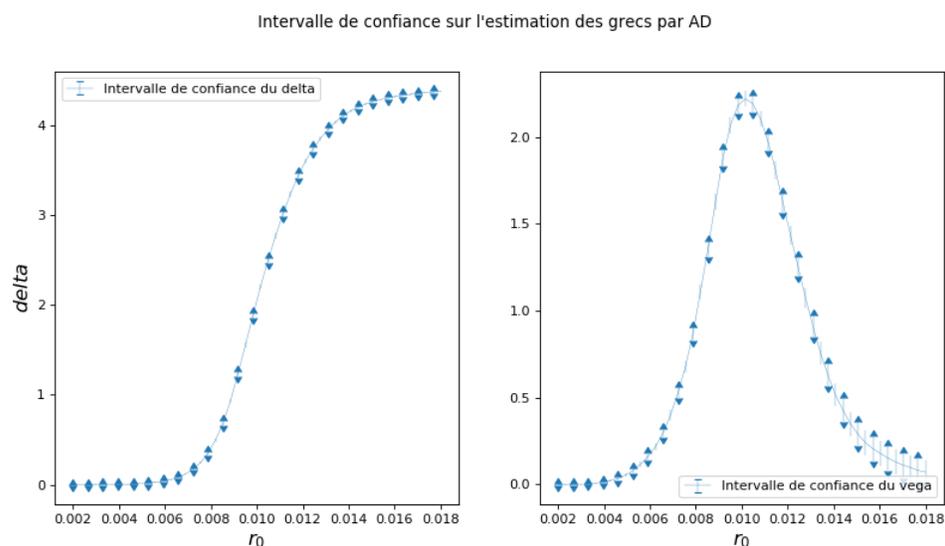


FIGURE 3: Intervalle de confiance sur l'estimation du delta et du vega de la Bermuda Swaption par AD

Produit	Caplet	Swap	Swaption	Bermuda Swaption
Temps de calcul AD	12	13	17	100
Temps de calcul différence finie sur AD	12	19	26	201

TABLE 1: Comparaison de temps de calcul (en secondes) de 50 valeurs de delta et de vega entre AD et différence finie

Comparaison des performances entre AD et différence finie Ci-dessus, on compare le temps de calcul des grecs pour les deux méthodes de dérivation utilisées. (1) permet de confirmer que l'AD est plus performant que les méthodes par différence finie, surtout pour les produits dont l'évaluation est coûteuse.

4 Credit Value Adjustment CVA

Cette partie est consacrée à la définition du Credit Value Adjustment (CVA) et à l'élaboration de méthodes d'estimation de ses dérivées pour chaque produit.

4.1 Définition de la CVA

La CVA (Credit Value Adjustment) peut être définie comme l'espérance des pertes dans un portefeuille de transactions en cas de défaut de la contrepartie.

Afin de définir la CVA de manière plus rigoureuse, considérons d'abord une unique contrepartie et notons τ la variable aléatoire définissant son temps de

défaut. τ est un temps d'arrêt.

En reprenant les notations des sections précédentes, $V_c(t)$ désigne la valeur de continuation de l'ensemble des contrats en cours avec cette contrepartie.

Soit un taux de recouvrement constant $0 \leq R \leq 1$. La CVA unilatérale, qu'on appelle simplement CVA, calculée au temps $t_0 = 0$ est définie par :

$$CVA = (1 - R)\mathbb{E}[\mathbb{1}_{0 \leq \tau < T} D(0, \tau) V_c(\tau)^+], \quad (14)$$

où T correspond à la date de la fin de tous les contrats.

En supposant dans un premier temps une absence de corrélation entre l'événement de défaut et les informations du marché et en procédant par une discrétisation temporelle $t_0 = 0 \leq T_1 \leq \dots \leq T_N = T$, on approxime la CVA par

$$CVA = (1 - R) \sum_{i=1}^{i=N} \mathbb{Q}(T_{i-1} \leq \tau < T_i) \mathbb{E}[D(0, T_i) V_c(T_i)^+], \quad (15)$$

où $\mathbb{Q}(T_{i-1} \leq \tau < T_i)$ est la probabilité d'occurrence de l'événement de défaut entre les instants T_{i-1} et T_i . Dans cette approximation, l'événement de défaut se situe en la fin de la période.

4.2 Vision générale sur le calcul de la CVA

Il y a plusieurs modèles mathématiques élaborés pour la probabilité de défaut. Néanmoins, on se place dans le cadre simple où les probabilités de défaut sont issues de l'interpolation linéaire de données historiques.

La quantité (15) est calculée par MC :

$$CVA = (1 - R) \sum_{i=1}^{i=N} \mathbb{Q}(T_{i-1} \leq \tau < T_i) \frac{\sum_{p=1}^{p=N_p} D(0, T_i)[p] V_c(T_i)[p]^+}{N_p}, \quad (16)$$

Le point clé pour le calcul de la CVA est donc de pouvoir calculer les valeur de continuation. Puisque l'AD développée ci-haut se fait au cours de la valorisation, Le calcul de la CVA est détaillé pour chaque produit en même temps que la procédure de l'AD pour la CVA.

Par la suite, on fixe $N = I$ et $T_i = t_{\phi(i)}$. En fait, cette discrétisation n'est pas importante dans le cadre de cette étude puisqu'elle ne s'intéresse pas à la probabilité de défaut spécifiquement.

On note que la suite $(T_i)_i$ doit être plus fine que $(t_{\phi(i)})_i$ et que la discrétisation de diffusion $(t_i)_i$ est plus fine que $(T_i)_i$.

4.3 Méthode de calcul des sensibilités de la CVA par AD

L'objectif de cette partie est d'adapter l'AD sur les produits financiers [2] vue précédemment afin de pouvoir l'appliquer sur la CVA en calculer de certains produits financiers de taux.

