

Calcul des sensibilités de la CVA (Credit Value Adjustment) via la méthode AAD (Adjoint Algorithmic Differentiation).

Hervé AZEVEDO CHAVES



14 septembre 2017

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage et qui m'ont aidé lors de la rédaction de ce rapport.

Tout d'abord, j'adresse mes remerciements à mes professeurs, Mr Aurélien Alfonsi et Mr Ahmed Kebaier qui ont su me conseiller dans ma recherche de stage. Leur écoute et leurs conseils m'ont permis d'orienter mes choix, et de trouver ce stage qui était en totale adéquation avec mes attentes.

Je remercie par la même occasion tous mes professeurs de l'École Sup Galilée, de L'École Nationale des Ponts et Chaussées, de l'Université de Paris XIII ainsi que de l'Université de Paris-Est Marne La Vallée.

Je tiens à remercier mon maître de stage, Mr Hicham El Mekeddem, Consultant Manager au sein de l'entreprise Quanteam, pour son accueil, le temps passé ensemble et le partage de son expertise. Grâce aussi à sa confiance j'ai pu m'accomplir totalement dans mes missions. Il fut d'une aide précieuse dans les moments les plus délicats.

Enfin, je tiens à remercier ma camarade de promotion Linda Chamakh pour tous ces conseils.

Table des matières

1	Introduction	4
2	CVA (Credit Valuation Adjustment)	5
2.1	Modélisation des probabilités de défauts	6
2.2	Modélisation des obligations zéro-coupon	6
2.3	Modélisation de l'exposition moyenne	7
3	Problème de corrélation	8
3.1	Higham	9
3.2	Houduo Qi et Defung Sun	10
3.3	McNeil, Frey et Embrechts	11
3.4	Synthèse des trois méthodes	11
4	Adjoint Algorithmic Differentiation	13
4.1	Le mode Forward	14
4.2	Le mode Adjoint	15
4.3	Calcul de l'adjoint d'une factorisation de Cholesky	15
5	Résultats Numériques	17
5.1	Un Swap	17
5.2	Plusieurs swaps et un Call Européen	21
6	Conclusion	24
A	Probabilités de défaut	25
B	Problème de Corrélation	25
C	Pseudo-code	27

1 Introduction

Ce stage s'est déroulé dans l'équipe de la cellule Finance Quantitative au sein du département R&D de Quanteam. Cette équipe est en charge de la mise en place des projets/réflexions sur des sujets d'actualités de la place financière. Le stage a eu pour objectif la mise en place en C++ de la méthode AAD (Adjoint Algorithmic Differentiation) pour le calcul des sensibilités de la CVA (Credit Value Adjustment).

Ce stage m'a permis de mettre en pratique mes compétences en mathématiques financières et de confronter mes connaissances dans un cadre concret sur un sujet d'actualité. Ce stage a également été l'occasion de développer un projet dans un environnement de qualité et d'acquérir des compétences en informatique.

Le calcul des sensibilités de la CVA est à la fois un pré requis pour le Hedging, le pilotage du desk CVA Front office et la mise en œuvre de la revue fondamentale du trading book (FRTB). La méthode des différences finies (approche traditionnelle) nécessite de vérifier un certain nombre de contraintes (continuité des payoffs, temps/puissance de calcul dont la globalité des structures ne disposent pas, ...). La première étape du stage consista en une veille bibliographique sur le sujet afin de recenser, comprendre et se familiariser avec les méthodes/modèles de calcul des sensibilités de la CVA. La seconde consista à implémenter et étudier les méthodes/modèles adaptés à cette problématique. Une réflexion sur l'extension et/ou mise en œuvre pratique de ces cadres théoriques a été menée avant d'appliquer ces différentes techniques à des produits traités sur le marché.

La différentiation algorithmique (AD) est une technique de programmation pour le calcul de dérivés dans un programme informatique (Griewank, 2000, [11]). Suite à l'introduction du mode adjoint de cette méthode en finance [10] (AAD), plusieurs travaux ont témoigné de l'efficacité de cette méthode de calcul de dérivés dans un contexte Monte Carlo [4], [5], [6]. Récemment, [7] utilise l'AAD pour le calcul de sensibilités de la CVA.

Dans un premier temps, nous donnons quelques définitions de la CVA ainsi que la façon dont nous la calculons (hypothèses de simplification), comment nous modélisons les probabilités de défauts, les facteurs d'actualisation ainsi que le portefeuille. Dans le cas où nous devons modéliser plus de trois mouvement Brownien corrélés nous nous sommes confrontés au problème des matrices de variance-covariance non symétrique définie positive et donc ne possédant pas de décomposition de Cholesky. Pour faire face à ce problème nous avons fait l'étude de trois méthodes de calcul de matrice de corrélation la plus proche. Nous décrivons ensuite le mode backward de l'AD, la méthode AAD, sa différence avec le mode forward et la différence de l'AD avec la méthode des différences finie. Enfin nous présentons nos résultats numériques. En annexe nous fournissons un algorithme de calcul de la CVA et de l'application de l'AAD pour le calcul du delta de la CVA (dérivé par rapport à la courbe de taux).

2 CVA (Credit Valuation Adjustment)

Nous suivons [2] et dénotons par τ le temps de défaut de la contrepartie. Nous nous plaçons dans un espace de probabilité $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{Q})$ où \mathbb{Q} est la probabilité risque neutre et \mathcal{G}_t la filtration modélisant le flux d'information du marché. \mathcal{G}_t est composée de deux filtrations continues à droite où l'une est générée par les événements de défauts et l'autre représente les quantités observables sur le marché. La CVA (Credit Valuation Adjustment) est la valeur de marché du risque de crédit de la contrepartie. Elle est la différence entre la valeur risque neutre d'un portefeuille et sa vraie valeur en prenant en compte le risque de contrepartie. Dans le cas où seule la qualité de crédit de la contrepartie est prise en compte, on parle de CVA unilatérale. Soit $\Pi(t)$ le processus modélisant le prix risque neutre d'un portefeuille de transactions $V(t)$ avec une seule contrepartie à la date t , on a $\Pi(t) = \mathbb{E}^{\mathbb{Q}}[V(t)|\mathcal{G}_t]$. Suivant [7], on modélise la CVA unilatérale à $t = 0$ par :

$$CVA = (1 - R) \mathbb{E}^{\mathbb{Q}} \left[1_{\{0 < \tau \leq T_{CVA}\}} D(0, \tau) (\Pi(\tau))^+ \right] \quad (1)$$

$0 \leq R \leq 1$ représente le taux de recouvrement, $D(0, t) = \exp \left(- \int_0^t r_s ds \right)$ le facteur d'actualisation et r_t le taux court. On rappelle que $x \mapsto (x)^+ = \max(x, 0)$.

Soit $(0 = t_0, t_1, \dots, t_{N_{CVA}} = T_{CVA})$ une discrétisation du temps $[0, T_{CVA}]$, on a

$$CVA = (1 - R) \sum_{i=1}^{N_{CVA}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{t_{i-1} < \tau \leq t_i\}} D(0, \tau) (\Pi(\tau))^+ \right].$$

Afin de simplifier le calcul numérique de (1) nous émettons quelques hypothèses : Premièrement, supposons que les défauts ne peuvent survenir qu'en date t_i , on a :

$$CVA \approx (1 - R) \sum_{i=1}^{N_{CVA}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{t_{i-1} < \tau \leq t_i\}} D(0, t_i) (\Pi(t_i))^+ \right].$$

Supposons de plus que l'évolution du marché est indépendante des défauts¹, on obtient :

$$CVA \approx (1 - R) \sum_{i=1}^{N_{CVA}} \mathbb{Q}(t_{i-1} < \tau \leq t_i) \mathbb{E}^{\mathbb{Q}} \left[D(0, t_i) (\Pi(t_i))^+ \right].$$

Pour la dernière espérance nous choisissons de passer de la mesure risque neutre \mathbb{Q} à la mesure t -forward \mathbb{Q}^t afin de pouvoir utiliser les prix zéro coupons d'aujourd'hui $P(0, t)$ que nous supposons connaître. Pour cela, nous posons $\frac{d\mathbb{Q}^t}{d\mathbb{Q}} = \frac{D(0, t)}{P(0, t)}$,

$$CVA \approx (1 - R) \sum_{i=1}^{N_{CVA}} \mathbb{Q}(t_{i-1} < \tau \leq t_i) P(0, t_i) EE(t_i), \quad (2)$$

où $P(0, t_i) = \mathbb{E}^{\mathbb{Q}}[D(0, t_i)]$ est le prix d'une obligation zéro-coupon et $EE(t_i) = \mathbb{E}^{\mathbb{Q}^{t_i}}[(\Pi(t_i))^+]$ est l'exposition moyenne.

Nous rappelons que $\mathbb{Q}(t_{i-1} < \tau \leq t_i) = \mathbb{Q}(\tau \leq t_i) - \mathbb{Q}(\tau \leq t_{i-1})$.

Dans la suite nous expliquons comment nous modélisons les trois termes apparaissant dans (2), à savoir, les probabilités de défauts, le prix d'une obligation zéro-coupon et l'exposition moyenne pour chaque date t_i , $i \in \{1, \dots, N_{CVA}\}$.

1. définir wrong-way-risk

2.1 Modélisation des probabilités de défauts

On suppose que le défaut est décrit par un processus de Poisson d'intensité $\lambda : t \mapsto \lambda(t)$ constante par morceaux :

$$\mathbb{Q}(\tau \leq t) = 1 - \exp\left(-\int_0^t \lambda(s)ds\right), \quad (3)$$

où les intensités sont données par :

$$\forall t \in [0, T_{CVA}], \quad 0 = \tilde{t}_0 < \tilde{t}_1 < \dots < \tilde{t}_{N_{PD}}, \quad \lambda(t) = \sum_{j=1}^{N_{PD}} \lambda_j 1_{\tilde{t}_{j-1} < t \leq \tilde{t}_j}.$$

Soient $(PD(0, \tilde{t}_1), PD(0, \tilde{t}_2), \dots, PD(0, \tilde{t}_{N_{PD}}))$, les probabilités de défauts historiques de la contrepartie pour N_{PD} ténors que nous possédons. La première étape du calcul des probabilités de défauts $\mathbb{Q}(\tau \leq t_i)$ pour chaque date t_i , $i \in \{1, \dots, N_{CVA}\}$ consiste à construire les intensités de défauts λ_i et d'en déduire itérativement les intensités de défauts λ_i pour tout $i \in \{1, \dots, N_{PD}\}$:

$$\lambda_i = \frac{-\log(1 - PD(0, \tilde{t}_i)) - \sum_{j=1}^{i-1} (\tilde{t}_j - \tilde{t}_{j-1})\lambda_j}{(\tilde{t}_i - \tilde{t}_{i-1})}.$$

Soit $I : t \mapsto I(t)$ l'indice inférieur du ténor des probabilités de défauts historiques que nous possédons le plus proche de t . Par exemple si t est compris entre \tilde{t}_{j-1} et \tilde{t}_j , alors $I(t) = j - 1$. Si $t \geq \tilde{t}_{N_{PD}}$, alors $I(t) = N_{PD}$. À partir des intensités de défaut, nous en déduisons le calcul des probabilités de défauts pour chaque date t_i , $i \in \{1, \dots, N_{CVA}\}$:

$$\mathbb{Q}(\tau \leq t_i) \approx 1 - \exp\left(-\sum_{j=1}^{I(t_i)} (\tilde{t}_j - \tilde{t}_{j-1})\lambda_j - (t_i - I(t_i))\lambda_{I(t_i)+1}\right). \quad (4)$$

Si $I(t) = N_{PD}$ (i.e. $t_i \geq \tilde{t}_{N_{PD}}$) il faut remplacer $\lambda_{I(t_i)+1}$ par $\lambda_{I(t_i)}$ dans la dernière ligne. Cela revient à supposer que $\lambda(t) = \lambda_{N_{PD}}$ pour tout $t \geq \tilde{t}_{N_{PD}}$. Une extrapolation serait mieux adapter. Une preuve de ses résultats est donnée en annexe A.

