

Mathématiques pour la finance

Gauthier Marteville

July 12, 2024

Contents

1	Propriétés utiles des probabilités	4
1.1	Loi et fonction de répartition	4
1.1.1	Simulation de lois	4
1.1.2	Méthodes de rejet	5
1.1.3	Algorithme de Box Muller	5
1.2	Espérances conditionnelles	5
1.3	Théorème Central Limite	6
1.4	Inégalité de Markov	7
1.5	Exercices	8
1.5.1	Variable définie par une fonction de répartition	8
2	Chaines de Markov	8
2.1	A propos des graphes	8
2.2	Processus de Markov	9
2.3	Hidden Markov Model (HMM)	9
3	Mouvements Browniens	9
3.1	Définition	9
3.2	Propriétés	9
3.2.1	Inversion temporelle - Stabilité	9
3.2.2	Propriété de Markov forte	9
3.2.3	Principe de reflection	10
3.2.4	Exercices	11
3.3	Lemme d'Itô	11
3.4	Intégrale d'Itô	12
3.5	Mouvements Browniens multidimensionnels	12
3.5.1	Définition et propriétés	12
3.5.2	Mouvements browniens corrélés	13

4	Martingales	13
4.1	Introduction	13
4.2	Martingales et mouvements browniens	14
4.3	Différence de martingales	14
4.4	Convergence des martingales	14
4.5	Stopping Times	15
4.5.1	Motivations	15
4.5.2	Définitions	15
4.5.3	Martingales locale	15
4.6	Théorème de Girsanov	15
4.7	Formule de Feynman–Kac	17
4.8	Quelques utilisations des martingales	17
4.8.1	Ruine d’un joueur	17
5	Processus de Lévy	17
5.1	Définition	17
5.2	Exemple : les processus de Poisson	18
6	Notions de finance	18
6.1	Les securities	18
6.1.1	Les obligations (<i>bonds</i>)	18
6.1.2	Les actions (<i>equity, stock</i>)	19
6.2	Les produits dérivés	19
6.2.1	Les options	20
6.2.2	Les futures	20
6.2.3	Les forwards	20
6.2.4	Les swaps	21
6.2.5	Tableau récapitulatif	21
6.3	Les indices	21
6.3.1	VIX volatility index	21
7	Vanilla Option Pricing	23
7.1	Stratégie simple – Achat ou vente d’option	23
7.2	Modèle de Black-Scholes	24
7.3	Modèle de Black-Scholes-Merton	24
7.3.1	Hypothèses	24
7.3.2	Equation différentielle stochastique	24
7.3.3	Parité call-put	25
7.3.4	Dividendes	26
7.4	Le modèle binomial (Cox–Ross–Rubinstein)	27
7.4.1	Présentation	27
7.4.2	Construction de l’arbre	27
7.4.3	Calcul de la valeur de l’option sur les noeuds	28
7.4.4	Exemple d’utilisation	29
7.5	PDF du prix et formule de BREEDEN–LITZENBERGER	29

8	Autres modèles de pricing	30
8.1	Evaluation de la volatilité	30
8.1.1	Modèle GARCH(1,1)	30
8.1.2	Modèle GARCH(1,1)	30
8.1.3	Volatilité Implicite	30
8.2	Modèle de volatilité locale	31
8.2.1	Volatilité de Dupire	32
8.2.2	Constant Elasticity of Variance Model (modèle CEV)	33
8.3	Modèle de volatilité stochastique	34
8.3.1	Modèle de Heston	34
8.3.2	Stochastic Alpha, Beta, Rho model (modèle SABR)	34
8.3.3	Modèle de Bergomi	34
9	Dérivés de taux	34
9.1	Zero-coupon bond	34
9.2	Short rate models	35
9.2.1	Modèle de Cox-Ingersoll-Ross	35
9.2.2	Modèle de Hull-White	35
9.2.3	Modèle de Vasicek	35
9.3	Health-Jarrow-Morton framework	36
9.4	Market models	36
9.5	Markov functional models	36
10	Dérivé Action (EQD)	36
10.1	Les types de dérivés sur action	36
10.1.1	Les options et warrants	36
10.1.2	Les Turbos	37
10.1.3	Les Compo et Quanto	37
10.1.4	Les futures contracts	38
10.1.5	Les obligations convertibles	38
10.1.6	Les Equity Swaps	38
10.2	Réplication d'une option digitale	39
10.3	Evaluation d'un contrat forward et future	40
10.3.1	Prix d'un future – Arbitrage <i>cash and carry</i>	40
10.3.2	Prix d'un forward	40
10.3.3	Détermination sous Black-Scholes	40
10.4	Les Greeks	40
10.4.1	Le delta δ	41
10.4.2	Le gamma γ	42
10.4.3	Le thêta θ	42
10.4.4	Le rho ρ	43
10.4.5	le véga ν	44
10.4.6	Exemple d'utilisation	44
11	Options exotiques	45
11.1	Options barrière	45
11.2	Baskets	45
11.2.1	Définition	45
11.2.2	Matrice de corrélation	46
11.2.3	Corrélation implicite	46

11.3 Autocallables	46
11.3.1 Description	46
11.4 Hedge d'une option barrière	46
12 Etude de cas – Stochastic repo hybrid model	47
12.1 Repo rate dynamics	47
12.2 Linear Gaussian Markov model (LGM)	47
12.3 Exchanges rate model	47
12.4 Pricing dans une devise étrangère	47
12.5 Calibration des corrélations	47
12.5.1 Quanto options	47
12.5.2 Hybrid stock/funding exchange	47
12.5.3 Hybrid basket	47
13 Dérivés de Crédit	47
14 Modèles de regression	47
14.1 Décomposition biais-variance	47
14.2 Moindres carrés (OLS)	48
14.3 Regression Ridge (L^2)	48
14.4 Regression Lasso (L^1)	50
14.4.1 Enoncé du problème	50
14.4.2 Forward Stagewise Algorithm	51
15 Méthodes numériques	52
15.1 Monte Carlo	52
15.1.1 Algorithme de base	52
15.1.2 Réduction de variance	52
15.1.3 Markov Chain Monte Carlo	52
15.2 Résolutions de PDE	52
15.2.1 Adjoint Algorithmic Differentiation	52
15.2.2 Finite Differences	52
16 Neural Networks	53
16.1 Artificial Neural Network	53
16.1.1 Motivations	53
16.1.2 Fonctionnement	54
16.1.3 Output units	56
16.1.4 Fonctions d'activation	56
17 Brainteasers	57

1 Propriétés utiles des probabilités

1.1 Loi et fonction de répartition

1.1.1 Simulation de lois

Définition 1 (Fonction quantile) Soit F une fonction de répartition. On appelle fonction quantile de F la fonction suivante, définie pour tout u dans

$]0; 1[$

$$F^{-1}(u) = \inf\{x | F(x) \geq u\}$$

La fonction quantile peut nous permettre d'établir un lien entre les loi de probabilités, qui est utile lorsque l'on veut simuler une loi. En effet, lorsqu'une variable aléatoire X est à densité, alors la fonction quantile F^{-1} est l'inverse de la fonction de répartition.

Propriété 1.0.1 *Soit U une loi uniforme sur un intervalle I . Alors la variable aléatoire X définie par $X = F^{-1}(U)$ a pour fonction de répartition F .*

On peut alors par exemple générer une loi exponentielle à partir de la loi uniforme sur $(0, 1)$. Il suffit de prendre pour F la fonction de répartition de la loi exponentielle, $F : x \rightarrow 1 - \exp(-\lambda x)$ et de trouver la fonction inverse. Pour un $\lambda > 0$ on a alors

$$X = -\frac{1}{\lambda} \log(1 - U) = -\frac{1}{\lambda} \log(U)$$

et X suit alors une loi exponentielle, avec U loi uniforme sur $(0, 1)$

1.1.2 Méthodes de rejet

1.1.3 Algorithme de Box Muller

C'est une méthode pour simuler un couple de variables aléatoires indépendantes de loi normale centrée réduite. Elle est basée sur le théorème suivant

Théorème 1.1 (Box Muller) *Soient R et Θ deux variables aléatoires indépendantes telles que R^2 suit une loi exponentielle de paramètre $1/2$ et Θ suit une loi uniforme sur $[0, 2\pi]$ Alors les variables définies par*

$$X = R \cos(\Theta) \text{ et } Y = R \sin(\Theta)$$

sont indépendantes et suivent tout deux une loi normale centrée réduite

Il suffit donc pour simuler deux variables aléatoires de loi normale centrée réduite de savoir simuler une loi exponentielle et une loi uniforme.

Démonstration 1

1.2 Espérances conditionnelles

Etant donné deux variables aléatoires, on peut définir l'espérance de X conditionnée par Y . Elle se note $\mathbb{E}[X|Y]$ et c'est une nouvelle variable aléatoire. Dans le cas où Y est une variable aléatoire discrète, elle est définie comme égale à $\varphi(Y)$ où φ est la fonction presque partout définie par : $\varphi(y) = \mathbb{E}[X|\{Y = y\}]$. L'espérance de X sachant B , si X est d'espérance finie, est définie par

$$\mathbb{E}[X|B] = \sum_{x_i} x_i \mathbb{P}_B(X = x_i) \tag{1}$$

1.3 Théorème Central Limite PROPRIÉTÉS UTILES DES PROBABILITÉS

Dans le cas général on note $\mathbb{E}[X|Y]$ l'espérance conditionnellement à une variable aléatoire, par le biais de la tribu engendrée par cette variable aléatoire $\mathbb{E}[X|\sigma(Y)]$.

Théorème 1.2 Soient X, Y et Z des variables aléatoires réelles sur un espace de probabilité. Alors

- (i) Espérance de l'espérance conditionnelle : $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$
- (ii) Si Z est \mathcal{F} -mesurable, $\mathbb{E}[XZ|\mathcal{F}] = Z\mathbb{E}[X|\mathcal{F}]$
- (iii) Si X est \mathcal{F} -mesurable, $\mathbb{E}[X|\mathcal{F}] = X$
- (iv) Inégalité de Jensen. Si ϕ est convexe et si $\phi(X)$ est intégrable alors, Si X est \mathcal{F} -mesurable, on a $\mathbb{E}[\phi(X)|\mathcal{F}] \geq \phi(\mathbb{E}[X|\mathcal{F}])$
- (v) **important. (propriété du double conditionnement)** Seul reste le conditionnement par **la tribu la plus petite**: Si $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ alors,

$$\mathbb{E}[\mathbb{E}(X|\mathcal{F}_{n+1})|\mathcal{F}_n] = \mathbb{E}[X|\mathcal{F}_n]$$

1.3 Théorème Central Limite

Ce théorème nous assure la convergence en loi d'une somme de variables aléatoires i.i.d. vers la loi normale centrée réduite. C'est-à-dire que cette somme tend en loi vers une variable aléatoire Gaussienne.