On a

$$\frac{\partial CVA}{\partial \theta} = (1 - R) \sum_{i=1}^{i=I} \frac{\partial [\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)}) \mathbb{E}[D(0, t_{\phi(i)}) V_c(t_{\phi(i)})^+]]}{\partial \theta}. \quad (17)$$

On suppose que l'événement de défaut est indépendant de θ , on a alors

$$\frac{\partial CVA}{\partial \theta} = (1 - R) \sum_{i=1}^{i=I} \mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)}) \mathbb{E}\left[\frac{\partial [D(0, t_{\phi(i)}) V_c(t_{\phi(i)})^+]}{\partial \sigma}\right]$$

Méthode pour un Swap La méthode est comme suit

- *Première étape* Initiation de la CVA : à l'instant finale $t_{\phi(I)}$, on a $CVA = 0$ et $\frac{\partial CVA}{\partial r_0}, \frac{\partial CVA}{\partial \theta} = 0$
- *Deuxième étape* : On se place au dernier instant de réception d'un flux financier. On se situe donc à l'instant $t_{\phi(I)}$, le flux reçu est $L(t_{\phi(I-1)}, t_{\phi(I)})$. On définit alors $F_{\phi(I)} = L(t_{\phi(I-1)}, t_{\phi(I)})$
- *Deuxième étape* : $F_{\phi(I)}$ est une fonction de $r_{\phi(I-1)}$ et σ . On se place donc directement à l'instant $t_{\phi(I-1)}$ et on définit

$$V_{\phi(I-1)} = D(t_{\phi(I-1)}, t_{\phi(I)}) F_{\phi(I)} \quad (18)$$

et ses dérivées

$$\frac{\partial V_{\phi(I-1)}}{\partial r_{t_{\phi(I-1)}}} = D(t_{\phi(I-1)}, t_{\phi(I)}) \frac{\partial V_{\phi(I)}}{\partial r_{t_{\phi(I-1)}}} + \frac{\partial D(t_{\phi(I-1)}, t_{\phi(I)})}{\partial r_{t_{\phi(I-1)}}} \Big|_{\mathcal{F}_{t_{\phi(I-1)}}} V_{\phi(I)} \quad (19)$$

$$\frac{\partial V_{\phi(I-1)}}{\partial \sigma} = D(t_{\phi(I-1)}, t_{\phi(I)}) \frac{\partial V_{\phi(I)}}{\partial \sigma} + \frac{\partial D(t_{\phi(I-1)}, t_{\phi(I)})}{\partial \sigma} \Big|_{\mathcal{F}_{t_{\phi(I-1)}}} V_{\phi(I)} \quad (20)$$

- *Troisième étape* De manière générale, supposons que l'on dispose de $\frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i)}}}$ et $\frac{\partial V_{\phi(i)}}{\partial \sigma}$ pour $1 \leq i \leq I - 1$. On utilise la méthode LSMC afin d'approximer

$$\mathbb{E}[V_{\phi(i)} | \mathcal{F}_{t_{\phi(i)}}]$$

Par la suite, on calcule

$$\frac{\partial \mathbb{E}[V_{\phi(i)} | \mathcal{F}_{t_{\phi(i)}}]}{\partial r_{t_{\phi(i)}}} \quad \text{et} \quad \frac{\partial \mathbb{E}[V_{\phi(i)} | \mathcal{F}_{t_{\phi(i)}}]}{\partial \sigma}$$

— *Quatrième étape* On calcule

$$\begin{aligned} V_c(t_{\phi(i)})^+ &= \mathbb{E}[V_{\phi(i)}|\mathcal{F}_{t_{\phi(i)}}]^+ \\ \frac{\partial V_c^+(t_{\phi(i)})}{\partial t_{\phi(i)}} &= \frac{\partial \mathbb{E}[V_{\phi(i)}|\mathcal{F}_{t_{\phi(i)}}]}{\partial r_{t_{\phi(i)}}} \mathbb{1}_{\mathbb{E}[V_{\phi(i)}|\mathcal{F}_{t_{\phi(i)}}] > 0} \\ \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma} &= \frac{\partial \mathbb{E}[V_{\phi(i)}|\mathcal{F}_{t_{\phi(i)}}]}{\partial \sigma} \mathbb{1}_{\mathbb{E}[V_{\phi(i)}|\mathcal{F}_{t_{\phi(i)}}] > 0} \end{aligned}$$

— *Cinquième étape* On applique l'AD pour calculer $\frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}}$ et $\frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_0}}$:

$$\begin{aligned} \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} &= \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} \\ \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} &= \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma} + \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial \sigma}|_{\mathcal{F}_{t_0}} \end{aligned}$$

Ensuite, on applique la transformation

$$\begin{aligned} \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_0}} &\leftarrow \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} \\ \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma} &\leftarrow \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} \end{aligned}$$

— *Sixième étape* On définit

$$EE(t_{\phi(i)}) = D(t_0, t_{\phi(i)})V_c(t_{\phi(i)})^+$$

et ses dérivées

$$\begin{aligned} \frac{\partial EE(t_{\phi(i)})}{\partial r_{t_0}} &= D(t_0, t_{\phi(i)}) \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_0} + \frac{\partial D(t_0, t_{\phi(i)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} V_c^+(t_{\phi(i)}) \\ \frac{\partial EE(t_{\phi(i)})}{\partial \sigma} &= D(t_0, t_{\phi(i)}) \frac{\partial V_c^+(t_{\phi(i)})}{\partial \sigma} + \frac{\partial D(t_0, t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} V_c^+(t_{\phi(i)}) \end{aligned}$$

— *Septième étape* On calcule les moyennes par Monte Carlo

$$\mathbb{E}\left[\frac{\partial EE(t_{\phi(i)})}{\partial r_{t_0}}\right] \quad \text{et} \quad \mathbb{E}\left[\frac{\partial EE(t_{\phi(i)})}{\partial \sigma}\right],$$

puis on met à jour la CVA et ses dérivées

$$\begin{aligned} CVA &\leftarrow CVA + (1-R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}[EE(t_{\phi(i)})] \\ \frac{\partial CVA}{\partial r_0} &\leftarrow \frac{\partial CVA}{\partial r_0} + (1-R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}\left[\frac{\partial EE(t_{\phi(i)})}{\partial r_{t_0}}\right] \\ \frac{\partial CVA}{\partial \sigma} &\leftarrow \frac{\partial CVA}{\partial \sigma} + (1-R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}\left[\frac{\partial EE(t_{\phi(i)})}{\partial \sigma}\right] \end{aligned}$$