2.2 Modélisation des obligations zéro-coupon

Cette partie est tirée de [3]. Pour rappel, le prix d'une obligation zéro-coupon est donné par :

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds} | \mathcal{G}_t\right].$$

Pour des questions de traçabilité nous choisissons de modéliser le taux court r par le modèle G2++.

Nous supposons détenir une courbe term-structure des taux déduit du marché.

Dans ce modèle à deux facteurs Gaussien, la dynamique du taux court instantané sous la probabilité risque neutre est donnée par :

$$r(t) = x(t) + y(t) + \phi(t), \quad r(0) = r_0,$$

où les processus $\{x(t) : t \geq 0\}$ et $\{y(t) : t \geq 0\}$ satisfont :

$$\begin{aligned} dx(t) &= -ax(t)dt + \sigma dW_1(t), \quad x(0) = 0, \\ dy(t) &= -by(t)dt + \eta \left(\rho dW_1(t) + \sqrt{1 - \rho^2} dW_2(t) \right), \quad y(0) = 0. \end{aligned}$$

$W_1(t)$ et $W_2(t)$ sont deux \mathcal{G}_t -mouvements browniens sous \mathbb{Q} . Les paramètres a , b , σ et η sont des constantes positives et $-1 < \rho < 1$. Une simple intégration de ces EDS donne :

$$\begin{aligned} x(t) &= e^{-a(t-s)}x(s) + \sigma \int_s^t e^{-a(t-u)}dW_1(u), \\ y(t) &= e^{-a(t-s)}y(s) + \eta \int_s^t e^{-a(t-u)}d\left(\rho W_1(u) + \sqrt{1-\rho^2}W_2(u)\right). \end{aligned}$$

En utilisant ces solutions on peut prouver que sachant \mathcal{G}_t , la somme des processus intégrés $x(\cdot)$ and $y(\cdot)$ sont normalement distribués avec moyenne $M(\cdot, T)$ et variance $V(\cdot, T)$:

$$\begin{aligned} \mathcal{I}(t, T) &:= \int_t^T (x(u) + y(u)) du \sim \mathcal{N}(M(t, T), V(t, T)), \\ M(t, T) &= \frac{x(t)}{a} (1 - e^{-a(T-t)}) + \frac{y(t)}{b} (1 - e^{-b(T-t)}), \\ V(t, T) &= \frac{\sigma^2}{a^2} \left((T-t) + \frac{4e^{-a(T-t)} - e^{-2a(T-t)} - 3}{2a} \right) \\ &\quad + \frac{\eta^2}{b^2} \left((T-t) + \frac{4e^{-b(T-t)} - e^{-2b(T-t)} - 3}{2b} \right) \\ &\quad + \frac{2\rho\sigma\eta}{ab} \left(\frac{e^{-a(T-t)} - 1}{a} + \frac{e^{-b(T-t)} - 1}{b} + \frac{e^{-(a+b)(T-t)} - 1}{a+b} \right). \end{aligned}$$

Connaissant la loi du taux court intégré il nous est possible de calculer le prix d'une obligation zéro-coupon de maturité T à la date t :

$$\begin{aligned} P(t, T) &= \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} | \mathcal{G}_t \right] \\ &= e^{-\int_t^T \phi(s) ds} \mathbb{E}^{\mathbb{Q}} \left[e^{-\mathcal{I}(t, T)} | \mathcal{G}_t \right] \\ &= \exp \left(-\int_t^T \phi(s) ds - M(t, T) + \frac{1}{2} V(t, T) \right). \end{aligned}$$

Nous supposons détenir une courbe des taux donnée par la fonction $T \mapsto P^M(0, T)$, le modèle G2++ concorde avec le marché si et seulement si $P(0, T) = P^M(0, T)$, i.e.

$$e^{-\int_t^T \phi(s) ds} = \frac{P^M(0, T)}{P^M(0, t)} \exp \left(-\frac{1}{2} (V(0, T) - V(0, t)) \right).$$

C'est la propriété exogène du modèle G2++.

Le prix d'une obligation zéro-coupon en fonction de la courbe des taux est donc :

$$P(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \exp \left(-M(t, T) - \frac{1}{2} [V(0, T) - V(0, t) - V(t, T)] \right).$$

2.3 Modélisation de l'exposition moyenne

Nous supposons que le portefeuille est constitué de Call Européen et de swap de taux. Pour des questions de simplicité nous nous modélisons le taux future avec la même courbe que pour les facteurs d'actualisation.

Introduisons quelques notations :

- T_{swap} : Maturité du swap.
- K : Taux fixe.
- N : Nominal du swap.

- N_{var}^i : Nombre de flux restant entre la date t_i et la maturité du swap pour la jambe variable.
- N_{fix}^i : Nombre de flux restant entre la date t_i et la maturité du swap pour la jambe fixe.
- $T_k^{i,var}$: date du k -ième flux de la jambe variable après la date t_i .
- $T_k^{i,fix}$: date du k -ième flux de la jambe fixe après la date t_i .
- $\tau(t, T)$: Base entre les dates t , et T .

Soit $F(t, T, S)$ le taux d'intérêt futur simplement composé à la date T évalué à la date t et de maturité S , on a :

$$F(t, T, S) = \frac{1}{\tau(T, S)} \left(\frac{P(t, T)}{P(t, S)} - 1 \right). \quad (5)$$

Soit $V_j(t_i)$ la valeur du j -ème produit dans le portefeuille à la date t_i : $V(t_i) = \sum_{j=1}^N V_j(t_i)$.

Le processus décrivant le prix du swap à la date t_i pour un payeur est donné par :

$$V_j(t_i) = N \sum_{k=1}^{N_{var}^i} \tau(T_{k-1}^{i,var}, T_k^{i,var}) P(t_i, T_k^{i,var}) F(t_i, T_{k-1}^{i,var}, T_k^{i,var}) - NK \sum_{k=1}^{N_{fix}^i} \tau(T_{k-1}^{i,fix}, T_k^{i,fix}) P(t_i, T_k^{i,fix}).$$

En remplaçant le taux variable par son expression (5), on retrouve une somme télescopique qui se simplifie :

$$V(t_i) = N \left(1 - P(t_i, T_{N_{var}^i}^{i,var}) \right) - NK \sum_{k=1}^{N_{fix}^i} \tau(T_{k-1}^{i,fix}, T_k^{i,fix}) P(t_i, T_k^{i,fix}) \quad (6)$$

Pour calculer les expositions moyennes à chaque date t_i , $i \in \{1, \dots, N_{CVA}\}$, $EE(t_i) = \mathbb{E}^{\mathbb{Q}^{t_i}} [(V(t_i))^+]$ on utilise un estimateur de Monte Carlo \hat{C}_i .

Soit $V(t_i, \omega_1), \dots, V(t_i, \omega_{N_{MC}})$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées selon la même loi que $V(t_i)$. Par application de la loi forte des grands nombres, on a :

$$\hat{C}_i = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} (V(t_i, \omega_j))^+ \xrightarrow[N \rightarrow +\infty]{p.s.} \mathbb{E}^{\mathbb{Q}} [(V(t_i))^+].$$

Soit σ_i l'écart-type de $V(t_i)$, par application du théorème central limite, on a :

$$\frac{\hat{C}_i - EE(t_i)}{\frac{\sigma_i}{\sqrt{N_{MC}}}} \xrightarrow[N \rightarrow +\infty]{loi} \mathcal{N}(0, 1),$$

en utilisant les tables de la loi normale, pour N grand, avec une probabilité de 0.95, on a :

$$\hat{C}_i \in \left[EE(t_i) - \frac{1.96\sigma_i}{\sqrt{N_{MC}}}, EE(t_i) + \frac{1.96\sigma_i}{\sqrt{N_{MC}}} \right]$$

3 Problème de corrélation

Dans notre modélisation nous supposons détenir une matrice de corrélation des mouvements Browniens. Pour N_{MB} mouvements browniens requis dans le calcul de la CVA, nous disposons d'une matrice symétrique de taille $N_{MB} \times N_{MB}$ telle que :

$$\Gamma = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \cdots & \rho_{1,N_{MB}} \\ \rho_{1,2} & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \rho_{N_{MB}-1,N_{MB}} \\ \rho_{1,N_{MB}} & \cdots & \cdots & \rho_{N_{MB}-1,N_{MB}} & 1 \end{pmatrix}.$$

i.e. $(\Gamma)_{i,j} = \rho_{i,j} = \langle dW_i, dW_j \rangle$, avec $\rho_{i,j} = \rho_{j,i}$ et $\rho_{i,i} = 1$ pour tous i et j dans $\{1, \dots, N_{MB}\}$. À partir de cette matrice de corrélation Γ , une des premières étapes de la modélisation de la CVA est de construire les mouvements browniens corrélés pour chaque diffusion à chaque pas de temps de discrétisation de $[0, T_{CVA}]$. Par exemple pour N_{MB} diffusions requises, pour une discrétisation de $[0, T_{CVA}]$ comprenant N_{CVA} points distincts et un nombre de N_{MC} simulations, nous devons simuler à chaque pas de temps $t_i, i \in \{1, \dots, N_{CVA}\}$, $N_{CVA} \times N_{MC}$ variables aléatoires suivant la loi normale. Supposons qu'il existe une matrice triangulaire inférieure L de taille $N_{MB} \times N_{MB}$ telle que $\Gamma = L L^T$, et X une matrice de taille $N_{MB} \times N_{MC}$ dont les entrées sont des variables aléatoires indépendantes de loi normale centrée et réduite. X contient la modélisation des mouvements browniens : $\sqrt{t}(X)_{i,j} = W_i^j(t), \forall i, j \in \{1, \dots, N_{MB}\} \times \{1, \dots, N_{MC}\}$. Alors $Y = L X$ est une matrice de taille $N_{MB} \times N_{MC}$ contenant les mouvements browniens corrélés à la date t . La première étape de la modélisation de la CVA est donc de construire pour chaque date $t_i, i \in \{1, \dots, N_{CVA}\}$, $N_{MB} \times N_{MC}$ une matrice $Y(t_i)$. Pour ce faire, il faut que Γ admette une décomposition de Cholesky, et donc que Γ soit symétrique définie positive. Ce qui n'est pas toujours le cas.

Dans ce qui suit nous introduisons la méthode que nous utilisons qui permet de trouver, dans le cas où Γ n'est pas définie positive, une matrice qui soit proche² de Γ et qui soit définie positive. Calculer la matrice de corrélation la plus proche d'une matrice symétrique est un problème de finance qui a été premierement étudié par Higham dans (Higham, 2002, [15]). Cependant sa méthode présentant une convergence lente et ne rendant pas la matrice résultante définie positive (les valeur propres sont au mieux nulles) elle nous ne nous permet pas de l'exploiter pour effectuer une décomposition de Cholesky. Pour y pallier nous utilisons la méthode de Houduo Qi et Defeng Sun (2006, [19]) qui converge de façon quadratique et qui permet de rendre la matrice résultante parfaitement définie positive. Nous comparons cette méthode avec celle de Higham (citée plus haut) et celle de McNeil, Frey et Embrechts (2015, [17]). Avant de les comparer dans des cas pratiques, nous décrivons les méthodes.