Théorème 1.3 (Théorème Central Limite) Soit un entier naturel N et soient $(X_i)_{i \in \mathbb{N}_N}$ des variables aléatoires i.i.d suivant une loi D . On suppose qu'il existe pour chaque X_i une espérance finie μ et un écart-type σ . On pose $S_n = \sum_{i \in \mathbb{N}_N} X_i$. Alors l'espérance de S_n est $n\mu$ et son écart-type vaut $\sqrt{n}\sigma$. De plus pour n assez grand la loi normale $\mathcal{N}(n\mu, \sqrt{n}\sigma)$ est une bonne approximation de la loi de S_n . Plus formellement, On pose

$$\overline{X_n} = \frac{S_n}{n} = \frac{\sum_{i=1}^N X_i}{n} \quad (2)$$

et

$$Z_n = \frac{\overline{X_n} - \mu}{\sqrt{n}\sigma} \quad (3)$$

Alors le théorème central limite nous assure que la suite de V.A. $(Z_n)_{n \in \mathbb{N}}$ converge en loi vers $\mathcal{N}(0, 1)$

Démonstration 2 (Démonstration) Supposons que X_1, X_2, \dots, X_n soient des variables aléatoires indépendantes et identiquement distribuées avec une moyenne

μ et une variance σ^2 . Alors, la somme $S_n = \sum_{i=1}^n X_i$ suit approximativement une distribution normale avec une moyenne $E[S_n] = n\mu$ et une variance $\text{Var}[S_n] = n\sigma^2$, lorsque n est grand. En d'autres termes, nous pouvons écrire :

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x)$$

où $\Phi(x)$ est la fonction de répartition de la distribution normale standard.

Pour démontrer le théorème central limite, commençons par calculer l'espérance et la variance de S_n :

$$E[S_n] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \mu = n\mu$$

De même, nous avons :

$$\text{Var}[S_n] = \text{Var}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \text{Var}[X_i] = \sum_{i=1}^n \sigma^2 = n\sigma^2$$

Nous pouvons maintenant utiliser la loi des grands nombres pour affirmer que la moyenne empirique $\bar{X} = \frac{S_n}{n}$ converge vers la moyenne théorique μ lorsque n devient grand. En d'autres termes, \bar{X} est un estimateur convergent de μ . Ainsi, nous pouvons écrire :

$$\lim_{n \rightarrow \infty} P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$$

où $\Phi(x)$ est la fonction de répartition de la distribution normale standard.

En multipliant numérateur et dénominateur par \sqrt{n} , nous obtenons :

$$\lim_{n \rightarrow \infty} P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) = \lim_{n \rightarrow \infty} P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq x\sqrt{n}\right) = \Phi(x\sqrt{n})$$

Maintenant, considérons la variable $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$. Nous avons :

1.4 Inégalité de Markov

Théorème 1.4 (Inégalité de Markov) Soit Z une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ supposée presque sûrement positive ou nulle. Alors

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a} \quad (4)$$

Corrolaire 1 (Inégalité de Markov généralisée) On suppose Φ une fonction positive croissante sur un intervalle non trivial I . Soit Y une variable aléatoire réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que $\mathbb{P}(Y \in I) = 1$. Alors

$$\forall b \in I, \Phi(b) > 0, \quad \mathbb{P}(Y \geq b) \leq \frac{\mathbb{E}[\Phi(b)]}{\Phi(b)} \quad (5)$$

Théorème 1.5 (Inégalité de Bienaymé-Tchebychev) Soit X une variable aléatoire de variance finie σ^2 et d'espérance μ . Alors pour tout réel strictement positif α ,

$$\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}. \quad (6)$$

1.5 Exercices

1.5.1 Variable définie par une fonction de répartition

On considère une loi normale centrée réduite $\mathcal{N}(0, 1)$ et deux réels a et b . Montrer que

$$\mathbb{E}[\Phi(aX + b)] = \Phi\left(\frac{b}{\sqrt{1 + a^2}}\right)$$

Avec Φ la fonction de répartition de la loi normale centrée réduite. **Démonstration.**

On peut écrire

$$\mathbb{E}[\Phi(aX + b)] = \mathbb{E}[\mathbb{P}(Y \leq ax + b | X = x)] = \mathbb{E}[\mathbb{P}(Y - ax \leq b | X = x)] = \mathbb{P}(Y - aX \leq b)$$

Or $Y - aX \sim \mathcal{N}(0, 1 + a^2)$ d'où le résultat.

2 Chaines de Markov

2.1 A propos des graphes

1. **Probabilité d'Existence d'une Arête :** Dans un graphe aléatoire $G(n, p)$, la probabilité qu'il existe une arête entre deux sommets donnés est p . Cela signifie que chaque paire de sommets est connectée indépendamment avec une probabilité p .
2. **Propriétés Asymptotiques :** Lorsque le nombre de sommets n devient très grand, certains résultats asymptotiques intéressants émergent. Par exemple, le modèle $G(n, p)$ peut exhiber des transitions de phase où des propriétés structurelles importantes changent de manière significative.
3. **Composantes Connexes :** En fonction de la probabilité p , un graphe aléatoire peut avoir une unique composante connexe géante lorsque p dépasse un certain seuil critique. En dessous de ce seuil, des petites composantes connexes isolées peuvent apparaître.
4. **Degré Moyen :** Le degré moyen d'un sommet dans un graphe aléatoire $G(n, p)$ est $(n-1)p$. Cela donne une indication de la densité des connexions dans le graphe.
5. **Propriétés de Phase de Transition :** Certains graphes aléatoires passent par des phases de transition, où des propriétés comme la connectivité globale changent brusquement en fonction de la probabilité p .

Les graphes aléatoires sont des outils puissants pour modéliser et analyser des systèmes complexes, et leur étude a des implications importantes dans divers domaines, y compris l'informatique, la physique statistique et les sciences sociales.

2.2 Processus de Markov

Définition 2 Un processus de Markov $(X_n)_{n \geq 0}$ à temps discret est un processus stochastique vérifiant la propriété de Markov.

$$P(X_{n+1}|X_n, X_{n-1}, \dots, t_0) = P(X_{n+1}|X_n) \quad (7)$$

2.3 Hiden Markov Model (HMM)

3 Mouvements Browniens

3.1 Définition

Définition 3 (Mouvement Brownien) Un mouvement brownien est un processus stochastique $(B_t)_{t \geq 0}$ continu qui modélise le mouvement aléatoire d'une particule dans un fluide. Il vérifie les trois propriétés suivantes

- (i) (Accroissements indépendants) Pour tous s, t tels que $t > s$ l'accroissement $B_t - B_s$ est indépendant du processus $(B_u)_{0 \leq u < s}$.
- (ii) (Accroissements stationnaires et gaussiens) Quels que soient les temps t et s tels que $t > s$, l'accroissement $B_t - B_s$ est une variable aléatoire normale de moyenne nulle et de variance $t - s$.
- (iii) $(B_t)_{t \geq 0}$ est presque sûrement continu, c'est-à-dire pour presque toute réalisation, la fonction $t \mapsto B_t(\omega)$ est continue.

Si $X_0 = 0$ on dit que le mouvement brownien est **standard**.

3.2 Propriétés

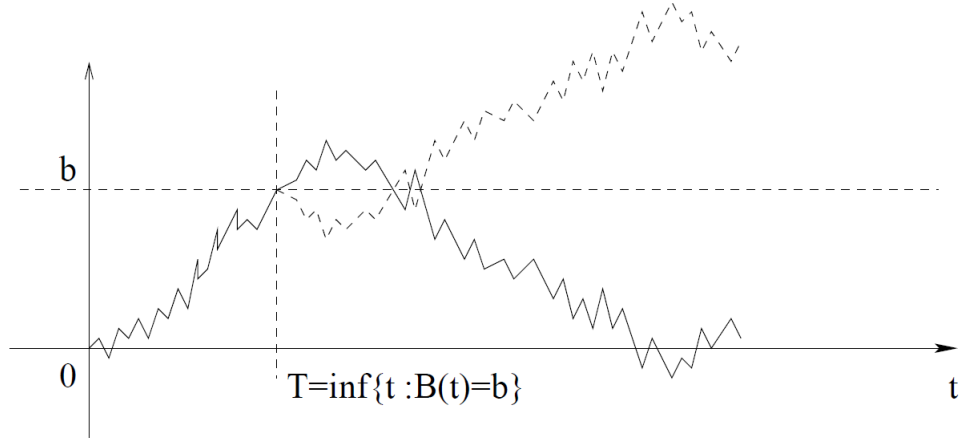
3.2.1 Inversion temporelle - Stabilité

Propriété 3.0.1 Le processus $(tB_{\frac{1}{t}})$, qui s'annule en $t = 0$, est un mouvement brownien.

Propriété 3.0.2 Pour tout $c > 0$, le processus $(cB_{\frac{t}{c^2}})$ est un mouvement brownien. On dit qu'il est stable d'indice 2.

3.2.2 Propriété de Markov forte

Propriété 3.0.3 Le mouvement brownien vérifie la propriété de Markov forte. Pour tout temps d'arrêt T , conditionnellement à $(T < +\infty)$, le processus $(B_t^T)_{t \geq 0} = (B_{t+T} - B_T)_{t \geq 0}$ est un mouvement brownien indépendant de $(B_s)_{0 \leq s < T}$

FIGURE 1 – Mouvement brownien réfléchi du temps d'arrêt T

3.2.3 Principe de reflection

Théorème 3.1 Soit T un temps d'arrêt et $(B_t)_t$ un Brownien standard. Alors le Reflected brownian motion at the stopping time T défini par:

$$B_T^* = B_t \mathbf{1}\{t \leq T\} + (2B_T - B_t) \mathbf{1}\{t > T\}$$

Est aussi un brownien standard.

Démonstration 3 On utilise la propriété de Markov forte aux deux browniens $(B_{t+T} - B_t)_t$ et $\left(-(B_{t+T} - B_t) \right)_t$

Application. On peut ainsi montrer la propriété suivante:

Propriété 3.1.1 Pour tout mouvement brownien $(B_s)_s$ on a la relation

$$\mathbb{P}\left(\max_{0 \leq s \leq t} B_s \geq a\right) = 2\mathbb{P}(B_t \geq a)$$

Démonstration 4 Posons $M_t = \max_{0 \leq s \leq t} B_s$. On peut déjà affirmer que $M_t \geq 0$ car $B_0 = 0$ et $(M_t)_t$ est décroissant en t . On pose $T = \inf\{s \geq 0, B_s = a\}$ et B_t^* le brownien réfléchi de T . Alors

$$\{M_t > a\} = \{M_t > a \cap (B_t > a \sqcup B_t \leq a)\} = \{(M_t > a \cap B_t > a) \sqcup (M_t > a \cap B_t \leq a)\}$$

On a donc

$$\mathbb{P}(M_t > a) = \mathbb{P}((M_t > a \cap B_t > a)) + \mathbb{P}((M_t > a \cap B_t \leq a)) = \mathbb{P}(B_t > a) + \mathbb{P}(B_t^* \geq a)$$

Par le principe de réflexion, on obtient le résultat.

3.2.4 Exercices

Ex.1 – Deutsche Bank SHL question Calculer $\mathbb{P}(B_1 > 0, B_2 > 0, \inf_{t \in [1,2]} B_t < 0)$
 0) On utilise le brownien réfléchi par le temps d'arrêt $\tau = \{\inf t, B_t = 0, t > 1\}$
 et on remarque que

$$\mathbb{P}(B_1 > 0, B_2 > 0, \inf_{t \in [1,2]} B_t < 0) = \mathbb{P}(B_1 > 0, B_2 < 0)$$

Et on obtient ainsi

$$\mathbb{P}(B_1 > 0, B_2 > 0, \inf_{t \in [1,2]} B_t < 0) = 1/8$$

Ex. 2 – Fonction de répartition Soit B un mouvement brownien et $x \in \mathbb{R}_+$.
 Calculons $\mathbb{P}(|B_t| > x)$.