- *Huitième étape* On se place à l'instant $t_{\phi(i-1)}$ et on applique l'AD pour calculer $\frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}}$ et $\frac{\partial V_{\phi(i)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}}$:

$$\begin{aligned}\frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} &= \frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} \\ \frac{\partial V_{\phi(i)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} &= \frac{\partial V_{\phi(i)}}{\partial \sigma} + \frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}}\end{aligned}$$

Ensuite, on applique la transformation

$$\begin{aligned}\frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} &\leftarrow \frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} + \frac{\partial L(t_{\phi(i-1)}, t_{\phi(i)})}{\partial r_{t_{\phi(i-1)}}} \\ \frac{\partial V_{\phi(i)}}{\partial \sigma} &\leftarrow \frac{\partial V_{\phi(i)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} + \frac{\partial L(t_{\phi(i-1)}, t_{\phi(i)})}{\partial \sigma}\end{aligned}$$

- *Neuvième étape* On calcule

$$V_{\phi(i-1)} = D(t_{\phi(i-1)}, t_{\phi(i)})V_{\phi(i)}$$

et ses dérivées

$$\begin{aligned}\frac{\partial V_{\phi(i-1)}}{\partial r_{t_{\phi(i-1)}}} &= D(t_{\phi(i-1)}, t_{\phi(i)}) \frac{\partial V_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} + \frac{\partial D(t_{\phi(i-1)}, t_{\phi(i)})}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} V_{\phi(i)} \\ \frac{\partial V_{\phi(i-1)}}{\partial \sigma} &= D(t_{\phi(i-1)}, t_{\phi(i)}) \frac{\partial V_{\phi(i)}}{\partial \sigma} + \frac{\partial D(t_{\phi(i-1)}, t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} V_{\phi(i)}\end{aligned}$$

- *Dixième étape* Si $i = 1$, on ignore l'étape précédente et on définit

$$V_0 = V_{\phi(0)} = D(t_{\phi(0)}, t_{\phi(1)})V_{\phi(1)}$$

et ses dérivées

$$\begin{aligned}\frac{\partial V_{\phi(0)}}{\partial r_{t_{\phi(0)}}} &= D(t_{\phi(0)}, t_{\phi(1)}) \frac{\partial V_{\phi(1)}}{\partial r_{t_{\phi(0)}}} + \frac{\partial D(t_{\phi(0)}, t_{\phi(1)})}{\partial r_{t_{\phi(0)}}}|_{\mathcal{F}_{t_{\phi(0)}}} V_{\phi(1)} \\ \frac{\partial V_{\phi(0)}}{\partial \sigma} &= D(t_{\phi(0)}, t_{\phi(1)}) \frac{\partial V_{\phi(1)}}{\partial \sigma} + \frac{\partial D(t_{\phi(0)}, t_{\phi(1)})}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(0)}}} V_{\phi(1)}\end{aligned}$$

- *Onzième étape* On calcule les grecs

$$\begin{aligned}\frac{\partial V}{\partial r_0} &= \mathbb{E}\left[\frac{\partial V_0}{\partial r_{t_0}}\right] \\ \frac{\partial V}{\partial \sigma} &= \mathbb{E}\left[\frac{\partial V_0}{\partial \sigma}\right]\end{aligned}$$

Méthode pour une Bermuda Swaption

- *Première étape* Initiation de la CVA : à l'instant finale $t_{\phi(I)}$, on a $CVA = 0$ et $\frac{\partial CVA}{\partial r_0}, \frac{\partial CVA}{\partial \sigma} = 0$
On définit la valeur à l'instant finale $V_{N_t} = 0$.
- *Deuxième étape* De manière générale, on suppose que $V_{\phi(i)}$ définie. On calcule d'abord

$$V_c(t_{\phi(i-1)}) = \mathbb{E}[D(t_{\phi(i-1)}, t_{\phi(i)})V(t_{\phi(i)})|\mathcal{F}_{t_{\phi(i-1)}}],$$

par LSMC. On calcule aussi ses dérivées

$$\frac{\partial V_c(t_{\phi(i-1)})}{\partial r_{t_{\phi(i-1)}}} \quad \text{et} \quad \frac{\partial V_c(t_{\phi(i-1)})}{\partial \sigma}$$

On calcule

$$V_c(t_{\phi(i-1)})^+ = \mathbb{E}[V_{\phi(i-1)}|\mathcal{F}_{t_{\phi(i-1)}}]^+$$

$$\frac{\partial V_c^+(t_{\phi(i-1)})}{\partial t_{\phi(i-1)}} = \frac{\partial \mathbb{E}[V_{\phi(i-1)}|\mathcal{F}_{t_{\phi(i-1)}}]}{\partial r_{t_{\phi(i-1)}}} \mathbb{1}_{\mathbb{E}[V_{\phi(i-1)}|\mathcal{F}_{t_{\phi(i-1)}}] > 0}$$

$$\frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma} = \frac{\partial \mathbb{E}[V_{\phi(i-1)}|\mathcal{F}_{t_{\phi(i-1)}}]}{\partial \sigma} \mathbb{1}_{\mathbb{E}[V_{\phi(i-1)}|\mathcal{F}_{t_{\phi(i-1)}}] > 0}$$

- *Troisième étape* On applique l'AD pour calculer $\frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}}$ et $\frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma}|_{\mathcal{F}_{t_0}}$:

$$\begin{aligned} \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} &= \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_{t_{\phi(i-1)}}} \frac{\partial r_{t_{\phi(i-1)}}}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} \\ \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} &= \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma} + \frac{\partial V_c^+(t_{\phi(i)})}{\partial r_{t_{\phi(i-1)}}} \frac{\partial r_{t_{\phi(i-1)}}}{\partial \sigma}|_{\mathcal{F}_{t_0}} \end{aligned}$$

Ensuite, on applique la transformation

$$\begin{aligned} \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_{t_0}} &\leftarrow \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} \\ \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma} &\leftarrow \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} \end{aligned}$$