3.1 Higham

Soit \mathcal{S}^n l'espace convexe des matrices symétriques de taille $n \times n$, \mathcal{S}_+^n le cône des matrices semi-définie positive dans \mathcal{S}^n , $U \in \mathcal{S}^n$ l'espace convexe des matrices symétriques de taille $n \times n$ et de diagonale unité et $\Gamma \in \mathcal{S}^n$ une matrice symétrique. Le problème que résout la méthode de Higham [15] (trouver la matrice de corrélation la plus proche de Γ) est le suivant :

$$\min_{X \in \mathcal{S}_+^n \cap U} \|\Gamma - X\|, \quad (7)$$

où $\|\cdot\|$ désigne la norme de Frobenius et $\langle \cdot, \cdot \rangle$ le produit scalaire induit par cette norme (i.e. $\forall X, Y \in \mathcal{S}^n, \|X\| = (\text{trace}(X^T X))^{\frac{1}{2}}, \langle X, Y \rangle = \text{trace}(X^T Y)$). \mathcal{S}_+^n et U étant deux espaces fermés et convexes, leur intersection l'est aussi. Ceci nous permet d'utiliser les résultats d'existence et d'unicité rappelés dans l'Annexe B. Le point de départ de la méthode de Higham est la condition nécessaire et suffisante (14) qui stipule que la solution du problème de minimisation (7) que l'on note $\tilde{\Gamma}$ est la projection orthogonale de Γ sur $\mathcal{S}_+^n \cap U$. Pour trouver la matrice de corrélation la plus proche de Γ qui soit à l'intersection de \mathcal{S}_+^n et U , l'idée proposée par Higham est donc de projeter itérativement une matrice initiale dans \mathcal{S}_+^n puis dans U . En reprenant ses notations, on note $P_U(A)$ la projection orthogonale de A dans U et $P_{\mathcal{S}_+^n}(B)$ la projection orthogonale de B dans \mathcal{S}_+^n . On a :

2. dans le sens de distance induite par une norme.

Théorèmes Soit A une matrice de taille $n \times n$,

$$P_U(A) = \left\{ \tilde{A} \in \mathcal{S}^n \text{ telle que } (\tilde{A})_{i,j} = \begin{cases} (A)_{i,j}, & i \neq j \\ 1, & i = j \end{cases}, i, j \in \{1, \dots, n\} \right\}.$$

Soit B une matrice de taille $n \times n$ et de décomposition spectrale $B = QDQ^T$, où D est une matrice diagonale contenant les valeurs propres de B : $(\lambda_i)_{i \in \{1, \dots, n\}}$ et Q une matrice orthogonale contenant les vecteurs propres associés aux valeurs propres de B ,

$$P_{\mathcal{S}_+^n}(B) = \{Q\tilde{D}Q^T \text{ où } \tilde{D} \text{ est une matrice diagonale telle que } (\tilde{D})_{i,i} = \max(\lambda_i, 0), i \in \{1, \dots, n\}\}. \quad (8)$$

(Preuves dans (Higham, 2002, [15])).

Cependant, projeter itérativement une matrice initiale dans \mathcal{S}_+^n et U n'implique pas la convergence. Dans (Boyle et Dykstra, 1985, [1]) et (Han, 1988, [13]) il est montré qu'en utilisant un pas de correction avant chaque projection on obtient une convergence linéaire de la méthode.

3.2 Houduo Qi et Defung Sun

Dans [19], Qi et Sun étudient une méthode de type Newton pour le problème de matrice de corrélation la plus proche. Ils prouvent la convergence quadratique de leur méthode (contre une convergence linéaire pour celle de Higham) et adapte leur méthode pour trouver la matrice de corrélation la plus proche de Γ qui soit de plus, définie positive. Condition nécessaire pour la décomposition de Cholesky de Γ . En effet, dans le cas où Γ contient n valeurs propres négatives, la méthode d'Higham trouve une matrice proche de Γ qui n'est pas inversible car la matrice contiendra au moins n valeurs propres nulles. Qi et Sun adapte leur méthode en ajoutant une condition à la matrice résultante afin d'en avoir l'inversibilité. Soit $\Gamma \in \mathcal{S}^n$ et $\alpha \in [0, 1]$, le problème à résoudre s'écrit :

$$\min_{\substack{X \in \mathcal{S}_+^n \cap U \\ X \succeq \alpha I}} \frac{1}{2} \|\Gamma - X\|^2 \quad (9)$$

Grâce au théorème 2.2 de (Micchelli et Utreras, 1988, [18]) et à la CNS (14) on sait que l'unique solution $\tilde{\Gamma}$ du problème (9) est de la forme :

$$\begin{aligned} \tilde{\Gamma} &= F(y^*), \\ \text{où } F(y^*) &= P_{\mathcal{S}_+^n}[\Gamma - \alpha I + \text{Diag}(y^*)], \end{aligned} \quad (10)$$

avec $P_{\mathcal{S}_+^n}(\cdot)$ la projection dans \mathcal{S}_+^n (8), $\text{Diag}(y^*)$ la matrice diagonale telle que

$$(\text{Diag}(y^*))_{i,j} = \begin{cases} 0, & i \neq j \\ (y^*)_{i,i}, & i = j \end{cases}, \quad \forall i, j \in \{1, \dots, n\},$$

et y^* l'unique solution du système :

$$(F(y))_{i,i} = (1 - \alpha), \quad \forall i \in \{1, \dots, n\}. \quad (11)$$

La méthode de Qi et Sun consiste à utiliser une méthode de Newton pour résoudre le système (11) et d'en déduire l'unique solution $\tilde{\Gamma}$ (10) du problème de matrice de corrélation la plus proche (9).

Le $(k + 1)$ -ième itéré de la méthode de Newton est donné par :

$$y^{k+1} = y^k - p^k,$$

où p^k est solution de

$$V_k p^k = \text{diag} \left(F(y^k) - (1 - \alpha)I \right), \quad V_k \in \partial \left((F(y^k))_{1,1}, \dots, (F(y^k))_{n,n} \right), \quad (12)$$

Avec $\partial\phi(x)$ le Jacobien généralisé au sens de (Clarke, 1983, [8]). ($\text{diag}(A)$ est le vecteur dont les composantes sont les éléments de la diagonale de A : $((\text{diag}(A))_i = (A)_{i,i}, \forall i \in \{1, \dots, n\})$. Pour le calcul du système (12), la méthode du Gradient Conjugué (Hestenes et Stiefel, 1952, [14]) est utilisée.

3.3 McNeil, Frey et Embrechts

La méthode de McNeil, Frey et Embrechts (2015, [17], page 231) est un peu plus simple que les méthodes précédentes mais permet de toujours obtenir une matrice de corrélation semi-définie positive en une seule itération. Soit $\Gamma \in \mathcal{S}^n$ une matrice de corrélation de taille $n \times n$ (symétrique, d'éléments diagonaux égaux à 1 et d'éléments extra-diagonaux compris entre -1 et 1) qui n'est pas semi-définie positive, l'algorithme proposé pour calculer une matrice $\tilde{\Gamma}$ qui soit proche de Γ consiste à :

1. Calculer la décomposition spectrale de Γ :

$$\Gamma = QDQ^T,$$

où D est une matrice diagonale contenant les valeurs propres de Γ : $(\lambda_i)_{i \in \{1, \dots, n\}}$ et Q une matrice orthogonale contenant les vecteurs propres associés aux valeurs propres de Γ .

2. Remplacer les valeurs propres négatives de Γ par un réel positif fixé $\epsilon > 0$:

$$\text{Calculer } (\tilde{D})_{i,i} = \max(D_{i,i}, \epsilon), \quad \forall i \in \{1, \dots, n\}.$$

3. Calculer $R = Q\tilde{D}Q^T$.

4. Renvoyez la matrice de corrélation $\tilde{\Gamma}$:

$$\Delta(R) = \text{diag} \left(\sqrt{(R)_{1,1}}, \dots, \sqrt{(R)_{n,n}} \right),$$

$$\tilde{\Gamma} = (\Delta(R))^{-1} R (\Delta(R))^{-1}$$

3.4 Synthèse des trois méthodes

Nous comparons les trois méthodes sur des cas pratiques, un premier test est effectué sur une matrice de corrélation de taille 162×162 et un second test sur une matrice de taille 500×500 et de diagonale unité créée à l'aide d'un générateur aléatoire. Dans le premier cas, la matrice contient 43/162 valeurs propres négatives ou nulles, tandis que dans le deuxième cas la matrice contient 238/500 valeurs propres négatives. Nous souhaitons évaluer la vitesse de calcul des trois méthodes, le nombre d'itérations avant la convergence, le nombre de valeurs propres strictement positives de la matrice résultante ainsi que la distance qui la sépare de la matrice d'entrée. Dans le tableau 1 (resp. tableau 2) on retrouve les résultats du premier test (resp. deuxième test) ainsi que le spectre des quatre matrices (matrice initiale Γ , matrices résultantes des trois méthodes) sur la figure 1 (resp. figure 2). Les tests sont réalisés sur un ordinateur portable Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz. Dans tous les tests réalisés, nous considérons qu'une valeur est numériquement égale à zéro lorsqu'elle est de l'ordre de 10^{-10} . Pour la méthode de Qi et Sun, nous avons fixé $\alpha = 10^{-4}$ et $\epsilon = 10^{-6}$ pour la méthode de MFE (McNeil, Frey et Embrechts).

	Higham	Qi et Sun	MFE
Temps	0.060748 s	0.098094 s	0.059387 s
Nb d'itérations	1	1	1
Valeurs propres strict. positives	119/162	162/162	162/162
Distance de Frobenius $\ \tilde{\Gamma} - \Gamma\ $	3.104e-09	0.007	114.6

TABLE 1 – Résultat 1er test - Matrice de corrélation 162×162

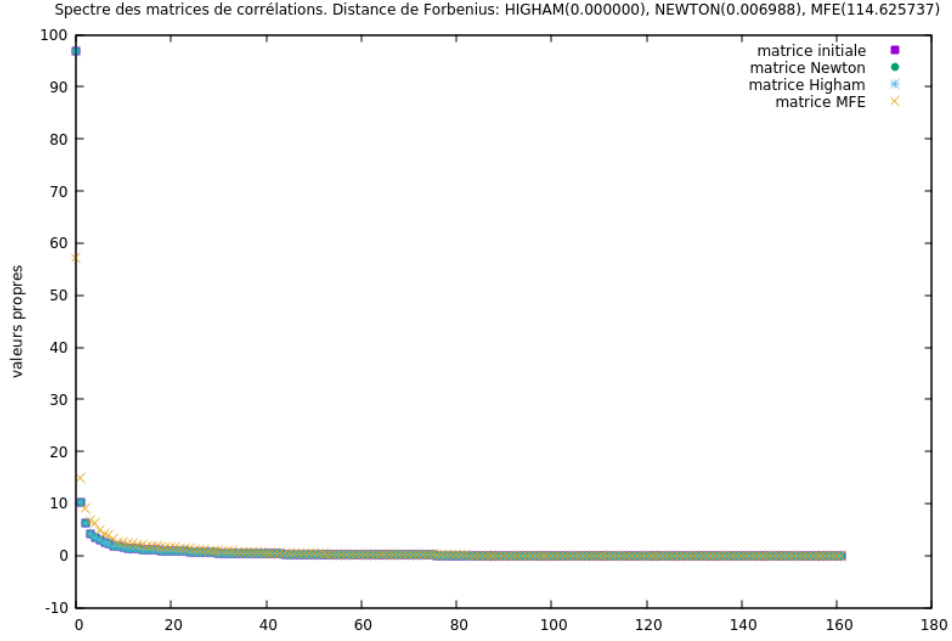


FIGURE 1 – Spectre de la matrice initiale Γ , et des matrices résultantes des trois méthodes pour le 1er test.