Si $X \sim \mathcal{N}(0, 1)$, On a $\mathbb{P}(|B_t| > x) = \mathbb{P}(|X| > \frac{x}{\sqrt{t}}) = 2N\left(-\frac{x}{\sqrt{t}}\right)$

On note $N(x)$ la fonction de répartition de la loi normale. On rappelle que

$$N(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x-\mu)^2}{2\sigma^2} \quad (8)$$

On remarque alors que les réalisations x de B_t vérifient $N(x) + N(-x) = 1$

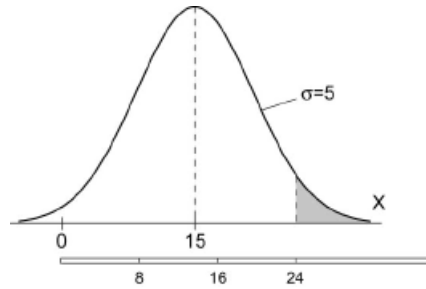


FIGURE 2 – Fonction de répartition de la loi normale

On a alors

$$\mathbb{E}[N(B_t)] + \mathbb{E}[N(-B_t)] = 1$$

Or B_t et $-B_t$ suivent la même loi. Alors $\mathbb{E}[N(B_t)] = \frac{1}{2}$.

3.3 Lemme d'Itô

Le lemme d'Itô permet de résoudre des équations différentielles stochastiques (EDS) très utile pour manipuler le modèle de Black-Scholes (cf. après).

Théorème 3.2 (Lemme d'Itô) Soit X_t un processus stochastique de la forme

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$$

On peut aussi formuler la version différentielle de l'énoncé i.e. le processus X_t vérifie

$$dX_t = \mu_t dt + \sigma_t dB_t$$

Avec σ_t et μ_t des processus aléatoire et B_t un mouvement brownien. Alors si $f(X_t, t)$ est une fonction de classe C^2 alors on peut écrire

$$d(f(X_t, t)) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma_t^2 dt \quad (9)$$

On peut écrire aussi sous la forme, grace à la relation entre X_t , dt et dW_t

$$d(f(t, X_t)) = \left(\frac{\partial f}{\partial t} + \mu X_t \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma_t^2 \right) dt + \sigma \frac{\partial f}{\partial x} X_t dW_t$$

Le lemme d'Ito donne une formule pour calculer la dérivée d'une fonction aléatoire en termes des dérivées du processus stochastique sous-jacent.

3.4 Intégrale d'Itô

Définition. Pour une fonction $f : [0, T] \rightarrow \mathbb{R}$ (peut être aléatoire) et W un brownien, l'Intégrale d'Itô est définie par:

$$\int_0^T f(t) dW_t = \lim_{N \rightarrow \infty} \sum_{n=0}^N f(t_n) (W_{n+1} - W_n) \quad (10)$$

Propriété 3.2.1 (Isométrie d'Itô) Pour toute fonction f, g non anticipative de W , de moment d'ordre 2 fini, alors

$$\mathbb{E} \left(\int_0^t f(s) dW_s \right) \left(\int_0^t g(s) dW_s \right) = \mathbb{E} \int_0^t f(s) g(s) ds \quad (11)$$

Exemple. On cherche à calculer, pour W un brownien standard, $\mathbb{E}(W_t \int_0^t s dW_s)$. On a alors

$$\mathbb{E} \left(W_t \int_0^t s dW_s \right) = \mathbb{E} \left(\int_0^t 1 dW_s \right) \left(\int_0^t dW_s \right) = \mathbb{E} \int_0^t 1 s ds = \quad (12)$$

3.5 Mouvements Browniens multidimensionnels

3.5.1 Définition et propriétés

Définition 4 Soit $B_t = (B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(n)})^T$ un processus n -dimensionnel. On dit que B_t est un mouvement brownien multidimensionnel si les $(B_t^{(i)})_i$ sont des mouvements browniens indépendants entre eux.

Théorème 3.3 Le processus à n -dimensions B_t est un mouvement brownien si et seulement si les processus $B^{(i)}$ et $B^{(i)} B^{(j)} - \delta_{i,j} t$ sont des martingales, où $\delta_{i,j}$ désigne le symbole de Kronecker.

Démonstration 5 *WIP...*

3.5.2 Mouvements browniens corrélés

Définition 5 (Mouvements browniens corrélés) On dit de deux mouvements browniens B_t^1 et B_t^2 qu'ils sont corrélés de coefficient de corrélation ρ si $B^1(t)B^2(t) - \rho t$ est une martingale.

Si on dispose de deux mouvements browniens indépendants W_t et \tilde{W}_t et de $\rho \in [-1, 1]$ On peut alors construire un processus \hat{W}_t corrélé comme suit:

$$\hat{W}(t) = \rho W(t) + \sqrt{1 - \rho^2} \tilde{W}(t)$$

A l'inverse, en disposant de deux mouvements corrélés W_t et \tilde{W}_t de coefficient de corrélation ρ , on peut construire un processus \hat{W}_t qui sera un mouvement brownien indépendants des deux premiers:

$$\hat{W}(t) = \frac{1}{\sqrt{1 - \rho^2}} (\tilde{W}(t) - \rho W(t))$$

4 Martingales

4.1 Introduction

Définition 6 (Martingale) Une séquence $Y = (Y_n, n \geq 0)$ est une martingale suivant une séquence $X = (X_n, n \geq 0)$ si pour tout $n \geq 0$

- (i) $\mathbb{E}[Y_n] < +\infty$
- (ii) $\mathbb{E}[Y_{n+1} | X_0, X_1, \dots] = Y_n$

Définition 7 (Filtration) Une filtration est une suite croissante (pour l'inclusion) de parties d'un ensemble. On définit alors une martingale suivant une filtration \mathcal{F} sur $(\Omega, \mathcal{F}, \mathbb{P})$ comme une paire $(Y, \mathcal{F}) = ((Y_n, \mathcal{F}_n), n \geq 0)$ telle que :

- (i) $\mathbb{E}[Y_n] < +\infty$
- (ii) $\mathbb{E}[Y_{n+1} | \mathcal{F}_n] = Y_n$

Soit Y une séquence de variables aléatoires adaptées à une filtration \mathcal{F} . En posant $X^+ = \max(0, X)$ et $X^- = \min(0, X)$ on peut définir les supermartingales et les submartingales comme suit:

- (i) **supermartingale** : $\mathbb{E}[Y_n] < +\infty$ et $\mathbb{E}[Y_{n+1} | \mathcal{F}_n] \leq Y_n$
- (ii) **submartingale** : $\mathbb{E}[Y_n] < +\infty$ et $\mathbb{E}[Y_{n+1} | \mathcal{F}_n] \geq Y_n$

Théorème 4.1 (Inégalité maximale de Doob) Soit $(M_n)_{n \geq 0}$ une sous-martingale positive. Pour tout $a > 0$ et tout $n \in \mathbb{N}$, on a

$$\mathbb{P}\left(\max_{0 \leq k \leq n} M_k \geq a\right) \leq \frac{\mathbb{E}[M_n]}{a} \quad (13)$$

Théorème 4.2 (Décomposition de Doob) Une submartingale (Y, \mathcal{F}) avec une espérance finie peut être décomposée de la forme suivante:

$$Y_n = M_n + S_n \quad (14)$$

Où (M, \mathcal{F}) est une martingale et (S, \mathcal{F}) est un processus croissant prédictible.

4.2 Martingales et mouvements browniens

Proposition 1. $(B_t)_t$ est une martingale

dem. Les moments d'ordre 1 et 2 sont finis car $B \sim \mathcal{N}(0, t)$. Ensuite,

$$\mathbb{E}[B_t|F_t] = \mathbb{E}[B_s + (B_t - B_s)|F_s] = B_s + \mathbb{E}[B_t - B_s|F_s] = B_s \quad (15)$$

car $B_t - B_s$ est indépendant de \mathcal{F}_n (suite croissante de tribus).

Proposition 2. $\text{cov}(B_s, B_t) = \inf(s, t)$

dem. De la même manière, $\text{cov}(B_t, B_s) = \text{cov}(B_s, B_s + (B_t - B_s)) = \text{var}(B_s) + \text{cov}(B_s - B_0, B_t - B_s) = \text{var}(B_s) = s$ si $s < t$ et on a le résultat avec $s > t$.

Proposition 3. $B_t^2 - t$ est une martingale.

dem. Même principe que pour les points précédents.

Proposition 4. $\exp\left(\lambda B_t - \frac{\lambda^2 t}{2}\right)$ est une martingale.

dem. On utilise le fait que si $X \sim \mathcal{N}(\mu, \sigma^2)$ alors

$$\mathbb{E}[e^{aX}] = e^{a\mu + \frac{a^2\sigma^2}{2}} \quad (16)$$

4.3 Différence de martingales

Théorème 4.3 (Inégalité de Hoeffding) Soit (Y, \mathcal{F}) une martingale. On suppose qu'il existe une suite de réels $(K_i)_{i \in \mathbb{N}}$ telle que pour tout n , on ait $\mathbb{P}(|Y_n - Y_{n-1}| \geq K_n) = 1$. Alors

$$\mathbb{P}(|Y_n - Y_0| \geq x) \leq 2 \exp\left(-\frac{1}{2}x^2 \Big/ \sum_{i=1}^n K_i^2\right) \quad (17)$$

C'est-à-dire que si la martingale est presque sûrement majorée. Il y a une faible probabilité de déviation de Y_n de sa valeur initiale Y_0 .

4.4 Convergence des martingales

Théorème 4.4 (Convergence des martingales) Soit une submartingale (Y, \mathcal{F}) .

On suppose que $\mathbb{E}[Y_n^+] \leq M$ pour un certain M et pour tout n . Alors il existe une variable Y_∞ tel que $Y_n \rightarrow Y_\infty$ presque sûrement. De plus,

1.

$$Y_\infty$$

est d'espérance finie si

$$\mathbb{E}|Y_0| < +\infty$$

2.

$$Y_n \rightarrow Y_\infty$$

si la séquence $(Y_n, n \geq 0)$ est uniformément intégrable.

4.5 Stopping Times

4.5.1 Motivations

Le principe des *stopping times* repose sur le fait qu'à chaque instant, on a l'information qu'on peut ou non prendre une décision. C'est-à-dire que si on considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et une filtration $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \dots)$, on attribue à \mathcal{F}_n l'information disponible à l'instant n . Plus précisément, la plus petite σ -algèbre (au sens de l'inclusion) contenant l'information disponible jusqu'au temps n . Les valeurs prises par cette variable aléatoire sont considérées comme des moments d'intérêt.

4.5.2 Définitions

Définition 8 *Un temps d'arrêt T est une variable aléatoire définie par rapport à une filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ si,*

$$\forall n, \{T = n\} \in \mathcal{F} \text{ ou encore } \{T \leq n\} \in \mathcal{F} \quad (18)$$

4.5.3 Martingales locale

Définition 9 *Un processus stochastique X_t est une **martingale locale** si il existe une suite de temps d'arrêts (τ_n) telle que $\mathbb{P}(\lim \tau_n = \infty) = 1$ and $X_{t \wedge \tau_n}$ est une martingale pour tout n .*

4.6 Théorème de Girsanov

le théorème de Girsanov indique comment un processus stochastique change si l'on change de mesure.