- *Quatrième étape* On définit

$$EE(t_{\phi(i-1)}) = D(t_0, t_{\phi(i-1)})V_c(t_{\phi(i-1)})^+$$

et ses dérivées

$$\begin{aligned} \frac{\partial EE(t_{\phi(i-1)})}{\partial r_{t_0}} &= D(t_0, t_{\phi(i-1)}) \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial r_0} + \frac{\partial D(t_0, t_{\phi(i-1)})}{\partial r_{t_0}}|_{\mathcal{F}_{t_0}} V_c^+(t_{\phi(i-1)}) \\ \frac{\partial EE(t_{\phi(i-1)})}{\partial \sigma} &= D(t_0, t_{\phi(i-1)}) \frac{\partial V_c^+(t_{\phi(i-1)})}{\partial \sigma} + \frac{\partial D(t_0, t_{\phi(i-1)})}{\partial \sigma}|_{\mathcal{F}_{t_0}} V_c^+(t_{\phi(i-1)}) \end{aligned}$$

— *Cinquième étape* On calcule les moyennes par Monte Carlo

$$\mathbb{E}\left[\frac{\partial EE(t_{\phi(i-1)})}{\partial r_{t_0}}\right] \quad \text{et} \quad \mathbb{E}\left[\frac{\partial EE(t_{\phi(i-1)})}{\partial \sigma}\right],$$

puis on met à jour la CVA et ses dérivées

$$\begin{aligned} CVA &\leftarrow CVA + (1 - R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}[EE(t_{\phi(i-1)})] \\ \frac{\partial CVA}{\partial r_0} &\leftarrow \frac{\partial CVA}{\partial r_0} + (1 - R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}\left[\frac{\partial EE(t_{\phi(i-1)})}{\partial r_{t_0}}\right] \\ \frac{\partial CVA}{\partial \sigma} &\leftarrow \frac{\partial CVA}{\partial \sigma} + (1 - R)\mathbb{Q}(t_{\phi(i-1)} \leq \tau \leq t_{\phi(i)})\mathbb{E}\left[\frac{\partial EE(t_{\phi(i-1)})}{\partial r_{t_0}}\right] \end{aligned}$$

— *Sixième étape* Calcul de PS(i-1) et ses dérivées (reprendre la procédure pour la Swaption)

$$\frac{\partial PS(i-1)}{\partial t_{\phi(i-1)}} \quad \text{et} \quad \frac{\partial PS(i-1)}{\partial \sigma}$$

— *Septième étape* Définition de $V_{\phi(i-1)}$ et de ses dérivées

$$\begin{aligned} V_{\phi(i-1)} &= V_c(t_{\phi(i-1)})\mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(i-1)} + PS(i-1)\mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(i-1)} \\ \frac{\partial V_{\phi(i-1)}}{\partial r_{t_{\phi(i-1)}}} &= \frac{\partial V_c(t_{\phi(i-1)})}{\partial r_{t_{\phi(i-1)}}}\mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(i-1)} + \frac{\partial PS(i-1)}{\partial t_{\phi(i-1)}}\mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(i-1)} \\ \frac{\partial V_{\phi(i-1)}}{\partial \sigma} &= \frac{\partial V_c(t_{\phi(i-1)})}{\partial \sigma}\mathbb{1}_{V_c(t_{\phi(i-1)}) \geq PS(i-1)} + \frac{\partial PS(i-1)}{\partial \sigma}\mathbb{1}_{V_c(t_{\phi(i-1)}) < PS(i-1)} \end{aligned}$$

— *Huitième étape* À l'instar des autres produits on aboutit aux valeurs des grecs

5 Résultats numériques sur la CVA

5.1 Calcul de la CVA

L'étape la plus importante dans le calcul de la CVA est l'estimation de $V_c(T_i)$ par LSMC. Ci dessous, on fait une étude simple permettant de valider pratiquement la méthode.

Considérons le Caplet qui permet à son détenteur de recevoir un flux financier à l'instant t_{N_t} égal à $L(t_{\phi(I-1)}, t_{\phi(I)}) - K$.

Dans un premier temps, on calcule la valeur de l'exposition, i.e $\mathbb{E}[V_c(t_i)^+]$ aux différents instants $t_{\phi(i)}$. On a :

$$V_c(t_{\phi(i)})^+ = \mathbb{E}[(L(t_{\phi(I-1)}, t_{\phi(I)}) - K)^+ | \mathcal{F}_{t_{\phi(i)}}] \quad (21)$$

L'estimation de (21) est faite d'abord par LSMC en utilisant, d'après la littérature, les 3 premiers polynômes de Laguerre comme fonctions de base.

Ensuite, on utilise la méthode MC^2 afin d'estimer (21). Le nombre de simulations et de simulations étant égal à $N_p = 10^4$, le nombre total de trajectoires générées est 10^8 . On voit donc que cette méthode est onéreuse.

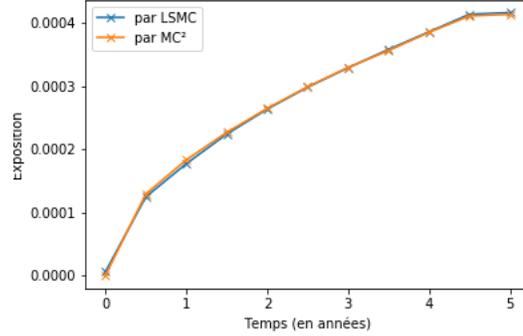


FIGURE 4: Comparaison de l'exposition $V_c(t_{\phi(i)})^+$ par MC^2 et LSMC.

La figure 4 permet de comparer les résultats des deux méthodes. On définit la largeur relative de l'intervalle de confiance de la quantité estimée par LSMC, à savoir $\frac{\text{largeur de l'intervalle de confiance à 95\%}}{\text{moyenne estimée}}$ par $2 \times 1.96 \frac{\text{écart-type}}{\sqrt{N_p} \text{moyenne}}$. Pour chaque instant $t_{\phi(i)}$ et pour chacune des deux méthodes, cette quantité est inférieure à 5.8%. Afin d'améliorer l'estimation par *LSMC*, il faut augmenter le nombre de fonctions de base, cependant, on se heurte à des problèmes d'inversion de la matrice (C_{nm}) . D'après la littérature, le nombre de simulations nécessaires pour la convergence de la méthode augmente très rapidement en fonction du nombre de fonctions de base. L'intervalle de confiance de l'estimation par MC^2 (Voir Annexe) contient l'estimation par *LSMC*, on peut donc confirmer les résultats de LSMC pour le Caplet.