	Higham	Qi et Sun	MFE
Temps	184.7 s	9.4 s	1.97 s
Nb d'itérations	93	4	1
Valeurs propres strict. positives	99/500	500/500	500/500
Distance de Frobenius $\ \tilde{\Gamma} - \Gamma\ $	256.7149	256.7151	289.63

TABLE 2 – Résultat 2ème test - Matrice de corrélation aléatoire 500×500

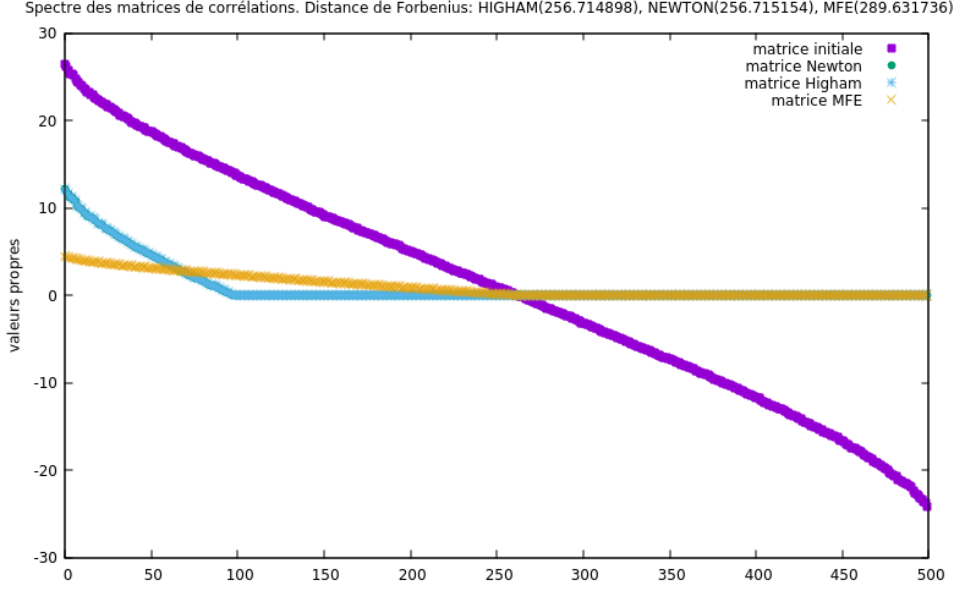


FIGURE 2 – Spectre de la matrice initiale Γ , et des matrices résultantes des trois méthodes pour le 2ème test.

La méthode de MFE est plus rapide car elle nécessite qu'une seule itération, en contrepartie, la matrice résultante est plus éloignée de la matrice d'entrée comparé aux deux autres méthodes. Sans aucune surprise, on remarque que la méthode de Qi et Sun converge plus vite et en moins d'itérations que la méthode de Higham (convergence quadratique contre convergence linéaire). Cependant, du fait qu'on requiert de la matrice résultante qu'elle soit définie positive (c'est la condition $X \succeq \alpha I$ dans (9), i.e. toutes les valeurs propres supérieures à α) par rapport à la méthode de Higham, la solution est plus éloignée de Γ que la solution de la méthode de Higham ne l'est. Soulignons qu'un temps de calcul conséquent parmi le temps total de chaque méthode est consacré aux décompositions spectrales. Ici nous utilisons une méthode de Jacobi pour calculer les valeurs propres et vecteurs propres de chaque matrice. D'un point de vue précision il semble plus judicieux d'utiliser la méthode de Qi et Sun. Cependant, l'algorithme de MFE étant plus simple, il en est d'autant plus pour calculer son adjoint dans le cadre de la méthode AAD. C'est ce que nous voyons dans la suite.

4 Adjoint Algorithmic Differentiation

L'AAD est une des deux méthodes d'Algorithmic Differentiation (AD) qui permet le calcul de dérivés avec un temps de calcul moindre et une meilleure précision comparé à la méthode des différences finies. Comme le montre Griewank et Walther dans [12], on peut quantifier le coût de calcul lorsqu'on applique l'AD pour calculer un prix (P) et ses dérivés (D) :

$$\frac{\text{Coût}(P + D)}{\text{Coût}(P)} \leq \omega_A, \quad \text{avec } \omega_1 \in [3, 4].$$

Tandis qu'en utilisant la méthode des différences finies, on a :

$$\frac{\text{Coût}(P + D)}{\text{Coût}(P)} \approx \text{nombre de dérivés} + 1.$$

Concernant l'erreur de convergence, pour calculer la dérivé d'un prix P en fonction d'un paramètre $\sigma \mapsto P(\sigma)$, avec la méthode des différences finies on a :

$$\frac{\partial P(\sigma)}{\partial \sigma} = \frac{P(\sigma + h) - P(\sigma)}{h} + \mathcal{O}(h).$$

L'erreur est égale à $\mathcal{O}(h) + \epsilon$, où ϵ désigne la précision machine, tandis qu'avec l'AD, l'erreur est seulement ϵ . (Pour les doubles en 64 bits, $\epsilon = 2.220446e - 16$).

Le mode adjoint de l'AD consiste à effectuer un premier balayage de l'algorithme (calculer la CVA à $t = 0$ ici) puis de calculer les adjoints des différentes étapes lors d'un second balayage. Le calcul des adjoints nécessite des valeurs calculées lors du premier balayage. Considérons un exemple. Soit $(S_t^{(i)}, 0 \leq t \leq T)$, $i \in \{1, \dots, N_S\}$ N_S actifs risqués solution d'une EDS sous probabilité risque neutre \mathbb{Q} . Nous discrétisons la grille de temps $[0, T]$ par N_T points équidistants tels que $t_k = k\Delta T$, $k \in \{1, \dots, N_T\}$, $\Delta T = \frac{T}{N_T}$. Et nous supposons disposer d'une fonction qui permet de passer d'un pas précédent à un pas suivant :

$$S_{t_{k+1}}^{(i)} = \Phi(S_{t_k}^{(i)}), \quad S_0^{(i)} \text{ donnés}, \quad \forall i \in \{1, \dots, N_S\}, \quad \forall k \in \{0, \dots, N_T - 1\}.$$

Φ peut être un schéma de discrétisation de type Euler-Maruyama, Milstein, ou une solution explicite. On note $(S_t, 0 \leq t \leq T)$ le vecteur composé des N_S actifs : $S_t = (S_t^{(1)}, \dots, S_t^{(N_S)})$, le but de cet exemple est de trouver la prime Π_0 d'une option de payoff $G(S_T)$, i.e. :

$$\Pi_0 = e^{-rT} \mathbb{E}^{\mathbb{Q}} [G(S_T)]$$

et de calculer les dérivés de Π_0 en fonction de $S_0^{(i)}$ pour tout i dans $\{1, \dots, N_S\}$ en utilisant la méthode des différences finies, de l'AD en Forward mode, et de l'AAD. Supposons la fonction G Lipschitzienne, il nous est ainsi possible de rentrer la dérivé dans l'espérance. Ainsi il nous suffit de construire un estimateur de l'espérance par Monte Carlo sur les dérivés :

$$\frac{\partial \Pi_0}{\partial S_0^{(i)}} = e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\frac{\partial G(S_T)}{\partial S_0^{(i)}} \right] \approx \frac{e^{-rT}}{N_{MC}} \sum_{j=1}^{N_{MC}} \frac{\partial G(S_{T,(j)})}{\partial S_0^{(i)}}$$

où $S_{T,(j)}, j \in \{1, \dots, N_{MC}\}$ désigne une simulation de S_T . La quantité qui nous intéresse est donc $\frac{\partial G(S_T)}{\partial S_0^{(i)}}$ (où nous avons volontairement retiré les indices j). Par différences finies il nous faut fixer $h > 0$ et approcher la dérivé par :

$$\frac{\partial G(S_T)}{\partial S_0^{(i)}} = \frac{1}{h} \left(G(S_T(S_0^{(i)} + h)) - G(S_T) \right) + \mathcal{O}(h).$$

4.1 Le mode Forward

Le mode Forward (ou Tangent) consiste à calculer les dérivées de chaque grandeur au fur et à mesure que le calcul avance.

$$\frac{\partial G(S_T)}{\partial S_0^{(i)}} = \frac{\partial G(S_T^{(1)}, \dots, S_T^{(N_S)})}{\partial S_0^{(i)}} = \sum_{j=1}^{N_S} \frac{\partial G(S_T)}{\partial S_T^{(j)}} \frac{\partial S_T^{(j)}}{\partial S_0^{(i)}}$$

Or

$$\frac{\partial S_T^{(j)}}{\partial S_0^{(i)}} = \frac{\partial S_{t_{N_T}}^{(j)}}{\partial S_0^{(i)}} = \frac{\partial \Phi(S_{t_{N_T-1}}^{(j)})}{\partial S_0^{(i)}} = \frac{\partial \Phi(S_{t_{N_T-1}}^{(j)})}{\partial S_{t_{N_T-1}}^{(j)}} \frac{\partial S_{t_{N_T-1}}^{(j)}}{\partial S_0^{(i)}} = \prod_{k=1}^{N_T} \frac{\partial \Phi(S_{t_{N_T-k}}^{(j)})}{\partial S_{t_{N_T-k}}^{(j)}} \frac{\partial S_0^{(j)}}{\partial S_0^{(i)}}$$

Si nous réécrivons l'égalité précédente sous forme matricielle, nous avons :

$$\begin{aligned}
\left(\frac{\partial G(S_T)}{\partial S_0}\right)^T &:= \begin{pmatrix} \frac{\partial G(S_T)}{\partial S_0^{(1)}} \\ \vdots \\ \frac{\partial G(S_T)}{\partial S_0^{(NS)}} \end{pmatrix}^T = \begin{pmatrix} \frac{\partial G(S_T)}{\partial S_T^{(1)}} \\ \vdots \\ \frac{\partial G(S_T)}{\partial S_T^{(NS)}} \end{pmatrix}^T \begin{pmatrix} \frac{\partial S_T^{(1)}}{\partial S_0^{(1)}} & \cdots & \frac{\partial S_T^{(1)}}{\partial S_0^{(NS)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial S_T^{(NS)}}{\partial S_0^{(1)}} & \cdots & \frac{\partial S_T^{(NS)}}{\partial S_0^{(NS)}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial G(S_T)}{\partial S_T^{(1)}} \\ \vdots \\ \frac{\partial G(S_T)}{\partial S_T^{(NS)}} \end{pmatrix}^T \begin{pmatrix} \frac{\partial \Phi_{N_T-1}^{(1)}}{\partial S_{t_{N_T-1}}^{(1)}} & & 0 \\ & \ddots & \\ 0 & & \frac{\partial \Phi_{N_T-1}^{(NS)}}{\partial S_{t_{N_T-1}}^{(NS)}} \end{pmatrix} \begin{pmatrix} \frac{\partial S_{t_{N_T-1}}^{(1)}}{\partial S_0^{(1)}} & \cdots & \frac{\partial S_{t_{N_T-1}}^{(1)}}{\partial S_0^{(NS)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial S_{t_{N_T-1}}^{(NS)}}{\partial S_0^{(1)}} & \cdots & \frac{\partial S_{t_{N_T-1}}^{(NS)}}{\partial S_0^{(NS)}} \end{pmatrix} \\
(\text{notation}) &:= \left(\frac{\partial G(S_T)}{\partial S_T}\right)^T (D_S \Phi_{N_T-1}) (D_{S_0} S_{t_{N_T-1}})
\end{aligned}$$

De ce calcul on peut facilement remarquer l'itération suivante : $(D_{S_0} S_{t_k}) = (D_S \Phi_{k-1}) (D_{S_0} S_{t_{k-1}})$, avec $(D_{S_0} S_0)$ la matrice identité car $\frac{\partial S_0^{(j)}}{\partial S_0^{(i)}} = \delta_{i,j}$. En utilisant cette remarque nous obtenons une formule explicite du calcul des dérivés :

$$\left(\frac{\partial G(S_T)}{\partial S_0}\right)^T = \left(\frac{\partial G(S_T)}{\partial S_T}\right)^T (D_S \Phi_{N_T-1}) \dots (D_S \Phi_1). \quad (13)$$

Dans le mode Forward de l'AD, au fur et à mesure que nous discrétisons $(S_{t_k}, 0 \leq t_k \leq T)$ pour calculer la prime de l'option, nous calculons la matrice $(D_S \Phi_k)$ correspondante (avec une technique de surcharge d'opérateur (voir DCO/C++³ par exemple) puis la multiplions par la matrice $(D_{S_0} S_{t_k})$ calculée au pas précédent pour obtenir $(D_{S_0} S_{t_{k+1}})$. Nous réitérons ce produit matrice matrice jusqu'à $k = N_T$ pour obtenir le vecteur de dérivés souhaités : $\left(\frac{\partial G(S_T)}{\partial S_0}\right)^T = \left(\frac{\partial G(S_T)}{\partial S_T}\right)^T (D_{S_0} S_T)$.