Il donne la manière de passer de la probabilité historique qui décrit la probabilité qu'un **actif sous-jacent** (comme le prix d'une action ou un taux d'intérêt) prenne dans le futur une valeur donnée à la probabilité risque neutre qui est un outil très utile pour évaluer la valeur d'un dérivé du sous-jacent.

Soit X_t une martingale locale continue par rapport à une filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfaisant les conditions usuelles. On définit l'exponentielle stochastique Z_t

$$Z_t = \exp \left(\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right) \quad (19)$$

On a montré que si θ était déterministe constant alors le processus Z_t est alors une martingale. On peut définir une mesure \mathbb{Q} équivalente à la restriction de la mesure \mathbb{P} à \mathcal{F}_t à partir de sa densité de Radon-Nikodym.

$$d\mathbb{Q} = Z_t d\mathbb{P} \quad (20)$$

Théorème 4.5 (Théorème de Girsanov) *Si on a les conditions suivantes*

- ❖ Z_t est une vraie martingale, la famille Q_t est induite par une mesure Q définie sur toute la tribu $\mathcal{F}_t : Q|_{\mathcal{F}_t} = Q_t$.
- ❖ W est un brownien sous \mathbb{P}

alors le processus

$$\tilde{W}_t = W_t - \int_0^t \theta_s ds \quad (21)$$

est un brownien sous \mathbb{Q} .

Ce théorème peut être utilisé pour trouver l'unique probabilité risque neutre dans le modèle de Black-Scholes. Ainsi, si un actif suit le processus de diffusion vérifiant :

$$\frac{dS}{S} = \mu dt + \sigma dW_t \quad (22)$$

Où W_t est un P -mouvement Brownien, en effectuant le changement de probabilité suivant :

$$\frac{dQ_t}{dP|_{\mathcal{F}_t}} = \exp \left(\int_0^t \frac{r - \mu}{\sigma} dW_s - \frac{1}{2} \left(\frac{r - \mu}{\sigma} \right)^2 ds \right) \quad (23)$$

Le note $\theta = \frac{r - \mu}{\sigma}$, c'est le **market price of asset risk**. On obtient une diffusion vérifiant :

$$\frac{dS}{S} = r dt + \sigma d\tilde{W}_t \quad (24)$$

Où \tilde{W}_t est un \mathbb{Q} -mouvement brownien. Si on note \tilde{S} la valeur actualisée de S , on a :

$$\frac{d\tilde{S}}{\tilde{S}} = \sigma d\tilde{W}_t \quad (25)$$

Sous la probabilité \mathbb{Q} la valeur de notre actif réactualisée est une martingale.

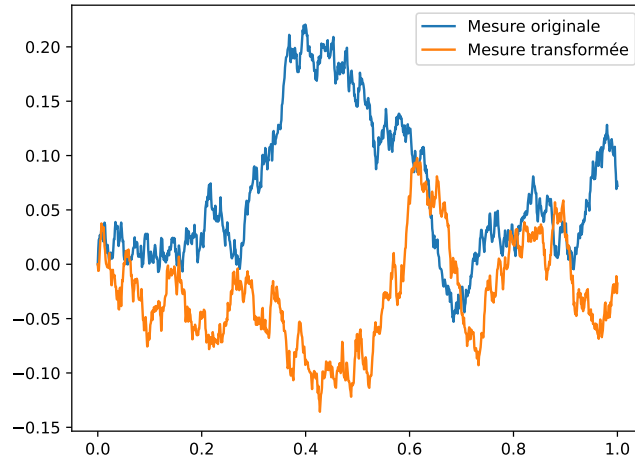


FIGURE 3 – Processus stochastiques sous la mesure transformée de Girsanov

Le premier processus X est généré selon la mesure de probabilité originale, en utilisant la méthode d'Euler pour simuler l'évolution du processus. Le deuxième

processus Y est généré en utilisant le théorème de Girsanov, en ajoutant un drift proportionnel à la différence entre la mesure transformée et la mesure originale. Enfin, les deux processus sont affichés sur un graphique, avec la fonction `plot` de `matplotlib`. Le graphique montre la différence entre les deux processus, qui est plus grande aux endroits où la mesure transformée est différente de la mesure originale.

4.7 Formule de Feynman–Kac

Cette formule relie la solution d'une équation aux dérivées partielles à une espérance conditionnelle impliquant un processus stochastique.

Théorème 4.6 (Formule de Feynman–Kac) *Soit f une fonction réelle bornée, Ψ une fonction de classe \mathcal{C}^2 à support compact K dans \mathbb{R} et q une fonction réelle bornée. Avec X_t vérifiant*

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$$

Alors pour $0 \leq t \leq T$, la formule de Feynman–Kac est

$$v(t, x) = \mathbb{E} \left[\int_t^T f(X_s) e^{-\int_t^s q(X_u) du} ds + e^{-\int_t^T q(X_s) ds} \Psi(X_T) | X_t = x \right] \quad (26)$$

et v est solution bornée unique de la SDE avec la condition $v(T, x) = \Psi(x)$

$$\frac{\partial v}{\partial t} + \frac{\sigma^2(x)}{2} \frac{\partial^2 v}{\partial x^2} + \mu(x) \frac{\partial v}{\partial x} - q(x)v(t, x) + f(t, x) = 0 \quad (27)$$

4.8 Quelques utilisations des martingales

4.8.1 Ruine d'un joueur

5 Processus de Lévy

5.1 Définition

Un processus stochastique $(X_t)_{t \geq 0}$ est dit *de Lévy* si il vérifie les conditions suivantes:

- (i) $X_0 = 0$ ps.
- (ii) (Accroissement indépendants) Pour tout $s < t$, l'accroissement $X_t - X_s$ est indépendant du processus $(X_u)_{0 \leq u \leq s}$.
- (iii) (Accroissement stationnaires) Pour tout $s < t$, l'accroissement $X_t - X_s$ est égal en loi au processus (X_{t-s}) .
- (iv) $t \mapsto X_t$ est continue ps et limitée à gauche

Remarque. On dit que $t \mapsto X_t$ est limitée à gauche lorsque si le processus X_t est discontinu en un certain t_0 alors sa valeur augmente instantanément mais ne peut pas diminuer.

Ils sont utilisés pour modéliser les variations des prix d'actifs financiers, comme les actions et les devises. Les sauts dans les prix, comme les krachs boursiers, peuvent être bien capturés par ces processus.

5.2 Exemple : les processus de Poisson

6 Notions de finance

6.1 Les securities

les titres financiers ou *securities* font référence à des instruments financiers négociables qui représentent une forme de propriété ou de créance. Les securities sont utilisées pour lever des fonds et pour investir dans des entreprises ou des gouvernements.

6.1.1 Les obligations (*bonds*)

<https://www.forbes.com/advisor/investing/what-is-a-bond/>

1. **Corporate bonds.** Investors who purchase corporate bonds become creditors of the company and receive periodic interest payments (coupon) and the return of the principal amount at the bond's maturity date.
2. **Treasury bonds.** Government bonds, often referred to as treasuries, are issued by governments to finance their operations. They are generally considered low-risk investments.
3. **Municipal bonds.** are issued by state and local governments or agencies to fund public projects such as infrastructure improvements or schools. They often offer tax advantages for investors.

On définit pour les obligations plusieurs quantités.

1. **Maturity.** La maturité désigne la date à laquelle l'obligation va expirer. i.e. la date à laquelle le *bond issuer* doit retourner l'argent au *bond investor*.
2. **Face Value, Par Value.** La valeur nominale est la valeur qui devra être remboursée à maturité. Généralement, on a $FV = \$1000$.
3. **Coupon.** Le coupon est le *rate of interest* que le bond issuer va payer à l'investisseur. Par exemple, un bond qui a 3% de coupon sur un FV de \$1000, le issuer va payer 3% of \$1000 = \$30 par an jusqu'à maturité. Lorsque le coupon est fixe on parle de **Fixed Income (FI)** et lorsqu'il est variable on parle de **Floating Rate Note (FRN)** ou encore *floater*
4. **Yield, rate of return.** le rendement d'un bond est variable, contrairement au FV ,

5. **Price.** La plupart des bonds sont tradés après leur émission, on considère deux prices pour les bons, *bid* et *ask*, le premier pour le **prix maximal que peut mettre un buyer**, le second pour le **prix minimal que le seller peut céder** le bond.
6. **Duration.** La duration d'une obligation est la durée de vie moyenne de ses flux financiers pondérée par leur valeur actualisée. **Plus la duration est élevée, plus le risque est grand.** Elle est fonction de toutes les autres quantités citées plus haut.

6.1.2 Les actions (*equity, stock*)

Investment growth potential, **dividends** : many companies pay dividends to investors corresponding to a portion of their profits distributed to shareholders. **Diversification** : By investing in stocks of different companies operating in various sectors, you reduce concentration risk. **Liquidity** : Stocks are typically liquid, meaning you can buy or sell them relatively easily on stock markets. This liquidity provides flexibility to adjust your portfolio based on market conditions. In summary, equities are one of the primary asset classes in financial markets, representing partial ownership in a company. They offer investment opportunities for those looking to participate in potential company growth and dividend income while being aware of the risks associated with market volatility. Investors should diversify their portfolios and have a long-term investment strategy to make the most of stocks.

6.2 Les produits dérivés

Un produit dérivé est un contrat financier dont la valeur dépend de l'évolution d'un actif sous-jacent tel qu'une *action*, une devise, une matière première, un indice boursier, une *obligation* ou un *taux d'intérêt*. Les produits dérivés permettent aux investisseurs de spéculer sur l'évolution future des prix ou des taux, d'assurer leurs positions existantes contre les risques de fluctuations des marchés, ou encore de se protéger contre les pertes potentielles.

Il existe différents types de produits dérivés, **notamment les contrats à terme, les options, les swaps et les contrats d'échange.**

- ❖ Les **contrats à terme** sont des accords d'achat ou de vente d'un actif à une date future à un prix déterminé à l'avance.
- ❖ Les **options** offrent à l'acheteur le droit, mais pas l'obligation, d'acheter ou de vendre un actif sous-jacent à un prix déterminé à l'avance, à une date donnée ou avant celle-ci.
- ❖ Les **swaps** sont des contrats d'échange d'actifs ou de flux de trésorerie entre deux parties.
- ❖ Les **contrats d'échange** permettent également aux parties d'échanger des flux de trésorerie en fonction de différents taux d'intérêt.

Les produits dérivés peuvent être utilisés pour diverses finalités, notamment pour couvrir les risques de marché, pour générer des rendements, pour se protéger

contre l'inflation ou pour se positionner sur des tendances de marché spécifiques. Toutefois, ils sont également associés à des risques de pertes potentiellement importantes et nécessitent une connaissance approfondie des marchés financiers et des instruments financiers pour être utilisés de manière efficace.

6.2.1 Les options

Les options sont des produits financiers dérivés donnant le droit (et non l'obligation) d'acheter (on dit alors **call**) ou de vendre (on dit alors **put**) une quantité d'actifs sous-jacents pendant une période et à un prix connu à l'avance. On paye une prime (prix d'exercice de l'option) pour acquérir ce droit.