Pour le calcul de la CVA, on prend $LGD = 0.6$ et les probabilités de défaut historiques spécifiées dans le graphe 5 ci-dessous

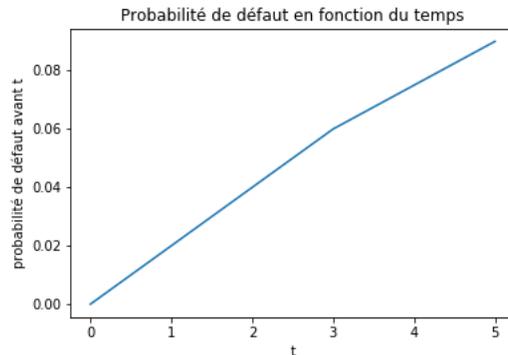


FIGURE 5: Valeurs des probabilités de défaut historiques.

Les données ci-dessus permettent d’estimer, par la méthode MC^2 , la valeur de la CVA à $1.51e - 05$.

On propose ensuite d’estimer la CVA à l’aide de la méthode LSMC. On choisit trois familles de fonctions de base : les polynômes de Laguerre, les polynômes de Legendre et les puissances de x ($x^i, i \geq 0$). Pour chaque famille de polynômes, on fait varier le nombre de fonctions de base utilisées dans la méthode et on calcule la valeur estimée de la CVA. Au-delà de 5 polynômes la qualité de l’estimation stagne pratiquement. Pour chacune des bases utilisées, la valeur de la CVA est environ $1.51e - 5$. Elle présente une erreur relative très petite par rapport à la valeur estimée par MC^2 .

On propose maintenant de quantifier la stabilité du calcul de la CVA pour les différentes bases à travers l’intervalle de confiance relatif à 95%. Ici encore, on constate une largeur relative presque invariante de l’ordre de 3.4%. On justifie le choix des 3 premiers polynômes de Laguerre pour les calculs par le fait que la largeur relative de l’intervalle de confiance obtenue est plus petite.

On s’aperçoit, suite aux résultats précédents, que la méthode LSMC fournit une bonne estimation de l’exposition et de la CVA.

5.2 Calcul des sensibilités de la CVA pour une Bermuda Swaption

En gardant les mêmes données numériques utilisées auparavant, on trace les graphes de delta et de vega pour la CVA d’une Bermuda Swaption. On aboutit au graphe

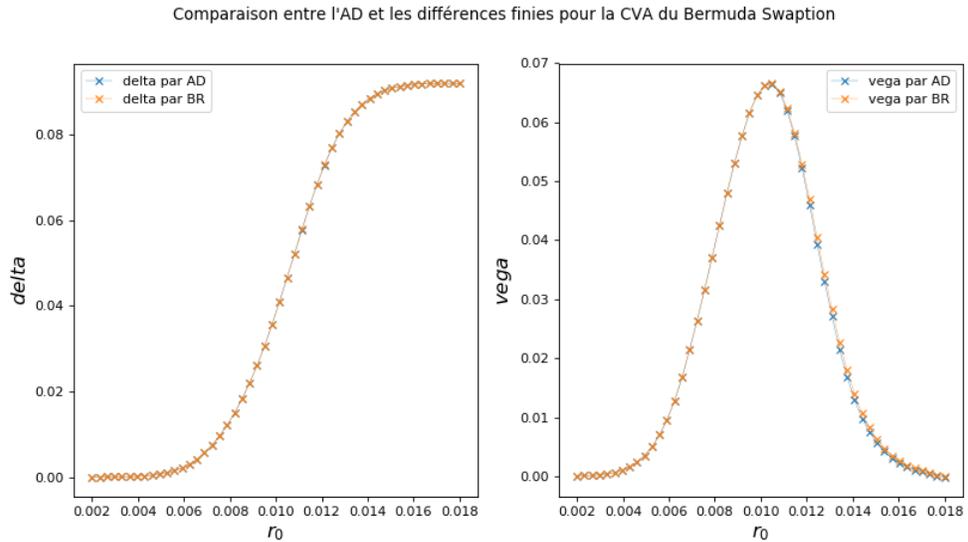


FIGURE 6: Courbes de delta et vega de la CVA de la Bermuda Swaption en fonction de r_0 par AD et par différence finie

Les deux méthodes fournissent des valeurs très proches confirmant ainsi la méthode d'AD. Quantitativement, on peut majorer la différence d'estimation par 10^{-3} pour chaque point à la fois pour le delta et le vega. On peut aussi tracer l'intervalle de confiance à 95% de l'estimation de chaque point. Les deux méthodes donnent des résultats très semblables. Ci-dessous les résultats de l'AD :

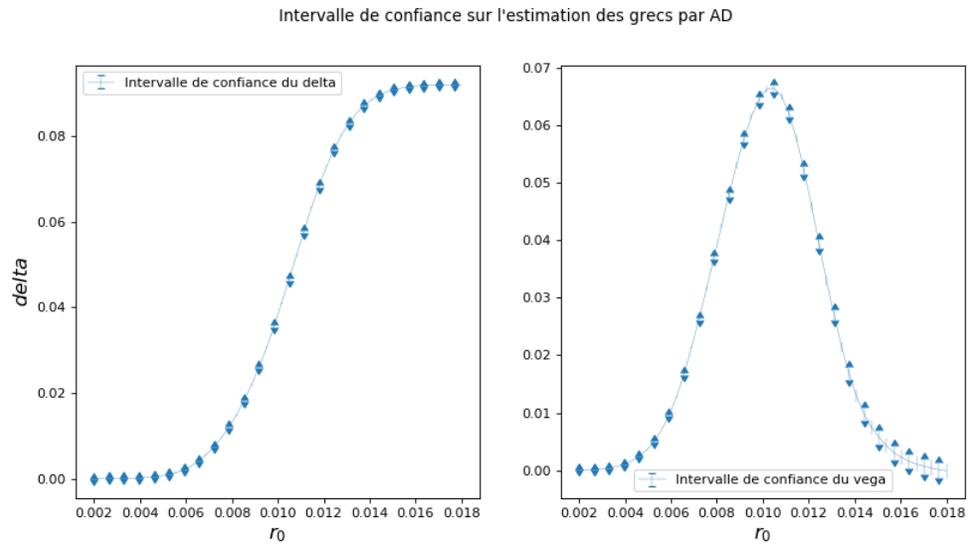


FIGURE 7: Intervalle de confiance sur l'estimation du delta du vega de la CVA de la Bermuda Swaption par AD