4.2 Le mode Adjoint

Pour ce mode, on utilise l'adjoint de l'équation matricielle (13) et on calcule les itérations dans le sens inverse : on calcule d'abord le produit vecteur matrice $a_1 = \left(\frac{\partial G(S_T)}{\partial S_T}\right)^T (D_S \Phi_{N_T-1})$, puis le produit vecteur matrice $a_2 = a_1 (D_S \Phi_{N_T-2})$ ainsi de suite pour arriver à :

$$\begin{aligned}
\left(\frac{\partial G(S_T)}{\partial S_0}\right)^T &= \left(\frac{\partial G(S_T)}{\partial S_T}\right)^T (D_S \Phi_{N_T-1}) \dots (D_S \Phi_0) \\
&= a_1 \dots a_{N_T}
\end{aligned}$$

L'inconvénient de ce mode par rapport au Forward mode est qu'il faut attendre d'avoir calculer $\left(\frac{\partial G(S_T)}{\partial S_0}\right)$, et l'avantage est que nous faisons que des produits vecteur matrice et non matrice matrice.

4.3 Calcul de l'adjoint d'une factorisation de Cholesky

Nous donnons un exemple de modification de code afin d'y ajouter les étapes du mode Adjoint de l'AD pour le calcul de sensibilité. Ici il s'agit de la factorisation de Cholesky mentionnée plus haut dans le cadre du calcul de la CVA ; étant donnée une matrice X de variable aléatoire

3. <http://www.nag.co.uk/content/downloads-dco-adw3231icl>

Gaussienne non corrélées, nous souhaitons calculer Y une matrice de variable aléatoire Gaussienne de matrice de corrélation égale à Γ alors $Y = LX$ où $\Gamma = LL^T$. Ci-dessous l'algorithme initial de calcul de Y :

Algorithme 1 Factorisation de Cholesky : $\Gamma = LL^T$ et calcul de $Y = LX$

```

1: Pour  $i$  de 1 à  $n$ 
2:   Pour  $j$  de 1 à  $i$ 
3:     Pour  $k$  de 1 à  $j - 1$ 
4:        $\Gamma_{i,j} = \Gamma_{i,j} - L_{i,k}L_{j,k}$ 
5:     Fin Pour
6:     Si  $j = i$ 
7:        $L_{i,i} = \sqrt{\Gamma_{i,i}}$ 
8:     Sinon
9:        $L_{i,j} = \frac{\Gamma_{i,j}}{L_{j,j}}$ 
10:    Fin Si
11:  Fin Pour
12: Fin Pour
13: Calcul de  $Y = LX$ 

```

Nous utilisons les notations de la littérature et notons par une barre au dessus d'une grandeur la dérivée de L par rapport à cette grandeur. Nous nous intéressons donc au calcul de $\bar{\Gamma}$ telle que $(\bar{\Gamma})_{i,j} = \sum_{l,m=1}^n \frac{\partial L_{l,m}}{\partial \Gamma_{i,j}} \bar{L}_{l,m}$. La méthode AAD consiste donc à effectuer un premier balayage de l'algorithme 1 puis un second balayage en partant de la fin pour le calcul des adjoints successifs :

Algorithme 2 Calcul de l'adjoint $\bar{\Gamma}$

```

1: Calcul de  $\bar{L} = \bar{Y}X^T$ 
2: Pour  $i$  de  $n$  à 1
3:   Pour  $j$  de  $i$  à 1
4:     Si  $j = i$ 
5:        $\bar{\Gamma}_{i,i} = \frac{1}{2} \frac{\bar{L}_{i,i}}{L_{i,i}}$ 
6:     Sinon
7:        $\bar{\Gamma}_{i,j} = \frac{\bar{L}_{i,j}}{L_{j,j}}$ 
8:        $\bar{L}_{j,j} = \bar{L}_{j,j} - \bar{L}_{i,j} \frac{L_{i,j}}{L_{j,j}}$ 
9:     Fin Si
10:    Pour  $k$  de 1 à  $j - 1$ 
11:       $\bar{L}_{i,k} = \bar{L}_{j,k} - \bar{\Gamma}_{i,j}L_{j,k}$ 
12:       $\bar{L}_{j,k} = \bar{L}_{j,k} - \bar{\Gamma}_{i,j}L_{i,k}$ 
13:    Fin Pour
14:  Fin Pour
15: Fin Pour

```

Application de l'AAD à la CVA

En annexe est présenté un pseudo code de l'algorithme de calcul de la CVA ainsi que son delta (sa dérivée par rapport aux valeurs de la courbe de taux donnée en entrée) pour un portefeuille constitué d'un swap de taux. On note

- N_{tenor} : Nombre de ténor de la courbe de taux en entrée.
- $P^M(0, tenor_k)$: Obligation zéro coupon à partir du marché pour le ténor $tenor_k$.
- $ZC^M(0, tenor_k)$: Taux zéro coupon à partir du marché pour le ténor $tenor_k$, $ZC^M(0, tenor_k) = \frac{-\log(P^M(0, tenor_k))}{\tau(0, tenor_k)}$.
- ω_j : Un événement de l'espace de probabilité $(\Omega, \mathcal{G}_t, \mathbb{Q})$.
- $P^j(t_i, t_i + tenor_k)$: Prix future d'un zéro coupon bond à la date t_i pour le ténor $tenor_k$ et selon l'événement ω_j .
- $ZC^j(t_i, t_i + tenor_k)$: Taux Zéro Coupon future à la date t_i pour le ténor $tenor_k$ et selon l'événement ω_j tel que $P^j(t_i, t_i + tenor_k) = e^{(-ZC^j(t_i, t_i + tenor_k) \times \tau(t_i, t_i + tenor_k))}$.
- $V(t_i, \omega_j)$: Prix future du swap à la date t_i selon le j -ième scénario.

5 Résultats Numériques

Nous présentons un premier résultat de calcul de sensibilité de la CVA par rapport à la courbe de taux avec un portefeuille constitué d'un seul swap. Nous présentons ensuite un second résultat avec un portefeuille constitué de plusieurs swap et d'un Call Européen. Premièrement nous déterminons de quel ordre doit être le pas utilisé dans la méthode des différences finie. Pour cela nous comparons la somme de dérivé de la CVA par rapport à la courbe des taux obtenue avec différence finie et la même dérivé obtenue avec AAD (i.e. $\sum_{i=1}^{N_{tenor}} \frac{\partial CVA}{\partial tenor_i}$) :

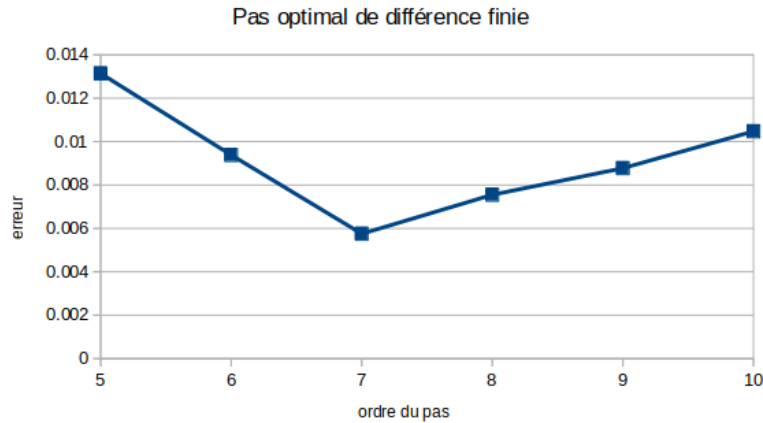


FIGURE 3 – Pas de différence finie optimal.

Nous utiliserons donc un pas égal à 10^{-7} pour la méthode de différence finie des deux tests suivant.

5.1 Un Swap

Nous supposons que le portefeuille est constitué d'un seul swap. Ci-dessous (Table 3) les paramètres utilisés pour le calcul de la CVA (maturité, nombre de pas de temps de discrétisation), les paramètres des facteurs de diffusions de la courbe de taux (coefficients de drift, de volatilité, corrélation), les paramètres du swap de taux (maturité, convention, fréquence, taux fixe), ainsi que la courbe de taux en entrée (Table 4).

T_{CVA}	15 (ans)
N_{CVA}	200
N_{MC}	1000
N_{tenor}	12
N_{PD}	30
Paramètres G2++	
a	0.32818
b	0.042382
σ	0.01418
η	0.012502
ρ	-0.99184
Paramètres Swap	
N	100 (€)
T_{swap}	10 (ans)
Fréquence Jambe Fixe	Annuelle
Fréquence Jambe Variable	Trimestrielle
K	0.012 (1.2%)

TABLE 3 – Tableau des paramètres

$tenor_k$ (en mois)	$ZC(0, tenor_k)$
1	-0.026751
3	0.002033
6	0.044294
9	0.047496
12	0.068156
24	0.133765
36	0.244568
60	0.565951
84	0.884501
120	1.262179
180	1.606225
360	1.756422

TABLE 4 – Courbe des taux en entrée

Sur la figure 4 on trace la diffusion des taux zéro-coupon pour plusieurs dates à ténor fixé ($tenor_7$). Sur la figure 6 et 5 on trace le Mark-to-Market ainsi que le profil d'exposition pour un payeur de swap. Enfin sur la figure 7 et le tableau 6 la dérivé de la CVA par rapport aux ténors de la courbe de taux. On remarque que la CVA est très sensible par rapport au ténor 10, ($\frac{\partial CVA}{\partial tenor_{10}} = 28.228231002$). Sachant que le portefeuille de la CVA est constitué d'un swap sur 10 ans et que le ténor 10 correspond à 10 ans, on pouvait s'y attendre. Concernant le temps de calcul, tableau 5, comme on pouvait s'y attendre le calcul des 12 dérivés via la méthode AAD (8 secondes) surpasse les Différences finies (22 secondes).

Méthode	Temps
Différence Finie	22 s
AAD	8 s

TABLE 5 – Temps de calcul de la CVA et des 12 dérivés - Différence Finie contre AAD.

CVA (IC - borne sup)	0.239397
CVA	0.232078
CVA (IC - borne inf)	0.22476
$\frac{\partial CVA}{\partial tenor_1}$	-0.00082572367
$\frac{\partial CVA}{\partial tenor_2}$	-0.00321228774
$\frac{\partial CVA}{\partial tenor_3}$	-0.0085042657
$\frac{\partial CVA}{\partial tenor_4}$	-0.0165295996
$\frac{\partial CVA}{\partial tenor_5}$	-0.1160029421
$\frac{\partial CVA}{\partial tenor_6}$	-0.4479407470
$\frac{\partial CVA}{\partial tenor_7}$	-2.025108287
$\frac{\partial CVA}{\partial tenor_8}$	-5.794385227
$\frac{\partial CVA}{\partial tenor_9}$	-7.122165199
$\frac{\partial CVA}{\partial tenor_{10}}$	28.228231002
$\frac{\partial CVA}{\partial tenor_{11}}$	1.953960903
$\frac{\partial CVA}{\partial tenor_{12}}$	0
$\Delta = \sum_{i=1}^{N_{tenor}} \frac{\partial CVA}{\partial tenor_i}$	14.647517627

TABLE 6 – Résultats.