On distingue les options européennes (on peut exercer son droit uniquement à la fin du contrat) des options américaines (on peut exercer son droit pendant toute la durée du contrat). **Variation du prix de l'option.** La prime varie

en fonction du sous-jacent (action, indice, matière première, devise, etc) mais également en fonction de l'offre et la demande. Il est donc composé de deux termes, la valeur *intrinsèque* et la valeur *temporelle*. (Source AMF)

6.2.2 Les futures

Le future (contrat à terme en français) est un contrat par lequel deux parties s'engagent à acheter (ou vendre) **une quantité déterminée d'un actif sous-jacent** (une action ou un indice boursier par exemple), à une date d'échéance et à un prix convenus à l'avance. Il permet d'anticiper les variations futures d'un actif sous-jacent et peut donc servir à couvrir un portefeuille contre les fluctuations à venir du marché. Il permet aussi de dynamiser les performances de son portefeuille. Le future constitue un engagement ferme, il doit être exécuté à sa date d'échéance par ses contreparties : l'acheteur du future doit acheter l'actif sous-jacent au prix convenu et le vendeur doit livrer l'actif.

Fonctionnement. Au moment de l'achat (On anticipe une hausse de l'actif sous-jacent) ou de la vente (On anticipe la baisse de l'actif sous-jacent) de future, un dépôt de garantie est requis. C'est une caution qui permet d'assurer la bonne fin des opérations en cas de défaillance d'une ou des parties. Puis, chaque jour, après la clôture du marché, vous recevez des notifications d'appels de marge. Ils correspondent aux profits ou pertes de votre position et dépendent de la réalisation (ou non) de votre anticipation. Ainsi, en cas de bonne anticipation, vous êtes bénéficiaire de l'appel de marge. Dans le cas contraire, vous êtes redevable de l'appel de marge. Dans l'hypothèse où le solde de votre compte ne suffit pas pour faire face à l'appel de marge, votre position peut être soldée et le dépôt de garantie servir à apurer l'appel de marge.

6.2.3 Les forwards

Un contrat forward est un contrat à terme **de gré à gré (Over The Counter – OTC)**, c'est-à-dire un accord entre un acheteur et un vendeur pour négocier un actif, *habituellement une devise*, à un prix et une date fixés d'un commun accord. Contrairement aux contrats futures, les contrats à terme sont des accords privés entre l'acheteur et le vendeur et, à ce titre, **ils ne sont pas négociés sur les marchés centralisés**, mais plutôt considérés comme faisant partie du marché

OTC. Cela rend les contrats forward plus risqués que les contrats futures. Les principaux facteurs qui déterminent le prix d'un contrat forward sont la valeur marchande de l'actif et le moment où le contrat prendra fin, ce qui est influencé par les taux swap.

6.2.4 Les swaps

Un swap est un échange entre deux parties, une sortie de crédit croisé consenti pour une période fixée dès le départ, de quelques jours à quelques mois ou années. Cet échange permet aux contreparties d'échanger une série de flux futurs. Outre la nature du flux échangé, le contrat swap doit spécifier plusieurs éléments : montant nominal, date de départ du contrat, durée, valeur du taux fixe, base de calcul et échéancier des versements, référence du taux variable et échéancier des versements, etc.

Fonctionnement. Les swaps ont pour vocation initiale de réduire l'exposition au risque d'une entreprise ou d'un individu. Aucune transaction n'est effectuée sur le capital : ce sont uniquement les flux d'intérêts qui sont échangés (swapés). Dans le cadre d'une opération de couverture, ils permettent souvent d'échanger un produit financier volatil contre un produit financier plus stable. Dans le cas d'un swap de taux, un taux variable pourra par exemple être échangé contre un taux fixe.

Principaux types de swap: On recense de multiples formes de swap. Parmi les plus fréquents :

- ❖ Les **swaps de taux** font figure de poids lourds du marché. Ils sont notamment utilisés par des entreprises endettées à taux fixe qui estiment que les taux vont baisser. Elle trouvent donc intérêt à échanger un taux fixe contre un taux variable pour profiter de cette baisse et inversement (pour la contrepartie).
- ❖ Les **Credit Default Swaps (CDS)** procurent une assurance contre un possible défaut, en contrepartie du versement périodique d'intérêt (une sorte d'assurance) jusqu'au terme de l'échange.
- ❖ Les **swaps de matières premières** permettent de troquer un prix fixe (par exemple celui du soja) contre un prix variable, le plus souvent calculé comme la moyenne d'un indice sur une période future.
- ❖ Les **swaps de devises** permettent d'échanger les intérêts et la valeur d'un sous-jacent dans une devise contre sa valeur dans une autre devise.

6.2.5 Tableau récapitulatif

On peut résumer

6.3 Les indices

6.3.1 VIX volatility index

Le VIX est un indicateur de volatilité établi tous les jours par le Chicago Board Options Exchange (CBOE). Il est calculé en prenant la moyenne des volatilités

UNDERLYING	CONTRACT TYPES				
	Exchange-traded futures	Exchange-traded options	OTC swap	OTC forward	OTC option
Equity	DJIA Index future Single-stock future	Option on DJIA Index future Single-share option	Equity swap	Back-to-back Repurchase agreement	Stock option Warrant Turbo warrant
Interest rate	Eurodollar future Euribor future	Option on Eurodollar future Option on Euribor future	Interest rate swap	Forward rate agreement	Interest rate cap and floor Swaption Basis swap Bond option
Credit	Bond future	Option on Bond future	Credit default swap Total return swap	Repurchase agreement	Credit default option
Foreign exchange	Currency future	Option on currency future	Currency swap	Currency forward	Currency option
Commodity	WTI crude oil futures	Weather derivative	Commodity swap	Iron ore forward contract	Gold option

FIGURE 4 – derivatives

annuelles sur les calls et puts sur indice S&P500. On a introduit aussi un indice de *vol de vol* VVIX qui mesure la volatilité du VIX.

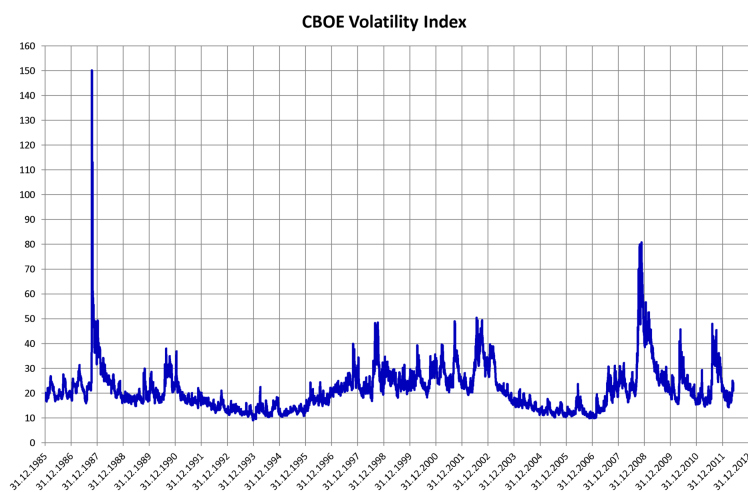


FIGURE 5 – Evolution du VIX de 1985 a 2012 (Wikipedia)

Des indices de vol sur d'autres indices existent aussi tels que le VXN (indice de vol sur NASDAQ), le VXD (Dow Jones) ou encore le VCAC sur le CAC40

7 Vanilla Option Pricing

7.1 Stratégie simple – Achat ou vente d'option

Prenons le cas d'un achat de Call au prix d'exercice K . Si l'option est exercée en T , l'acheteur du Call achète le sous-jacent au prix K , et peut le revendre au prix du jour S . L'acheteur n'exercera l'option que si cela lui procure un gain, ainsi son gain peut s'écrire $\max(S_T - K, 0)$. A ce gain il faut soustraire la prime p qui a été payée pour acquérir l'option, le profil des gains de l'acheteur de l'option devient donc

$$R_{call} = \max(S_T - K, 0) \quad (28)$$

L'acheteur d'une option d'achat a un risque de perte limité (le montant de la prime). Il réalise des profits dès que la valeur du sous-jacent est supérieure ou égale au prix d'exercice + la prime (appelé point mort ou seuil critique). Pour la vente d'un call, le résultat sera $R_{call} = -\max(S - K, 0) - p$. De manière similaire, le résultat pour l'acheteur d'un put sera $R_{put} = \max(0, K - S) - p$ et, symétriquement, le résultat pour le vendeur d'un put sera $R_{put} = -\max(0, K - S) + p$.

On peut voir l'évolution du prix de l'option pour les quatres scénarios précédents dans les graphes suivants:

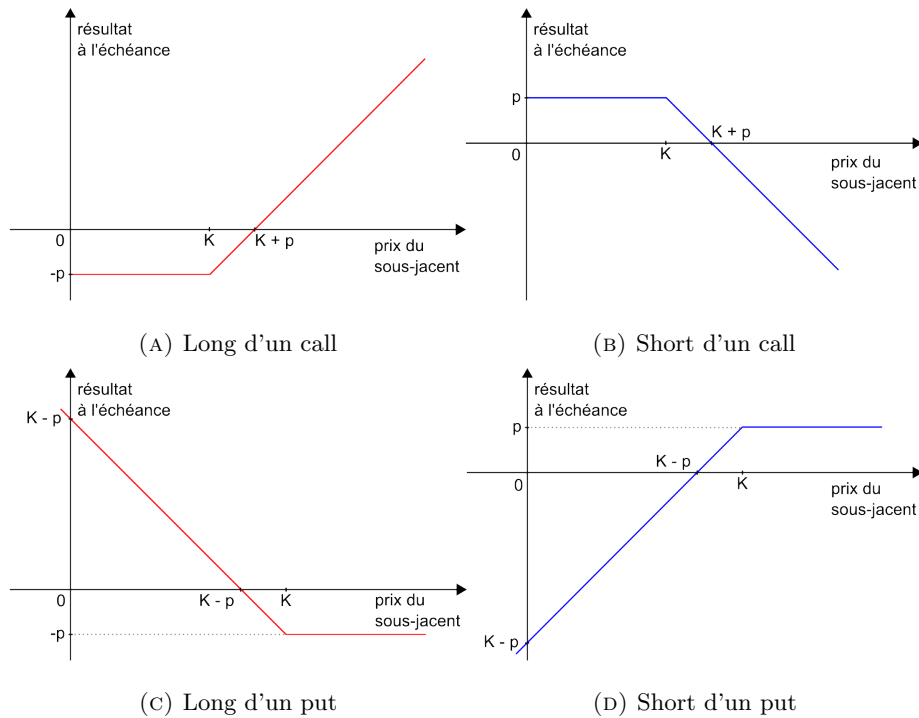


FIGURE 6 – Evolution du prix de l'option

7.2 Modèle de Black-Scholes

Ce modèle permet premièrement de pricer les options européennes (European Call Option). On considère deux actifs : une option (bond/money market account) dont la valeur évolue continuellement avec un intérêt fixé r (rate) et une action (stock) dont le prix par unité (price per unit) est modélisé par un processus stochastique $(S_t, t \geq 0)$.