Les résultats ci-dessous confirment l'exactitude de l'estimation par AD de la CVA du Bermuda Swaption. Par ailleurs, les calculs par AD sont deux fois plus rapides.

Conclusion

Cet article a permis de voir en théorie et en pratique comment calculer les sensibilités pour les instruments financiers de taux suivants : Swap et Bermuda Swaption. Les méthodes par différence finie présentent le défi du choix de h . Par ailleurs, pour les instruments exotiques, un bon choix du nombre et de la nature des fonctions de base est crucial pour la bonne estimation du payoff à l'aide la méthode Least Squares Monte Carlo. Une bonne estimation du payoff est, quant à elle, primordiale pour une bonne évaluation des grecs par Adjoint Différentiation et par différence finie.

À l'instar des produits financiers, l'AD permet aussi d'approximer la valeur des sensibilités de la CVA, même pour des produits de taux exotiques. On a aboutit à des estimations du delta et du vega d'une très bonne concordance entre les deux méthodes. L'estimation des dérivées de second ordre, étant moins précise, est un point d'amélioration crucial.

Globalement, l'AD est avantageuse par rapport aux méthodes par différence finie lorsque la valorisation est coûteuse, du fait que l'AD requiert la valorisation pour un seul jeu de paramètres. Ceci permet de réduire considérablement le temps de calcul par rapport aux méthodes par différence finie qui requièrent deux valorisations.

Références

1. Glasserman, P. (2013). Monte Carlo methods in financial engineering (Vol. 53). Springer Science & Business Media.
2. Antonov, A. (2017). Algorithmic differentiation for callable exotics. Available at SSRN 2839362.
3. Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American options by simulation : a simple least-squares approach. *The review of financial studies*, 14(1), 113-147.

A Least Square Monte Carlo

La méthode LSMC a été développée par Longstaff et Schwartz [3] pour calculer la valeur des options américaines. La méthode permet d'estimer numériquement l'espérance conditionnelle.

Justification de la méthode Soit $0 \leq i \leq N_t - 1$ et soit $j > i$. On sait que $r_{t_j} = G(r_{t_{j-1}}, Z_{j-1}, \theta)$. Par récurrence, on arrive à exprimer r_{t_j} en fonction de r_{t_i} , $(Z_k, i \leq k \leq j - 1)$ et θ . On déduit donc que $\mathbb{E}[r_{t_j} | \mathcal{F}_{t_i}]$ est une fonction de r_{t_i} .

De manière générale, soit f une fonction quelconque de $\mathbb{R}^{N_t - i + 1}$ vers \mathbb{R} , on peut facilement prouver que $\mathbb{E}[f(r_{t_i}, \dots, r_{t_{N_t}}) | \mathcal{F}_{t_i}]$ est une fonction de r_{t_i} et θ .

Principe de la méthode Soit $\{\phi_n, 1 \leq n \leq N_b\}$ un ensemble de fonctions réels. La méthode *LSMC* consiste à approximer $\mathbb{E}[f(r_{t_i}, \dots, r_{t_{N_t}}) | \mathcal{F}_{t_i}]$ par

$$\sum_{n=1}^{N_b} \alpha_n^* \phi_n(r_{t_i}), \quad (22)$$

où $\alpha^* = (\alpha_n^*)_{1 \leq n \leq N_b}$ est tel que

$$\alpha^* = \underset{\alpha \in \mathbb{R}^{N_b}}{\operatorname{argmin}} \mathbb{E}[(f(r_{t_i}, \dots, r_{t_{N_t}}) - \sum_{n=1}^{N_b} \alpha_n \phi_n(r_{t_i}))^2].$$

On démontre alors que

$$\alpha_n^* = \sum_m C_{nm}^{-1} \mathbb{E}[\phi_m(r_{t_i}) f(r_{t_i}, \dots, r_{t_{N_t}})],$$

où la matrice C est telle que

$$C_{nm} = \mathbb{E}[\phi_n(r_{t_i}) \phi_m(r_{t_i})].$$

La méthode LSMC permet de valoriser les Bermuda Swaption. Elle intervient d'avantage dans les calculs de sensibilités. Pour ce, il faut expliciter les formules des dérivées de (22) par rapport à r_{t_i} et θ .

Soit la fonction g telle que $\mathbb{E}[f(r_{t_i}, \dots, r_{t_{N_t}}) | \mathcal{F}_{t_i}] = \mathbb{E}[g(r_{t_i}, \theta, (Z_k, i \leq k < j)) | \mathcal{F}_{t_i}]$. Supposons que l'on dispose des dérivées totales de la fonction f par rapport r_{t_i} et θ , i.e. les dérivées

$$\frac{\partial g(r_{t_i}, \theta, (Z_k, i \leq k < j))}{\partial r_{t_i}} \quad \text{et} \quad \frac{\partial g(r_{t_i}, \theta, (Z_k, i \leq k < j))}{\partial \theta}$$

[3] explique comment calculer

$$\frac{\partial \alpha_n^*}{\partial r_{t_i}}, \quad (23)$$

en utilisant $\frac{\partial g(r_{t_i}, \theta, (Z_k, i \leq k < j))}{\partial r_{t_i}}$. Par ailleurs, on calcule

$$\frac{\partial \alpha_n^*}{\partial \theta} \quad (24)$$

en utilisant $\frac{\partial g(r_{t_i}, \theta, (Z_k, i \leq k < j))}{\partial \theta}$. (A) et (24) permettent de calculer facilement (22).