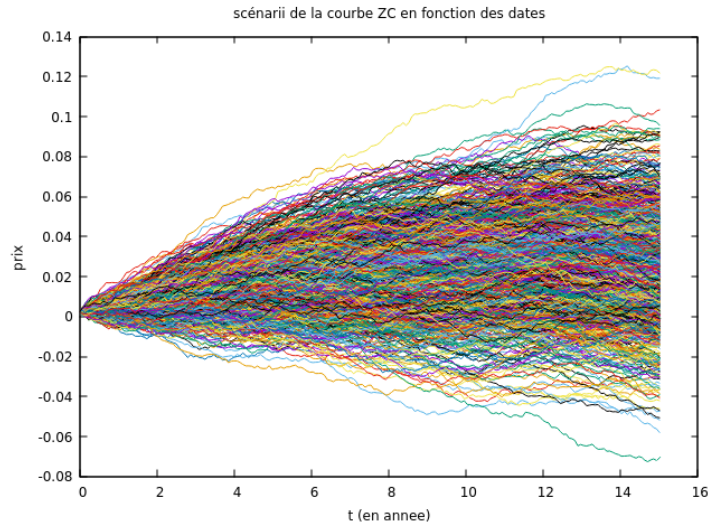


FIGURE 4 – taux ZC diffusés : $ZC^j(t, t + tenor_7)$ en fonction de t pour plusieurs scénarios ω_j .

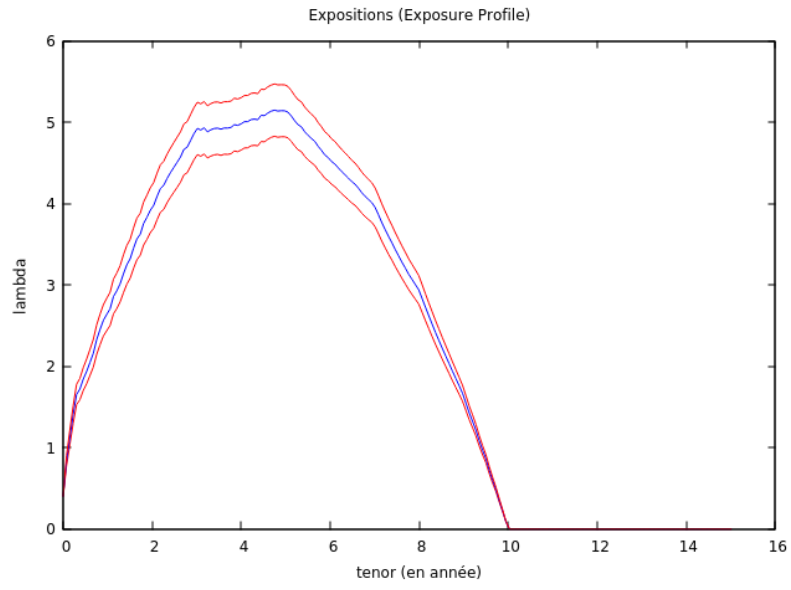


FIGURE 5 – Profil d'expositions : $EE(t)$ en fonction de t et son intervalle de confiance.

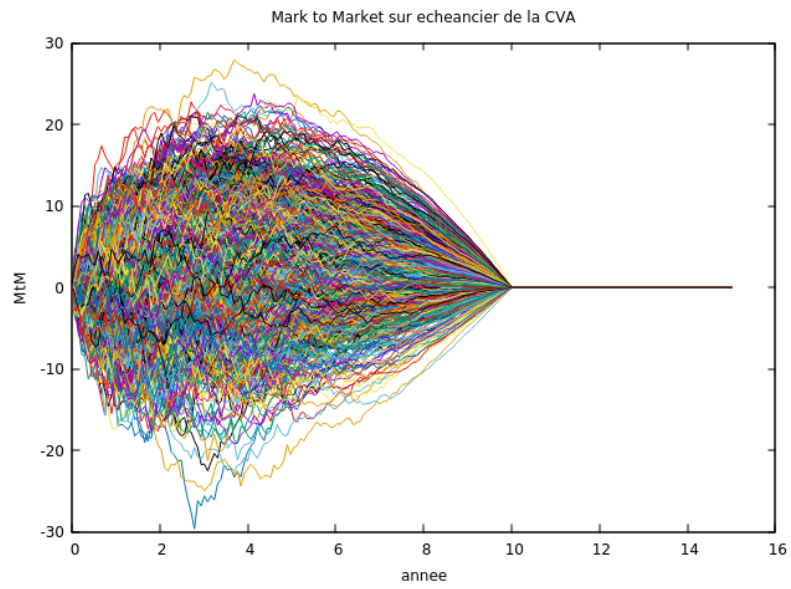


FIGURE 6 – Mark to Market : $V(t, \omega_j)$ en fonction de t et pour plusieurs scénarios ω_j .

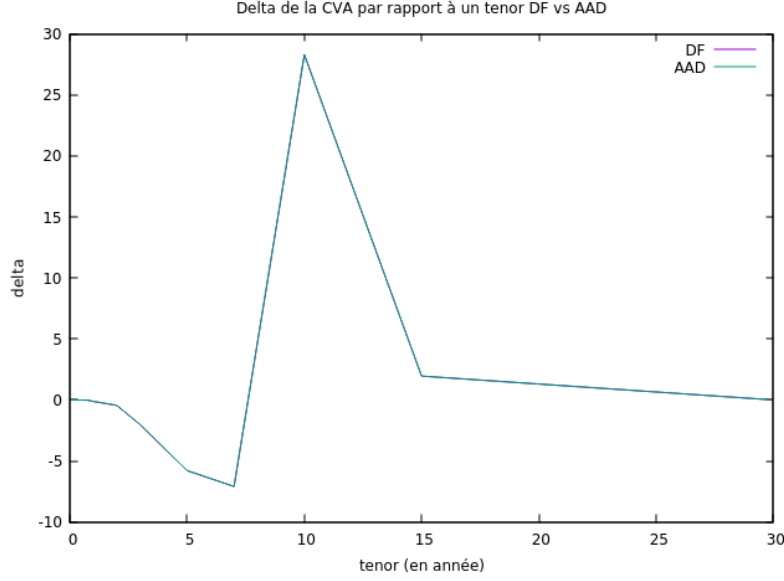


FIGURE 7 – Delta : $\frac{\partial CVA}{\partial \text{tenor}}$ en fonction des *tenors* calculés par Différences Finie et via l'AAD.

La dérivé de la CVA par rapport au taux à 10 ans est important car le swap mature à cette date. Le fait que nous n'observons pas une grande différence dans le calcul des sensibilités par la méthode des différences finie et par l'AAD est d'une part parce que nous avons choisi un pas de différence finie optimal et d'autre part grâce à la continuité de l'exposition moyenne.

5.2 Plusieurs swaps et un Call Européen

Dans ce second test nous modélisons un portefeuille constitué de 10 swaps et un Call Européen dont la diffusion est corrélée avec la diffusion du taux court. Pour les swaps nous utilisons la même configuration que précédemment Table 3 avec pour différence la maturité de chaque swap : Le 1er swap mature à 1 an, le 2ème à 2 ans, ..., le 10ème à 10 ans. Pour le Call Européen nous supposons que le sous-jacent est solution de l'EDS suivante :

$$dS_t = r_t S_t dt + \sigma S_t dW_3(t),$$

où $W_3(t)$ est un \mathcal{G}_t -mouvement brownien sous \mathbb{Q} , r_t le taux court diffusé par le modèle G2++ . La matrice de variance covariance des 3 mouvements Brownien est donnée par :

$$\Gamma = \begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} \\ \rho_{1,2} & 1 & \rho_{2,3} \\ \rho_{1,3} & \rho_{2,3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & -0.99184 & 0.1 \\ -0.99184 & 1 & 0.1 \\ 0.1 & 0.1 & 1 \end{pmatrix}$$

de vecteurs propres 1.99184, 1.01977 et -0.0116105 . C'est cette dernière valeur propre négative qui implique la non existence de décomposition de Cholesky de Γ . Une application de la méthode de Qi et Sun [19] sur Γ fournit la matrice de covariance définie positive $\tilde{\Gamma}$:

$$\tilde{\Gamma} = \begin{pmatrix} 1 & -0.980851 & 0.0978497 \\ -0.980851 & 1 & 0.0978497 \\ 0.0978497 & 0.0978497 & 1 \end{pmatrix}$$

de vecteurs propres 1.98085, 1.01915 et $1.07492e - 06$. Le spectre des matrices résultantes des 3 méthodes de calcul de $\tilde{\Gamma}$ est affiché sur la figure 8.

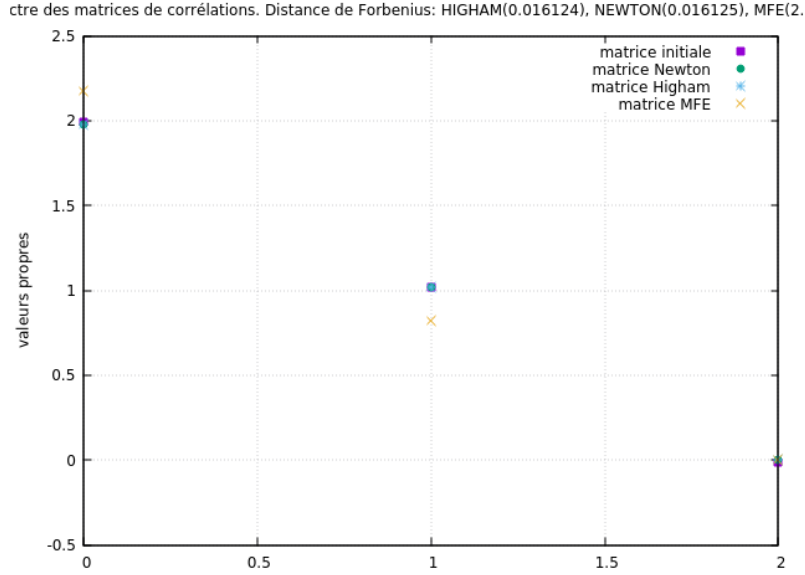


FIGURE 8 – Spectre de la matrice initiale Γ , et des matrices résultantes des trois méthodes.

Sur la figure 10 et 9 on trace le Mark-to-Market ainsi que le profil d'exposition. Enfin sur la figure 7 et le tableau 8 la dérivé de la CVA par rapport aux ténors de la courbe de taux. Concernant le temps de calcul, tableau 7, on remarque encore un gain de temps avec la méthode AAD.

Méthode	Temps
Différence Finie	39.367 s
AAD	22.6119 s

TABLE 7 – Temps de calcul de la CVA et des 12 dérivés - Différence Finie contre AAD.

CVA (IC - borne sup)	0.681027
CVA	0.669063
CVA (IC - borne inf)	0.657098
$\frac{\partial CVA}{\partial tenor_1}$	-0.00253255094
$\frac{\partial CVA}{\partial tenor_2}$	-0.01516874573
$\frac{\partial CVA}{\partial tenor_3}$	-0.0499098883
$\frac{\partial CVA}{\partial tenor_4}$	-0.1024719361
$\frac{\partial CVA}{\partial tenor_5}$	-0.735042353
$\frac{\partial CVA}{\partial tenor_6}$	-2.683617743
$\frac{\partial CVA}{\partial tenor_7}$	-9.5190282824
$\frac{\partial CVA}{\partial tenor_8}$	-14.32508823
$\frac{\partial CVA}{\partial tenor_9}$	17.379419095
$\frac{\partial CVA}{\partial tenor_{10}}$	49.453400499
$\frac{\partial CVA}{\partial tenor_{11}}$	1.90610530927
$\frac{\partial CVA}{\partial tenor_{12}}$	0
$\Delta = \sum_{i=1}^{N_{tenor}} \frac{\partial CVA}{\partial tenor_i}$	41.3060651

TABLE 8 – Résultats.