On note M_t le prix d'une unité de bond à l'instant t en considérant $M_0 = 1$ on a $dM_t = rM_t dt$ soit $M_t = e^{rt}$. La première assertion du modèle de Black-Scholes est que S satisfait l'équation différentielle stochastique suivante:

$$dS_t = S_t(\mu dt + \sigma dW_t) \text{ soit } S_t = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \quad (29)$$

où W représente un mouvement Brownien standard. S représente donc un mouvement Brownien géométrique de paramètres μ et σ (qui est nommée la *volatilité* du prix).

7.3 Modèle de Black-Scholes-Merton

7.3.1 Hypothèses

Suite au krach de 1987, on trouve ci-après un modèle qui permet de prendre en compte les taux et les volatilités non constantes et qui permet de pricer des options qui payent des dividendes.

On cherche à pricer un contrat européen, i.e. on cherche la fonction $P(t, X_t)$.

Hypothèse de non-arbitrage. on suppose qu'il n'y a pas d'opportunité d'arbitrage. i.e. il existe des processus (a_t) et (b_t) tels que

$$a_t X_t + b_t e^{rt} = P(t, X_t) \text{ pour } 0 \leq t \leq T$$

7.3.2 Equation différentielle stochastique

La propriété de self-financing nous donne

$$d(a_t X_t + b_t e^{rt}) = a_t dX_t + r b_t e^{rt} dt$$

On a alors grace au lemme d'Ito,

$$d(P(X_t, t)) = \left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2} \frac{\partial^2 P}{\partial x^2} \sigma_t^2 \right) dt + \sigma \frac{\partial P}{\partial x} X_t dW_t$$

donc en ré-injectant les deux équations,

$$\left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2} \frac{\partial^2 P}{\partial x^2} \sigma_t^2 \right) dt + \sigma \frac{\partial P}{\partial x} X_t dW_t = a_t dX_t + r b_t e^{rt} dt$$

Soit,

$$\left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2} \frac{\partial^2 P}{\partial x^2} \sigma_t^2 \right) dt + \sigma \frac{\partial P}{\partial x} X_t dW_t = a_t \mu dt + r b_t e^{rt} dt + a_t \sigma X_t dW_t$$

On peut alors identifier les termes en dW_t :

$$a_t = \frac{\partial P}{\partial x}(t, X_t) \quad (30)$$

On a alors avec la condition de non-arbitrage

$$b_t = e^{-rt}(P(t, X_t) - a_t X_t) = e^{-rt}\left(P(t, X_t) - \frac{\partial P}{\partial x} X_t\right) \quad (31)$$

en ré-injectant dans l'équation de départ, on a en divisant par dt :

$$\boxed{\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2} + r\left(X_t \frac{\partial P}{\partial x} - P\right) = 0} \quad (32)$$

C'est l'équation de Black-Scholes, qui est vérifiée pour chaque stock price X_t .

Exemple. Pour pricer un European call option de strike K et de maturité T , on a alors

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + r\left(S \frac{\partial C}{\partial S} - C\right) = 0$$

Avec les conditions aux limites:

- ❖ *Expiration de l'option.* à $t = T$, $C(S, T) = (S - K)_+$ car le prix de l'option est égal à la valeur intrinsèque, qui est la différence entre le prix de l'actif sous-jacent S et le prix d'exercice K , si elle est positive, sinon, elle est nulle.
- ❖ *Limite à l'infini.* Lorsque S tend vers l'infini: $C(S, t) \rightarrow S$ car **le prix de l'option d'achat augmente à mesure que le prix de l'actif sous-jacent augmente.**
- ❖ *Limite en 0.* Lorsque S tend vers zéro: $C(S, t) \rightarrow 0$ car le prix de l'option d'achat **ne peut pas être négatif.**

Montrons que la solution de cette équation (prix d'un call européen sans dividendes) est:

$$C(S, t) = e^{-r(T-t)} N(d_2) \text{ avec } d_2 = \frac{1}{\sigma\sqrt{T-t}} \quad (33)$$

7.3.3 Parité call-put

entre un put et un call avec les memes maturity et strike price. On a tout d'abord,

$$(S_T - K)_+ + (K - S_T)_- = S_T - K$$

Il suffit alors d'appliquer des deux cotés de l'équation $e^{-r(T-t)}\mathbb{E}^\mathbb{Q}$ et on obtient

$$C(t, S_t) - P(t, S_t) = S_t - Ke^{-r(T-t)} \quad (34)$$

Car S_t est une martingale.

7.3.4 Dividendes

On considère que lors d'une période de temps de $[t, t + dt]$ le paiement des dividendes s'écrit $qS_t dt$ avec q constant. Le sous-jacent suit alors un mouvement brownien géométrique de dynamique:

$$dS_t = (r - q)S_t dt + \sigma S_t dW_t^{\mathbb{Q}} \quad (35)$$

On forme alors un portefeuille comme suit:

- ❖ On long 1 call C
- ❖ On short Δ stock

On a alors

$$d\Pi_t = dC - \Delta dS_t - q\Delta S_t dt \quad (36)$$

Avec la formule d'Ito on obtient:

$$dV_t = \left(\frac{\partial V}{\partial t} + (r - q)S_t \frac{\partial V}{\partial S} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma_t^2 \right) dt + \sigma \frac{\partial V}{\partial S} S_t dW_t^{\mathbb{Q}} \quad (37)$$

L'absence d'opportunité d'arbitrage donne que la valeur du portefeuille sans-risque doit être la même que celle du portefeuille sans risque zéro-coupon i.e. $d\Pi_t = r\Pi_t dt$. En identifiant les termes en dt et en $dW_t^{\mathbb{Q}}$, la forme modifiée de l'équation de Black-Scholes pour une option d'achat européenne avec des dividendes est la suivante :

$$\boxed{\frac{\partial C}{\partial t} + (r - q)S \frac{\partial C}{\partial x} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial x^2} - rC = 0}$$

Le modèle de Black-Scholes nous permet d'écrire le prix d'un call C et celui d'un put P :

$$C(S, T) = e^{-qT} S_0 N(d_1) - e^{-rT} K N(d_2) \text{ et } P(S, T) = e^{-rT} K N(-d_2) - e^{-qT} S_0 N(-d_1) \quad (38)$$

Dans ces équations,

- ❖ S représente le prix actuel de l'actif sous-jacent, S_0 le prix initial
- ❖ T représente la date d'échéance de l'option,
- ❖ r représente le taux d'intérêt sans risque,
- ❖ q représente le rendement attendu de l'actif sous-jacent,
- ❖ K représente le prix d'exercice de l'option,
- ❖ N est la fonction de répartition de la loi normale standard, on note N (ou alors Φ) la fonction de répartition de la loi Normale centrée réduite:

$$N(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad \text{et} \quad N'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

- ❖ d_1 et d_2 sont des variables qui dépendent de ces différents paramètres. sont utilisés pour calculer la probabilité que l'option soit exercée (ou non) à l'expiration. Plus précisément, d_1 mesure l'écart entre le prix actuel de l'actif sous-jacent et le prix d'exercice de l'option, ajusté pour prendre en compte les dividendes et la volatilité. D'autre part, d_2 mesure la même chose, mais sans prendre en compte les dividendes.

$$d_1 = \frac{1}{\sigma\sqrt{T}} \ln \left[\left(\frac{S}{K} \right) + \left(r - q + \frac{1}{2}\sigma^2 \right) T \right] \quad \text{et} \quad d_2 = d_1 - \sigma\sqrt{T}$$

Et on peut calculer le prix le prix à terme (ou le prix modifié en avance) de l'actif sous-jacent F :

$$F = e^{(r-q)T} S_0 \quad (39)$$

Ce modèle permet de calculer le prix d'une option en supposant que les mouvements du prix de l'actif sous-jacent suivent une marche aléatoire géométrique avec une volatilité constante. En d'autres termes, *elle suppose que les rendements futurs de l'actif sous-jacent sont distribués de manière normale et que les investisseurs sont rationnels et ont des attentes homogènes concernant les rendements futurs*. Cette formule est largement utilisée dans la finance pour évaluer le prix des options, bien qu'elle ait ses limites et ses critiques.

7.4 Le modèle binomial (Cox–Ross–Rubinstein)

7.4.1 Présentation

Le modèle de Cox-Ross-Rubinstein (CRR) est une méthode de modélisation de la dynamique des prix des options d'achat *call* et de vente *put* dans le cadre de la théorie des options.

Le modèle CRR utilise une approche discrète pour décrire l'évolution des prix d'un actif sous-jacent. Il suppose que l'actif sous-jacent peut prendre deux valeurs distinctes à chaque période de temps. La variation de l'actif sous-jacent est représentée par un arbre binomial qui permet de calculer le prix de l'option à chaque période.

La formule CRR est particulièrement utile pour évaluer les *options européennes*, qui peuvent être exercées uniquement à une date d'expiration spécifiée. Elle permet également de calculer les valeurs d'autres options complexes, comme les *options américaines*, qui peuvent être exercées à tout moment avant la date d'expiration.

7.4.2 Construction de l'arbre

On part de la date à laquelle on souhaite valoriser l'option on finit à la date de d'expiration de l'option. A chaque étape, le sous-jacent augmente *up* ou diminue *down* en fonction d'un facteur spécifique d ou u (avec $u \geq 1$ et $0 < d \leq 1$). Si S est le prix actuel, alors le prix au pas suivant sera $S_u = S \times u$ ou $S_d = S \times d$. Les facteurs spécifiques sont calculés en prenant en compte la *volatilité* σ du sous-jacent, le pas de temps t (en année, selon la convention du nombre de jours

du sous-jacent). En considérant que $V(\log S) = \sigma^2 t$, on a :

$$u = e^{\sigma\sqrt{t}} \quad \text{et} \quad d = e^{-\sigma\sqrt{t}} = \frac{1}{u} \quad (40)$$

La méthode CRR assure le fait que l'arbre est recombinant, i.e. si les mouvements du sous-jacent sont d'abord une augmentation puis une diminution, le prix sera le même que si les mouvements suivis avaient été d'abord une diminution puis une augmentation. Dans ce cas là, les deux branches fusionnent et se recombinent. Cette propriété réduit donc le nombre de nœuds de l'arbre et par conséquent accélère le calcul du prix de l'option. Cette propriété permet également de calculer la valeur de l'actif sous-jacent à chaque nœud directement par des formules, sans passer par la création d'un arbre. La valeur d'un nœud est donc $S_n = S_0 \times u^{N_u - N_d}$

7.4.3 Calcul de la valeur de l'option sur les noeuds

Calcul pour chaque noeud final. La valeur de l'option est sa valeur intrinsèque i.e.

- ❖ $\max[(S - K), 0]$ pour un call
- ❖ $\max[(K - S), 0]$ pour un put

Où K est le strike (prix d'exercice de l'option) et S est le spot (cours au comptant) du sous-jacent.