B La méthode MC^2

Afin d'estimer la valeur d'une espérance conditionnelle d'une quantité qui dépend exclusivement du processus $(r_t)_t$, sachant l'information disponible jusqu'au temps t_i , on commence par simuler $(r_t)_t$ jusqu'au temps t_i inclus. Ensuite, on se sert de chaque valeur simulée de r_{t_i} comme point initial de resimulation (un nouveau r_0) de trajectoires de taux simulées jusqu'à l'instant final t_{N_t} . La figure 8 ci-dessous illustre le principe de resimulation.

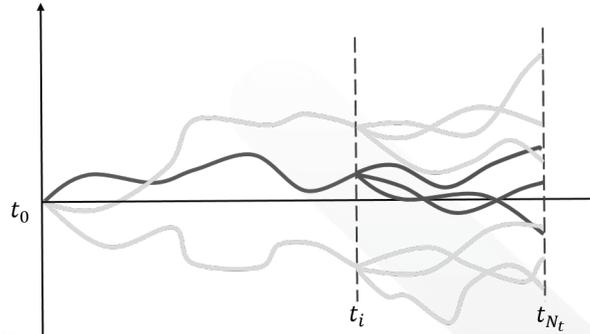


FIGURE 8: Illustration du principe de la méthode MC^2 .

Pour chacune trajectoires resimulées, un payoff est calculé et la moyenne des payoff fournit la valeur l'espérance conditionnelle. La méthode MC^2 assure les résultats fournis par LSMC.

C Modèle de diffusion des taux courts : Vasicek

Le taux court est un outil théorique permettant de simuler l'évolution de certaines quantités dans le marché financier. Il existe plusieurs modèles de taux courts, chaque modèle possède ses propres paramètres et sa propre équation différentielle stochastique (EDS). Le modèle de Vasicek pour le taux court $(r_t)_{t \geq 0}$ a pour EDS

$$dr_t = a(b - r_t)dt + \sigma dW_t, \quad \forall t \geq 0. \quad (25)$$

$\Theta = (a, b, \sigma)$ est le vecteur de paramètres du modèle, a est appelé coefficient de retour à la moyenne b et σ constitue la volatilité du modèle. Par ailleurs, $(W_t)_t$

est un mouvement brownien. Le processus de détermination des paramètres est appelé calibrage du modèle. Le but du calibrage est de pouvoir générer des taux courts fidèles à l'état du marché financier, c'est à dire de pouvoir répliquer les données du marché à l'instar de l'évolution temporelle du prix de l'obligation zéro coupon.

Pour les simulations des taux courts avec le modèle de Vasicek, on utilise le vecteur de paramètres $(10^{-2}, 10^{-3}, 10^{-3})$. Le but étant de mettre en place une méthode de dérivation mathématique sans décrire correctement le marché, ces paramètres ne sont pas issus d'un travail de calibrage. Ils sont néanmoins raisonnables.

Soit $N_t > 0$, on note $(t_i)_{0 \leq i \leq N_t}$ la suite équidistante telle que $t_0 = 0$ et $t_{N_t} = T$. Un schéma d'Euler pour (25) est

$$r_{t_{i+1}} = r_{t_i} + a(b - r_{t_i})\Delta t + \sigma\sqrt{\Delta t}Z_i \quad (26)$$

où $(Z_i)_{0 \leq i \leq N_t}$ est une suite de variables aléatoires indépendantes identiquement distribuées suivant la loi normale centrée réduite et $\Delta t = t_{i+1} - t_i$ est le pas de temps. Cette discrétisation temporelle de l'EDS permet de simuler numériquement des trajectoires probables d'évolution du taux court dans le temps. Un autre schéma de discrétisation possible est le schéma de Milstein qui est plus précis. Il est néanmoins équivalent au schéma d'Euler dans le cas du modèle de Vasicek.

L'équation (26) présente un schéma d'évolution Markovien au sens où la valeur du taux court à l'instant suivant $r_{t_{i+1}}$ ne dépend que de l'information à l'instant précédent (r_{t_i}, Z_i) . On peut ainsi définir la fonction G telle que

$$r_{t_{i+1}} = G(r_{t_i}, Z_i, \theta), \quad 0 \leq i < N_t, \quad (27)$$

où $\theta \in \Theta$ est le paramètre du modèle auquel s'intéresse l'étude. Les autres paramètres, ainsi que le pas de discrétisation dt , sont considérés constants.

C.1 Prix de l'obligation zéro coupon

Le prix de l'obligation zéro coupon entre deux instants $t < s$ est défini par

$$ZCB(t, s) = \mathbb{E}[e^{-\int_t^s r_u du} | \mathcal{F}_t],$$

où $(\mathcal{F}_t)_t$ est une filtration telle que \mathcal{F}_t rassemble toutes les informations du marché disponibles jusqu'à l'instant t , incluant l'évolution du taux court. Le processus $(r_t)_t$ est \mathcal{F}_t -adapté.

Sous le modèle de Vasicek, $P(t, s)$ a pour expression la formule fermée

$$ZCB(t, s) = e^{-R_\infty(s-t) + (R_\infty - r_t) \frac{1 - e^{-a(s-t)}}{a} - \frac{\sigma^2}{4a^3} (1 - e^{-a(s-t)})^2},$$

avec $R_\infty = b - \frac{\sigma^2}{2a^2}$.

Le prix de l'obligation zéro coupon permet de définir les Cash Flow (flux de trésorerie) des instruments financiers.

C.2 Discount factor

Le discount factor $D(t, s)$ permet de calculer la valeur, actualisée à l'instant t , d'une quantité Q perçue à l'instant s . Le discount factor entre les instants t et s est

$$D(t, s) = e^{-\int_t^s r_u du}.$$

La valeur actuelle de Q est donc

$$D(t, s)Q.$$

On suppose que $t = t_q$ et $s = t_r$ pour certains indices $0 \leq q < r \leq N_t$. Dans la littérature, on approxime souvent $D(t, s)$ par

$$e^{-dt \sum_{k=q}^{r-1} r_{t_k}}.$$

Cette approximation est utilisée tout au long du rapport.