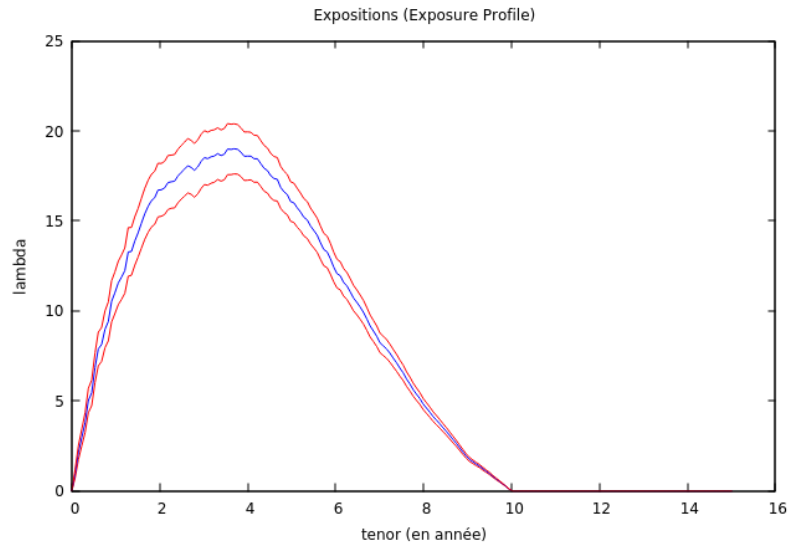


FIGURE 9 – Profil d'expositions : $EE(t)$ en fonction de t et son intervalle de confiance.

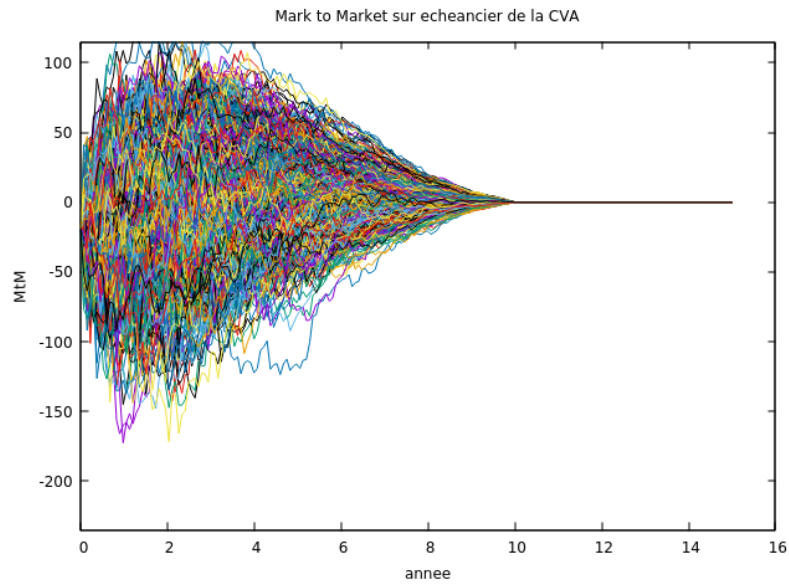


FIGURE 10 – Mark to Market : $V(t, \omega_j)$ en fonction de t et pour plusieurs scénarios ω_j .

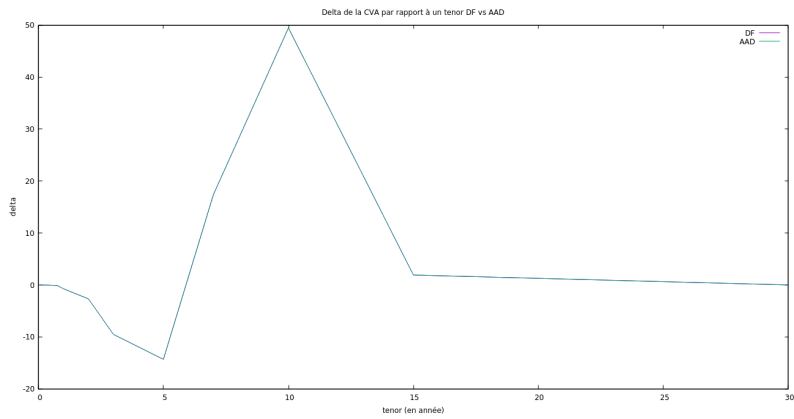


FIGURE 11 – Delta : $\frac{\partial CVA}{\partial tenor}$ en fonction des *tenors* calculés par Différences Finie et via l'AAD.

6 Conclusion

Dans ce rapport nous avons appliqué la méthode de l'AAD au calcul des sensibilités de la CVA par rapport à la courbe des taux. Une comparaison avec la méthode standard des différences finie permet d'observer un gain de temps. Dans la modélisation de la CVA nous utilisons un algorithme de calcul de matrice de corrélation symétrique définie positive la plus proche pour faire face au problème de matrice de variance covariance ne possédant pas de décomposition de Cholesky. une amélioration possible serait d'appliquer la méthode de l'AAD pour le calcul des sensibilités des la CVA par rapport aux sensibilités. Le fait de ne pas avoir recourt à des librairies externe et d'implémenter chaque méthode permet d'en calculer leur adjoint. Cependant l'algorithme de Qi et Sun [19] étant très complexe, il en est d'autant plus difficile de calculer son adjoint. L'algorithme de MFE [17] étant plus simple, c'est naturellement vers celui-ci que nos travaux se pencheront.

Dans la modélisation du swap de taux, nous utilisons la même courbe de taux pour l'actualisation et le taux forward. Une différente approche (multicurving) consisterait à modéliser le spread de base dans l'évaluation du swap de taux d'intérêts.

Enfin, une différente approche de calcul des expositions consisterait à utiliser une régression (American Monte Carlo) pour le calcul de l'espérance conditionnelle [9].

A Probabilités de défaut

La première étape consiste à construire les intensités de défauts λ_j pour tout $j \in \{1, \dots, N_{PD}\}$. Pour cela on utilise l'expression (3) avec les probabilités de défauts historiques $(PD(0, \tilde{t}_j))_{j \in \{1, \dots, N_{PD}\}}$ que nous connaissons :

$$\begin{aligned}
PD(0, \tilde{t}_0) &= 1, & (\text{par convention}) \\
PD(0, \tilde{t}_1) &= 1 - \exp\left(-\int_0^{\tilde{t}_1} \lambda(s) ds\right) = 1 - \exp(-\tilde{t}_1 \lambda_1), \\
PD(0, \tilde{t}_2) &= 1 - \exp\left(-\tilde{t}_1 \lambda_1 - (\tilde{t}_2 - \tilde{t}_1) \lambda_2\right), \\
&\vdots \\
PD(0, \tilde{t}_i) &= 1 - \exp\left(-\int_0^{\tilde{t}_i} \lambda(s) ds\right) = 1 - \exp\left(-\sum_{j=1}^{N_{PD}} \lambda_j \int_0^{\tilde{t}_i} 1_{\tilde{t}_{j-1} < s \leq \tilde{t}_j} ds\right) \\
&= 1 - \exp\left(-\sum_{j=1}^i (\tilde{t}_j - \tilde{t}_{j-1}) \lambda_j\right),
\end{aligned}$$

On en déduit donc le calcul itératif des intensités de défauts λ_i pour tout $i \in \{1, \dots, N_{PD}\}$:

$$\begin{aligned}
\lambda_1 &= \frac{-\log(1 - PD(0, \tilde{t}_1))}{\tilde{t}_1}, \\
\lambda_2 &= \frac{-\log(1 - PD(0, \tilde{t}_2)) - \tilde{t}_1 \lambda_1}{(\tilde{t}_2 - \tilde{t}_1)}, \\
&\vdots \\
\lambda_i &= \frac{-\log(1 - PD(0, \tilde{t}_i)) - \sum_{j=1}^{i-1} (\tilde{t}_j - \tilde{t}_{j-1}) \lambda_j}{(\tilde{t}_i - \tilde{t}_{i-1})}.
\end{aligned}$$

Nous en déduisons ainsi le calcul des probabilités de défauts pour pour chaque date t_i , $i \in \{1, \dots, N_{CVA}\}$:

$$\begin{aligned}
\mathbb{Q}(\tau \leq t_i) &= 1 - \exp\left(-\int_0^{t_i} \lambda(s) ds\right) \\
&= 1 - \exp\left(-\int_0^{I(t_i)} \lambda(s) ds - \int_{I(t_i)}^{t_i} \lambda(s) ds\right) \\
&= 1 - \exp\left(-\int_0^{I(t_i)} \lambda(s) ds - \int_{I(t_i)}^{t_i} \lambda(s) ds\right) \\
&= 1 - \exp\left(-\sum_{j=1}^{I(t_i)} (\tilde{t}_j - \tilde{t}_{j-1}) \lambda_j - \int_{I(t_i)}^{t_i} \lambda(s) ds\right) \\
&\approx 1 - \exp\left(-\sum_{j=1}^{I(t_i)} (\tilde{t}_j - \tilde{t}_{j-1}) \lambda_j - (t_i - I(t_i)) \lambda_{I(t_i)+1}\right).
\end{aligned}$$

B Problème de Corrélation

Nous reprenons un résultat théorique important en analyse qu'on peut retrouver dans (Luenberger, 1969, [16]) et qui permet de démontrer l'existence et l'unicité du minimum de (7). On note $K = \mathcal{S}_+^n \cap U$. Montrons qu'il existe un unique $\tilde{\Gamma} \in K$ tel que :

$$\tilde{\Gamma} = \min_{X \in K} \|\Gamma - X\|$$

Et qu'une condition nécessaire et suffisante pour que $\tilde{\Gamma}$ existe est que l'angle entre Γ et n'importe quel point de K par rapport à la solution $\tilde{\Gamma}$ soit obtus :

$$\langle \tilde{\Gamma} - \Gamma, \tilde{\Gamma} - X \rangle \leq 0, \quad \forall X \in K. \quad (14)$$

Existence Soit $(\Gamma_n)_{n \in \mathbb{N}}$ une suite dans K telle que $\lim_{n \rightarrow \infty} \|\Gamma - \Gamma_n\| = \inf_{X \in K} \|\Gamma - X\| = \delta$. Considérons un parallélogramme dans K formé des quatre points suivant : Γ , Γ_i , Γ_j et $(\Gamma_i + \Gamma_j - 2\Gamma)$. Par l'identité du parallélogramme (la somme des deux diagonales est égale à la somme des quatre côtés), on a :

$$\begin{aligned} 2 \|\Gamma_i - \Gamma\|^2 + 2 \|\Gamma_j - \Gamma\|^2 &= \|\Gamma_j - \Gamma_i\|^2 + \|\Gamma_i - \Gamma + \Gamma_j - \Gamma\|^2 \\ \Rightarrow \|\Gamma_j - \Gamma_i\|^2 &= 2 \|\Gamma_j - \Gamma\|^2 + 2 \|\Gamma_i - \Gamma\|^2 - 4 \left\| \Gamma - \frac{\Gamma_i + \Gamma_j}{2} \right\|^2. \end{aligned} \quad (15)$$

Or K est un espace convexe (i.e. $\forall \Gamma_i, \Gamma_j \in K, \forall \theta \in [0, 1]$ tel que $\theta\Gamma_i + (1 - \theta)\Gamma_j \in K$).