Calcul pour les noeuds antérieurs. La valeur de l'option pour chaque noeud est trouvée en utilisant la valeur du noeud précédent, en remontant vers le premier noeud de l'arbre (i.e. la valeur de l'option à la date où l'on souhaite la valoriser). En prenant en compte l'hypothèse de neutralité du risque, le prix, aujourd'hui, d'un instrument dérivé est égal à la valeur de ses bénéfices futurs actualisés en fonction du taux sans risque. Par conséquent, la valeur attendue est calculée en utilisant la **valeur de l'option lors des deux derniers nœuds** (appelés ici option *up* et option *down*) pondérés par leurs probabilités respectives. Soient p la probabilité d'une variation à la hausse de la valeur du sous-jacent et $(1 - p)$ la probabilité d'une variation à la baisse. La valeur de l'option est ensuite actualisée avec r le taux sans risque correspondant à la durée de l'option.

$$C_{t-\Delta t, i} = e^{-r\Delta t} \left(p C_{t, i+1} + (1 - p) C_{t, i-1} \right) \quad (41)$$

Où $C_{t, i}$ représente la valeur de l'option pour le i -ème noeud au temps t , p représente la probabilité de la loi binomiale qui simule le mouvement brownien géométrique de l'action sous-jacente avec les paramètres r et σ et q est le rendement du dividende du sous-jacent correspondant à la durée de vie de l'option i.e.

$$p = \frac{e^{(r-q)\Delta t} - d}{u - d}$$

Pour qu'on ait $p < 1$, il faut avoir $\Delta t < \frac{\sigma^2}{(r-q)^2}$.

$C_{t-\Delta t, i}$ représente le juste prix du dérivé à un point donné dans le temps (c'est-à-dire chaque nœud), étant donné l'évolution du prix du sous-jacent à ce point. C'est la valeur de l'option à ce point – opposée à la valeur d'exercice.

7.4.4 Exemple d'utilisation

L'exemple d'utilisation du code affiche l'arbre de prix pour un actif sous-jacent initial de 100, un prix d'exercice de 110, un taux d'intérêt sans risque de 5%, une durée de vie de l'option de 1 an, 3 périodes de temps dans l'arbre et une volatilité de 20%.

```
import numpy as np

# Fonction pour calculer l'arbre de Cox-Ross-Rubinstein
def CRR_tree(S0, K, r, T, N, sigma):
    dt = T / N
    u = np.exp(sigma * np.sqrt(dt))
    d = 1 / u
    p = (np.exp(r * dt) - d) / (u - d)
    # Initialisation de l'arbre
    tree = np.zeros((N + 1, N + 1))
    for i in range(N + 1):
        for j in range(i + 1):
            tree[j, i] = S0 * (u ** (i - j)) * (d ** j)
    return tree

# Fonction pour afficher l'arbre de Cox-Ross-Rubinstein
def print_tree(tree):
    N = tree.shape[0] - 1
    for i in range(N + 1):
        print(' ' * (N - i), end='')
        for j in range(i + 1):
            print('{0:8.2f}'.format(tree[i - j, j]), end='')
        print()
```

Voici ce que retourne le code:

```
S0 = 100
K = 110
r = 0.05
T = 1
N = 3
sigma = 0.2

tree = CRR_tree(S0, K, r, T, N, sigma)
-----
[In] > print_tree(tree)
[Out]
      100.00
    0.00    112.24
  0.00    89.09  125.98
0.00    0.00  100.00  141.40
```

7.5 PDF du prix et formule de BREEDEN–LITZENBERGER

On cherche à calculer la densité de probabilité du prix. BREEDEN–LITZENBERGER (1978) ont montré le lien entre la dérivée de la formule de FEYNMANN–KAC et la densité du prix.

Théorème 7.1

8 Autres modèles de pricing

On cherche alors à introduire issus de Black-Scholes mais qui incluent des termes correctifs pour tenter de modéliser au mieux le comportement réels de la volatilité implicite observée sur le marché. On distingue alors deux types de modèles, ceux à volatilité "locale" et ceux à volatilité stochastique. *Les modèles à volatilité locale ont l'avantage de conserver l'aspect de marché complet car ceux-ci n'introduisent pas, à la différence des modèles à volatilité stochastique, de dynamique supplémentaire dans la modélisation. L'inconvénient vient du fait que les dynamiques des volatilités implicites observés dans le marché ne sont pas en adéquation avec les dynamiques prévues dans le modèle* [?] La complétude du marché signifie qu'on peut identifier chaque source d'aléa par un portefeuille de produit. Autrement dit on peut se couvrir ou isoler chaque source de risque. Cette complétude des marchés ne peut être conservée qu'en introduisant des produits comme le VIX ou les variances forward qui apportent une méthode supplémentaire de couverture

8.1 Evaluation de la volatilité

8.1.1 Modèle GARCH(1,1)

8.1.2 Modèle GARCH(1,1)

8.1.3 Volatilité Implicite

Dans le modèle de Black-Scholes, la volatilité est considérée comme constante. La volatilité implicite est alors définie pour refléter la perception du marché quant à la fluctuation future des prix d'un actif sous-jacent. Contrairement à la volatilité historique, qui se base sur les mouvements passés des prix, la volatilité implicite est déduite du prix actuel des options. Mathématiquement, la volatilité implicite ($\sigma_{\text{implicite}}$) est souvent déterminée en inversant la formule de prix de l'option. Pour une option d'achat européenne, par exemple, la formule de Black-Scholes est souvent utilisée :

$$C = S_0 N(d_1) - X e^{-rT} N(d_2) \quad (42)$$

où :

$$d_1 = \frac{\ln\left(\frac{S_0}{X}\right) + \left(r + \frac{\sigma_{\text{implicite}}^2}{2}\right) T}{\sigma_{\text{implicite}} \sqrt{T}}$$

$$d_2 = d_1 - \sigma_{\text{implicite}} \sqrt{T}$$

Dans ces équations, C est le prix de l'option d'achat, S_0 est le prix actuel de l'actif sous-jacent, X est le prix d'exercice de l'option, r est le taux d'intérêt sans risque, T est le temps restant jusqu'à l'expiration de l'option, et N est la fonction de distribution cumulative de la loi normale standard. Inverser cette

équation pour résoudre $\sigma_{\text{implicite}}$ donne une estimation de la volatilité implicite basée sur le prix actuel de l'option et d'autres paramètres du marché.

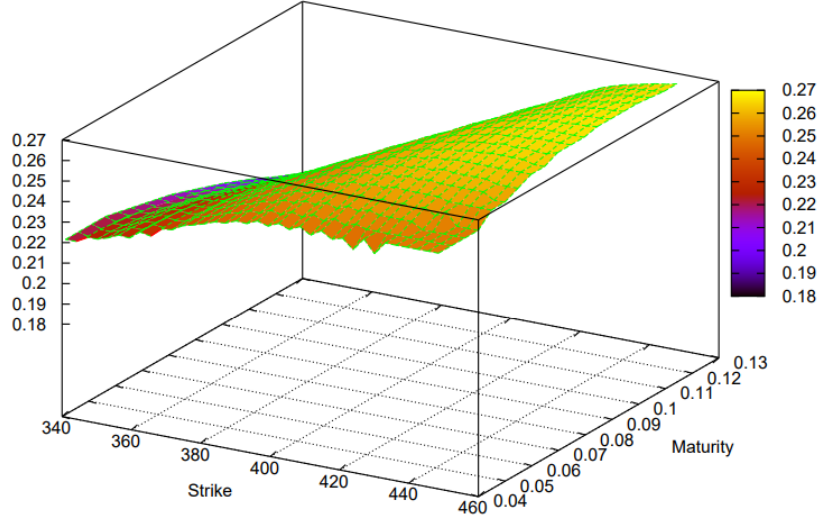


FIGURE 7 – The volatility surface

On constate sur les données de marché que la volatilité implicite ne peut être considérée comme constante. Elle dépend du strike et de la maturité. On peut alors supposer plusieurs dynamiques afin de modéliser au mieux la volatilité implicite face aux données de marché. On peut par exemple supposer qu'elle suit une dynamique stochastique propre, avec un brownien corrélé au brownien du sous-jacent. De plus, les phénomènes exceptionnels tels que les krachs ou les défauts font présenter des queues de distributions plus épaisses et imposent une distribution non gaussienne (kurtosis positif).

8.2 Modèle de volatilité locale

Dans les modèles à volatilité locale, la volatilité est une fonction non plus constante comme dans Black-Scholes, mais une fonction du temps et du spot S_t . La dynamique du sous-jacent devient alors

$$dS_t = \mu S_t dt + \sigma(t, S_t) dB_t$$

Dans les modèles de Black-Scholes, la fonction de volatilité est une fonction affine de S_t avec σ constant pour chaque option (chaque strike et chaque maturité). On distingue deux modèles de volatilité locale, la volatilité de Dupire, où la fonction de volatilité est explicite, et le modèle CEV (Constant Elasticity of Variance) où la fonction de volatilité n'est pas explicite.

8.2.1 Volatilité de Dupire

Dans ce modèle, la volatilité est construite afin de calibrer au mieux le smile (volatilité implicite en fonction du strike).

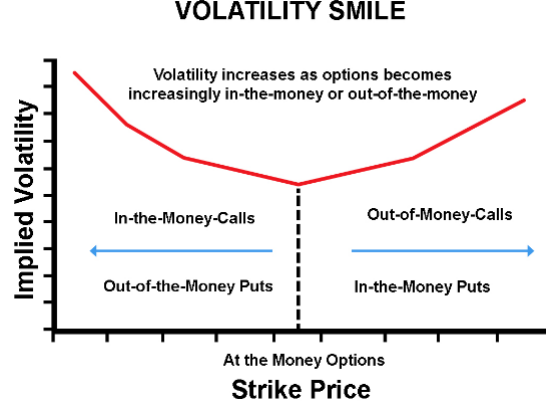


FIGURE 8 – Smile de la volatilité implicite

Afin de simplifier le modèle, on considère que les taux sont nuls et que donc la dynamique de S_t est

$$dS_t = \sigma(t, S_t)dB_t$$

On définit alors la fonction σ comme la **projection markovienne** de la dynamique de S_t :

$$\sigma(t, S_t) = \sqrt{\mathbb{E} \left[\frac{d\langle S_t \rangle}{dt} \middle| S_t = S \right]} \quad (43)$$

En écrivant le prix d'un call $C(T, K) = \mathbb{E}_Q[(S_T - K)_+]$ sous la forme intégrale et en dérivant par rapport à la maturité on obtient

$$\frac{\partial C(T, K)}{\partial T} = \int_K^{+\infty} (x - K)_+ \frac{\partial v(dx)}{\partial T}$$

Avec $v(dx)$ la densité de probabilité du prix S_t . On pose la fonction f qui associe à chaque pay-off $h(S_T)$ son prix $\mathbb{E}[h(S_T)|S_t = x] = f(x, t)$. On écrit alors l'espérance conditionnelle et on obtient:

$$\mathbb{E}[h(S_T)] = \mathbb{E} \left[\mathbb{E}[h(S_T)|S_T = x] \right] = \int f(x, t) v_t(x) dx$$

Le terme de gauche ne dépend pas de t , alors en dérivant l'équation par rapport à t il vient

$$\int \frac{\partial f(x, t)}{\partial t} v_t(x) dx + \int f(x, t) \frac{\partial v_t(x)}{\partial t} dx = 0$$

En appliquant la formule d'Ito à f on obtient

$$\frac{\partial f(x, t)}{\partial t} + \frac{1}{2} \sigma^2(x, t) x^2 \frac{\partial^2 f(x, t)}{\partial x^2} = 0$$

Après intégration par parties on peut réécrire cette relation en fonction de $f(x, t)$. On obtient alors que pour tout f issu de tout pay-off $h(S_T)$ on a :

$$\int \left(\frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2 x^2 v) - \frac{\partial v_t(x)}{\partial t} \right) f(x, t) dx = 0$$