C.3 Taux Libor

Créé par la British Banker's Association, le London Interbank Offer Rate (Libor) est une moyenne des taux d'intérêts prévalant sur le marché monétaire londonien. Il est indicatif du coût de financement des grandes banques mondiales. Le taux Libor entre les instants t et s est défini par

$$L(t, s) = \frac{1}{s-t} \left(\frac{1}{ZCB(t, s)} - 1 \right).$$

Les quantités décrites ci-dessus servent de base pour définir les instruments financiers se basant sur les taux d'intérêt.

D Procédure d'AD pour un Swap

Pour un swap de taux, la procédure AD est la suivante :

- *Première étape* : On se place à l'instant $t_{\phi(I)}$, le flux reçu est $L(t_{\phi(I-1)}, t_{\phi(I)})$. On définit alors $P_{\phi(I)} = L(t_{\phi(I-1)}, t_{\phi(i)})$
- *Deuxième étape* : $P_{\phi(I)}$ est une fonction de $r_{\phi(I-1)}$ et θ . On se place donc directement à l'instant $t_{\phi(I-1)}$ et on définit

$$P_{\phi(I-1)} = D(t_{\phi(I-1)}, t_{\phi(I)})P_{\phi(I)} \quad (28)$$

et ses dérivées

$$\frac{\partial P_{\phi(I-1)}}{\partial r_{t_{\phi(I-1)}}} = D(t_{\phi(I-1)}, t_{\phi(I)}) \frac{\partial P_{\phi(I)}}{\partial r_{t_{\phi(I-1)}}} + \frac{\partial D(t_{\phi(I-1)}, t_{\phi(I)})}{\partial r_{t_{\phi(I-1)}}} \Big|_{\mathcal{F}_{t_{\phi(I-1)}}} P_{\phi(I)} \quad (29)$$

$$\frac{\partial P_{\phi(I-1)}}{\partial \sigma} = D(t_{\phi(I-1)}, t_{\phi(I)}) \frac{\partial P_{\phi(I)}}{\partial \sigma} + \frac{\partial D(t_{\phi(I-1)}, t_{\phi(I)})}{\partial \sigma} \Big|_{\mathcal{F}_{t_{\phi(I-1)}}} P_{\phi(I)} \quad (30)$$

- *Troisième étape* Supposons que l'on dispose de $\frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i)}}}$ et $\frac{\partial P_{\phi(i)}}{\partial \sigma}$ pour $1 \leq i \leq I - 1$. On se place à l'instant $t_{\phi(i-1)}$.

On applique l'AD pour calculer $\frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i)}}}$ et $\frac{\partial P_{\phi(i-1)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}}$:

$$\frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} = \frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} \quad (31)$$

$$\frac{\partial P_{\phi(i)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} = \frac{\partial P_{\phi(i)}}{\partial \sigma} + \frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i)}}} \frac{\partial r_{t_{\phi(i)}}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} \quad (32)$$

Ensuite, on applique la transformation

$$\frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} \leftarrow \frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i-1)}}} + \frac{\partial L(t_{\phi(i-1)}, t_{\phi(i)})}{\partial r_{t_{\phi(i-1)}}} \quad (33)$$

$$\frac{\partial P_{\phi(i)}}{\partial \sigma} \leftarrow \frac{\partial P_{\phi(i)}}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} + \frac{\partial L(t_{\phi(i-1)}, t_{\phi(i)})}{\partial \sigma} \quad (34)$$

- *Quatrième étape* On calcule

$$P_{\phi(i-1)} = D(t_{\phi(i-1)}, t_{\phi(i)})P_{\phi(i)} \quad (35)$$

et ses dérivées

$$\frac{\partial P_{\phi(i-1)}}{\partial r_{t_{\phi(i-1)}}} = D(t_{\phi(i-1)}, t_{\phi(i)}) \frac{\partial P_{\phi(i)}}{\partial r_{t_{\phi(i-1)}}} + \frac{\partial D(t_{\phi(i-1)}, t_{\phi(i)})}{\partial r_{t_{\phi(i-1)}}}|_{\mathcal{F}_{t_{\phi(i)}}} P_{\phi(i)} \quad (36)$$

$$\frac{\partial P_{\phi(i-1)}}{\partial \sigma} = D(t_{\phi(i-1)}, t_{\phi(i)}) \frac{\partial P_{\phi(i)}}{\partial \sigma} + \frac{\partial D(t_{\phi(i-1)}, t_{\phi(i)})}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(i-1)}}} P_{\phi(i)} \quad (37)$$

- *Cinquième étape* Si $i = 1$, on ignore l'étape précédente et on définit

$$P_0 = P_{\phi(0)} = D(t_{\phi(0)}, t_{\phi(1)})P_{\phi(1)} \quad (38)$$

et ses dérivées

$$\frac{\partial P_{\phi(0)}}{\partial r_{t_{\phi(0)}}} = D(t_{\phi(0)}, t_{\phi(1)}) \frac{\partial P_{\phi(1)}}{\partial r_{t_{\phi(0)}}} + \frac{\partial D(t_{\phi(0)}, t_{\phi(1)})}{\partial r_{t_{\phi(0)}}}|_{\mathcal{F}_{t_{\phi(0)}}} P_{\phi(1)} \quad (39)$$

$$\frac{\partial P_{\phi(0)}}{\partial \sigma} = D(t_{\phi(0)}, t_{\phi(1)}) \frac{\partial P_{\phi(1)}}{\partial \sigma} + \frac{\partial D(t_{\phi(0)}, t_{\phi(1)})}{\partial \sigma}|_{\mathcal{F}_{t_{\phi(0)}}} P_{\phi(1)} \quad (40)$$

- *Sixième étape* On calcule les grecs

$$\frac{\partial V}{\partial r_0} = \mathbb{E}\left[\frac{\partial P_0}{\partial r_{t_0}}\right] \quad (41)$$

$$\frac{\partial V}{\partial \sigma} = \mathbb{E}\left[\frac{\partial P_0}{\partial \sigma}\right] \quad (42)$$