En prenant $\theta = \frac{1}{2}$, on a donc $\frac{\Gamma_i + \Gamma_j}{2}$ qui est un élément de K et par définition de δ (qui est la distance⁴ minimale entre Γ et tout élément de K),

$$\left\| \Gamma - \frac{\Gamma_i + \Gamma_j}{2} \right\| \geq \delta. \quad (16)$$

En utilisant l'égalité (15) et l'inégalité (16), on a :

$$\|\Gamma_j - \Gamma_i\|^2 \leq 2 \|\Gamma_j - \Gamma\|^2 + 2 \|\Gamma_i - \Gamma\|^2 - 4\delta^2.$$

Le membre de droite tendant vers 0 lorsque i et j tendent vers l'infini, on en conclut que $(\Gamma_n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans un espace complet donc converge vers un élément de K qu'on note $\tilde{\Gamma}$ (i.e. $\lim_{n \rightarrow \infty} \|\Gamma_n - \tilde{\Gamma}\| = 0$). De plus, on a que $\delta = \|\Gamma - \tilde{\Gamma}\|$. En effet, par inégalité triangulaire,

$$\begin{aligned} \|\Gamma - \Gamma_n\| &= \|\Gamma - \tilde{\Gamma} + \tilde{\Gamma} - \Gamma_n\| \leq \|\Gamma - \tilde{\Gamma}\| + \|\tilde{\Gamma} - \Gamma_n\| \\ \Rightarrow \left(\|\Gamma - \Gamma_n\| - \|\Gamma - \tilde{\Gamma}\| \right)^2 &\leq \|\tilde{\Gamma} - \Gamma_n\|^2. \end{aligned}$$

Le terme de droite tendant vers 0, nous obtenons le résultat souhaité.

Unicité Supposons qu'il existe une deuxième solution $\Theta \in K$ telle que $\Theta = \delta = \inf_{X \in K} \|\Gamma - X\| = \lim_{n \rightarrow \infty} \|\Gamma - \Gamma_n\|$ et posons

$$\Gamma_n = \begin{cases} \tilde{\Gamma}, & n \text{ pair} \\ \Theta, & n \text{ impair} \end{cases}.$$

On a $\|\Gamma - \Gamma_n\| \xrightarrow{n \rightarrow \infty} \delta$, donc $(\Gamma_n)_{n \in \mathbb{N}}$ est une suite de Cauchy qui converge. On a nécessairement $\Theta = \tilde{\Gamma}$.

4. dans le sens de la norme de Frobenius.

Preuve de la CNS (14)

$$\langle \tilde{\Gamma} - \Gamma, \tilde{\Gamma} - X \rangle \leq 0, \forall X \in K.$$

Par l'absurde, supposons qu'il existe $\Lambda \in K$ tel que $\langle \tilde{\Gamma} - \Gamma, \tilde{\Gamma} - \Lambda \rangle = \epsilon > 0$.

On pose $\Gamma_\theta = \theta\Lambda + (1 - \theta)\tilde{\Gamma}$, par convexité de K , $\Gamma_\theta \in K$.

On note $f : \theta \mapsto f(\theta) = \|\Gamma - \Gamma_\theta\|$. Après développement et dérivation on trouve

$$f'(\theta) = 2 \langle \Gamma - \tilde{\Gamma}, \tilde{\Gamma} - \Lambda \rangle = -2\epsilon < 0$$

Ce qui implique que pour tout $0 \leq \theta \leq 1$, on a $\|\Gamma - \Gamma_\theta\| \leq \|\Gamma - \tilde{\Gamma}\|$ ce qui est absurde puisque $\tilde{\Gamma} = \operatorname{argmin}_{X \in K} \|\Gamma - X\|$.

Le point de départ de la méthode de Higham et de la méthode de Qi et Sun est la condition nécessaire et suffisante (14) qui permet de représenter l'unique solution que l'on note $\tilde{\Gamma}$. $\tilde{\Gamma}$ est la projection de Γ sur $\mathcal{S}^n \cap U$. En notant $\partial(\mathcal{S}^n \cap U)(\tilde{\Gamma})$ le cône normal de $\mathcal{S}^n \cap U$ en $\tilde{\Gamma}$, cette condition se réécrit $\Gamma - \tilde{\Gamma} \in \partial(\mathcal{S}^n \cap U)(\tilde{\Gamma})$.

C Pseudo-code

De la 1ère à la 19ème ligne, il s'agit des étapes du calcul de la CVA, incluant la diffusion de la courbe de taux, la modélisation des probabilités de défauts, la modélisation du portefeuille et de son exposition. C'est le premier balayage. Pour le calcul des dérivés via l'AAD, on effectue un deuxième balayage (lignes 20 à 37) pour le calcul des adjoints des étapes du premier balayage. En sortie nous avons la dérivé de la CVA par rapport à chaque ténor de la courbe des taux.

Algorithme 3 Calcul de la CVA et de ses sensibiliés via l'AAD

Entrées : courbe term structure $(P^M(0, tenor_1), P^M(0, tenor_2), \dots, P^M(0, tenor_{N_{tenor}}))$, paramètres de diffusions du taux court, paramètres du swap de taux.

Premier balayage : calcul de la CVA

diffusion des ZC :

1: **for** $i = 1 : N_{CVA}$ **do**

2: **for** $j = 1 : N_{MC}$ **do**

3: **for** $k = 1 : N_{tenor}$ **do**

4: $P^j(t_i, t_i + tenor_k) = \frac{P^M(0, t_i + tenor_k)}{P^M(0, t_i)} \exp(f(\omega_j, t_i, t_i + tenor_k))$

5: **end for**

6: **end for**

7: **end for**

Calcul des prix du swap :

8: **for** $i = 1 : N_{CVA}$ **do**

9: **for** $j = 1 : N_{MC}$ **do**

10: $V(t_i, \omega_j) = N \sum_{k=1}^{N_{var}^i} \tau(T_{k-1}^{i,var}, T_k^{i,var}) P^j(t_i, T_k^{i,var}) F(t_i, T_{k-1}^{i,var}, T_k^{i,var}) -$
 $NK \sum_{k=1}^{N_{fix}^i} \tau(T_{k-1}^{i,fix}, T_k^{i,fix}) P^j(t_i, T_k^{i,fix})$

11: **end for**

12: **end for**

Calcul des expositions futures :

13: **for** $i = 1 : N_{CVA}$ **do**

14: $EE(t_i) = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \max(V(t_i, \omega_j), 0)$

15: **end for**

Calcul de la CVA :

16: **for** $i = 1 : N_{CVA}$ **do**

17: $\alpha_i = P^M(0, tenor_i) EE(t_i)$

18: **end for**

19: $CVA = (1 - R) \sum_{i=1}^{N_{CVA}} \alpha_i (PD(t_i) - PD(t_{i-1}))$

Deuxième balayage : Calcul des sensibilités via l'AAD

Calcul des adjoints des α :

20: **for** $i = 1 : N_{CVA}$ **do**

21: $\bar{\alpha}_i = \frac{\partial CVA}{\partial \alpha_i} = (1 - R) (PD(t_i) - PD(t_{i-1}))$

22: **end for**

Calcul des adjoints des EE :

23: **for** $i = 1 : N_{CVA}$ **do**

24: $\bar{EE}(t_i) = \frac{\partial CVA}{\partial EE(t_i)} = \frac{\partial CVA}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial EE(t_i)} = \bar{\alpha}_i P^M(0, t_i)$

25: **end for**

Calcul des adjoints des P^M :

26: **for** $i = 1 : N_{CVA}$ **do**

27: $\bar{P}^M(0, t_i) = \frac{\partial CVA}{\partial P^M(0, t_i)} = \frac{\partial CVA}{\partial \alpha_i} \frac{\partial \alpha_i(P^M(0, t_i), EE(t_i))}{\partial P^M(0, t_i)} = \bar{\alpha}_i EE(t_i)$

28: **end for**

Calcul des adjoints des V :

29: **for** $i = 1 : N_{CVA}$ **do**

30: **for** $j = 1 : N_{MC}$ **do**

31: $\bar{V}(t_i, \omega_j) = \frac{\partial CVA}{\partial V(t_i, \omega_j)} = \frac{\partial CVA}{\partial EE(t_i)} \frac{\partial EE(t_i)}{\partial V(t_i, \omega_j)} = \bar{EE}(t_i) \frac{1}{N_{MC}} \mathbb{1}_{V(t_i, \omega_j) > 0}$

32: **end for**

33: **end for**

Calcul du delta de la CVA \bar{ZC} :

34: $\bar{ZC} = \sum_{i=1}^{N_{CVA}} (\bar{P}^M(0, t_i) \frac{\partial P^M(0, t_i)}{\partial ZC} + \sum_{j=1}^{N_{MC}} \bar{V}(t_i, \omega_j) \frac{\partial V(t_i, \omega_j)}{\partial ZC})$

Sortie : \bar{ZC}

Références

- [1] James P. Boyle and Richard L. Dykstra. A method for finding projections onto the intersection of convex sets in hilbert spaces. *Advances in Order Restricted Inference 37, Lecture Notes in Statistics. Berlin : Springer*, 1985. pp 28–47.
- [2] Damiano Brigo and Agostino Capponi. Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps. *eprint arXiv :0812.3705*, Decembre 2008.
- [3] Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice*. Springer-Verlag Berlin Heidelberg, 2006.
- [4] Luca Capriotti. Fast greeks by algorithmic differentiation. *Journal of Computational Finance*, 2011.
- [5] Luca Capriotti and Michael Giles. Fast correlation greeks by adjoint algorithmic differentiation. *Risk*, Avril 2010.
- [6] Luca Capriotti and Michael Giles. Algorithmic differentiation : adjoint greeks made easy. *available at <http://ssrn.com/abstract=1801522>*, Avril 2011.
- [7] Luca Capriotti, Shinghoi (Jacky) Lee, and Matthew Peacock. Real-time counterparty credit risk management in monte carlo. *Risk Magazine*, 2011. Available at SSRN : <https://ssrn.com/abstract=1824864> or <http://dx.doi.org/10.2139/ssrn.1824864>.
- [8] Frank H. Clarke. Optimization and nonsmooth analysis. *Canadian Mathematical Society Series of Monographs and Advanced Texts. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York*, 1983. xiii+308.
- [9] Christian P. Fries. Automatic backward differentiation for american monte-carlo algorithms (conditional expectation). *eprint arXiv :1707.04942*, 07/2017.
- [10] Michael Giles and Paul Glasserman. Smoking adjoints : fast monte carlo greeks. *Risk*, Janvier 2006.
- [11] Andreas Griewank. *Evaluating Derivatives : Principles and Techniques of Algorithmic Differentiation*. 2000.
- [12] Andreas Griewank and Andrea Walther. *Evaluating Derivatives : Principles and Techniques of Algorithmic Differentiation, Second Edition*. 2008.
- [13] Shih-Ping Han. A successive projection method. *Mathematical Programming*, 40, 1988. pp 1-14.
- [14] E. Hestenes, M. R. et Stiefel. Methods of conjugate gradients for solving linear systems. *J. Research Nat. Bur. Standards* 49, 1952. 409–436.
- [15] Nicholas John Higham. Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis* 22, 2002. p. 329 - 343.
- [16] David G. Luenberger. *Optimization by Vector Space Methods*. Wiley-Interscience, 1969.
- [17] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management : Concepts, Techniques and Tools Revised Edition*. Princeton University Press, 2015.
- [18] C. A. Micchelli and F. I. Utreras. Smoothing and interpolation in a convex subset of a hilbert space. *SIAM Journal on Matrix Analysis and Applications* 9, 1988. pp 728-746.
- [19] Houduo Qi and Defung Sun. A quadratically convergent newton method for computing the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications* 28, 2006. 360–385.

Abstract

The aim of this report is to explain how we have applied the AAD method (Adjoint Algorithmic Differentiation) to compute sensibilities of the CVA (Credit Value Adjustment). The computation of the CVA's sensibilities is a meadow required for Hedging, piloting of the CVA desk and the implementation of the Fundamental Review of the Trading Book (FRTB). The traditional approach consists in using the finite difference method but requires to verify a number of constraints (continuity of payoffs, time / power of calculation which all the financial institutions does not have).