Il vient alors la relation

$$\frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2 x^2 v) = \frac{\partial v_t(x)}{\partial t}$$

En injectant ce résultat dans (8.2.1) on obtient alors

$$\begin{aligned} \frac{\partial C(T, K)}{\partial T} &= \int_K^{+\infty} (x - K)_+ \frac{\partial v(dx)}{\partial T} \\ &= \frac{1}{2} \int_K^{+\infty} (x - K)_+ \frac{\partial^2}{\partial x^2} (\sigma^2 x^2 v) \\ &= \frac{1}{2} \sigma^2(T, K) K^2 v_T(K) \\ &= \frac{1}{2} \sigma^2(T, K) K^2 \frac{\partial^2 C(T, K)}{\partial K^2} \end{aligned}$$

On a supposé que $\sigma^2 x v$ admet une asymptote à l'infini qui est nulle afin de pouvoir calculer l'intégrale (???). On obtient alors la formule de volatilité de Dupire

Théorème 8.1 (Formule de volatilité de Dupire) *Dans le cadre du modèle de volatilité locale de Dupire, on a la relation entre la volatilité et le prix du call à une maturité T et un strike K donné dans le cas où le taux d'intérêt r_t est nul :*

$$\sigma^2(T, K) = \frac{2 \frac{\partial \mathbb{E}[(S_T - K)_+]}{\partial T}}{K^2 \frac{\partial^2 \mathbb{E}[(S_T - K)_+]}{\partial K^2}} \quad (44)$$

8.2.2 Constant Elasticity of Variance Model (modèle CEV)

Le modèle CEV, introduit dans [3], est un cas particulier de modèle à volatilité locale. Il suppose l'existence d'une forme particulière de volatilité et donc de dynamique du sous-jacent. L'intérêt de ce modèle est l'existence de formules approchées et de méthode de calibration qui en font un très bon modèle proxy. Un modèle proxy est un modèle de première approximation qu'on enrichit ensuite à l'aide de méthode de perturbation. La dynamique du modèle s'écrit :

$$dS_t = \sigma S_t^\gamma dW_t \quad (45)$$

σ représente la volatilité initiale et on peut jouer aussi sur le paramètre γ , qui est pris en générale entre 0 et 1.

8.3 Modèle de volatilité stochastique

8.3.1 Modèle de Heston

8.3.2 Stochastic Alpha, Beta, Rho model (modèle SABR)

8.3.3 Modèle de Bergomi

9 Dérivés de taux

Un **interest rate option** ou option sur taux d'intérêt, donne au *holder* le droit (mais pas l'obligation) donne le droit à l'acheteur d'emprunter un montant déterminé (Cap) ou d'en prêter un (Floor) à un taux d'intérêt fixé (taux d'intérêt d'exercice) pour une durée spécifique. On distingue plusieurs types:

- (i) **Cap option** Elle donne le droit au *holder* de recevoir un paiement si le taux d'intérêt dépasse un certain montant fixé à l'avance.
- (ii) **Floor option** Elle donne le droit au *holder* de recevoir un paiement si le taux d'intérêt passe en dessous un certain montant fixé à l'avance.
- (iii) **Collar option** Correspond à une combinaison d'achat de cap et de vente de floor (ou inversement) permettant d'encadrer le taux d'intérêt.

9.1 Zero-coupon bond

Le prix B d'un bond et le yield y sont reliés par la formule suivante. On suppose que le bond rémunère au détenteur un cash flow c_i au temps t_i pour $(0 \leq i \leq n)$

$$B = \sum_{i=1}^n c_i e^{-y t_i} \quad (46)$$

On définit alors la **duration** du bond par la relation:

$$D = \frac{1}{B} \sum_{i=1}^n c_i e^{-y t_i} = \sum_{i=1}^n t_i \left(\frac{c_i e^{-y t_i}}{B} \right) \quad (47)$$

Le prix du bond peut être aussi exprimé comme la somme de tous les futures coupons et de la valeur nominale, et le yield à maturité y_n

$$P = \sum_{i=1}^n \frac{c_i}{(1 + y_n)^i} + \frac{FV}{(1 + y_n)^n} \quad (48)$$

On peut essayer de trouver une relation entre la variation du prix du bond et celle du yield comme suit.

$$dP = \left[\sum_{i=1}^n \frac{-i c_i (1 + y_n)^{i-1}}{(1 + y_n)^{2i}} + \frac{n FV (1 + y_n)^{n-1}}{(1 + y_n)^{2n}} \right] dy_n = \frac{-D P dy_n}{1 + y_n} \quad (49)$$

On définit la *modified duration* D^* comme la duration divisée par $1 + y_n$. On obtient alors la relation

$$\frac{dP}{P} = -D^* dy_n$$

En supposant que le yield suit un modèle log-normal $dy_n = y_n \sigma_y dW_t$, On obtient alors la relation vérifiée par le prix du bond P ,

$$\frac{dP}{P} = -\sigma_y D^* y_n dW_t$$

D est l'**élasticité** du prix de l'obligation par rapport à l'interest rate.

$$D = -\frac{d(\log P)}{d(\log 1 + r)}$$

Pour évaluer le prix d'une obligation en pratique, on utilise la relation, dans le cas de coupon fixes (Fixed Income)

$$P = cFV \times \frac{1 - (1 + r)^{-n}}{r} + \frac{FV}{(1 + r)^n} \quad (50)$$

Ou

- ❖ Coupon = Par Value \times Coupon Payment
- ❖ r désigne le *rate of return*
- ❖ n désigne le nombre de *years to maturity* dans le cas de *annual coupon paying bonds*

9.2 Short rate models

Le SRM donne une dynamique pour le *interest rate* de la forme:

$$dr = \alpha(t, S_t)dt + \beta(t, S_t)dW_t$$

9.2.1 Modèle de Cox-Ingersoll-Ross

Un exemple de *short rate model* est le modèle de COX-INGERSOLL-ROSS model. Il est utilisé lors du pricing d'un *interest rate derivative*. La dynamique est de la forme:

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t$$

Le *drift* est composé de deux paramètres a et b qui peuvent être interprétés par a le *speed of adjustment* à la moyenne b . Pour éviter d'avoir des valeurs négative on a aussi la **Condition de Feller**

$$2ab \geq \sigma^2$$

9.2.2 Modèle de Hull-White

9.2.3 Modèle de Vasicek

Définition. Un processus d'Orstein-Uhlenbeck est un processus stochastique vérifiant l'équation différentielle stochastique

$$dX_t = -\theta(S_t - \mu)dt + \sigma dW_t \quad (51)$$

Où σ , μ et θ sont déterministes.

Pour résoudre cette équation, on peut appliquer le lemme d'Itô à la fonction $f(t, X_t) = X_t e^{\theta t}$. Il vient

$$df(t, X_t) = \theta X_t e^{\theta t} + e^{\theta t} dX_t = \theta \mu e^{\theta t} dt + \sigma e^{\theta t} dW_t \quad (52)$$

On obtient alors

$$X_t = X_0 + \mu \theta \int_0^t e^{\theta s} ds + \sigma \int_0^t e^{\theta s} dW_s \quad (53)$$

On peut montrer que

(i)

$$\mathbb{E}[X_t] = X_0 e^{-\theta t} + \mu(1 - e^{-\theta t})$$

(ii)

$$Var(X_t) = \frac{\sigma^2}{2\theta} e^{-2\theta t} (e^{2\theta t} - 1)$$

(iii)

$$Cov(X_t, X_s) = \frac{\sigma^2}{2\theta} e^{-\theta(s+t)} (e^{2\theta \min(s,t)} - 1)$$

9.3 Health–Jarrow–Morton framework

9.4 Market models

Le market model est utilisé pour les quantités observables telles que les **forward rates** ou encore les **swap rates**. Par exemple le modèle LIBOR (London Interbank Offered Rate) donne la dynamique pour le j -ème forward rate f_j :

$$\frac{df_j}{f_j} = \mu_j dt + \sigma_j dW_t^{(j)}$$

9.5 Markov functional models

10 Dérivé Action (EQD)

10.1 Les types de dérivés sur action

10.1.1 Les options et warrants

Un call warrant (put warrant) sur actions donne à son acquéreur le droit et non l'obligation d'acheter (de vendre) lesdites actions à un prix d'exercice convenu à l'avance et ceci à tout moment d'ici l'échéance prévue au contrat.

Mais les warrants sur actions se distinguent des options sur actions sur un certain nombre de points :

- ❖ Les warrants sont des **valeurs mobilières (securities)**, créées par un émetteur (**établissement financier**) pour être proposées à la vente aux investisseurs, alors que les options sont des **contrats (derivative)**, créés par une entreprise de marché (**Euronext par exemple**), s'agissant des options échangées sur un marché organisé ou par des établissements financiers s'agissant des options OTC. Les warrants figurent sur le même compte titres que les actions, tandis que la négociation d'options donne lieu à la création d'un compte distinct ;
- ❖ Leurs caractéristiques (échéances, prix d'exercice, quotités) sont choisies par l'émetteur alors que celles des options sont standardisées ;
- ❖ Les warrants sont de type **américain** (on peut les exercer à tout moment entre la date d'achat et la date d'échéance) alors que les options peuvent être de type **américain ou européen**.
- ❖ Un nombre illimité d'options peut être créé pour une même classe mais le nombre de warrants émis sur une maturité et un prix d'exercice donnés est en revanche limité ;
- ❖ Si les options peuvent être négociées à l'unité, il y a souvent une taille minimale de transaction sur les warrants. Exemple : si cette taille minimale est de 1000 unités, on ne peut acheter ce warrant que par multiple de 1000 ;
- ❖ La liquidité des warrants (proposition permanente de prix à l'achat et à la vente) est généralement assurée par le seul émetteur alors que plusieurs établissements « teneurs de marché » proposent généralement des prix dans le cas des options. **Lorsque tous les warrants déjà émis sont aux mains des investisseurs et que l'émetteur ne souhaite pas en émettre de nouveaux, il ne peut donc coter que des prix à l'achat (« bid only ») ;**

10.1.2 Les Turbos

L'élément essentiel du Turbo est la **barrière désactivante (knock out)** ou « seuil de sécurité ». Lorsque le cours de l'actif sur lequel est basé le Turbo atteint ce seuil, le Turbo est immédiatement « désactivé », c'est-à-dire qu'il est radié de la cotation. Il ne peut plus être échangé. Dans le cas des Turbos « classiques », il perd alors toute valeur. Les Turbos « classiques » ont une durée de vie fixée à l'avance : la maturité. Les Turbos Infinis / Illimités n'ont pas de date d'échéance, ils ont une durée de vie infinie.

10.1.3 Les Compo et Quanto

Ici, la devise utilisée dans l'option est différente de celle du sous-jacent. Le produit est donc exposé au risque de change. Le payoff pour ces deux produits est

$$QC = FX_0 \max(S_T - K, 0)$$

Ce qui distingue le compo du quanto est

10.1.4 Les futures contracts

Le future (contrat à terme en français) est un contrat par lequel deux parties s'engagent à acheter (ou vendre) **une quantité déterminée d'un action**, à une date d'échéance et à un prix convenus à l'avance. Il permet d'anticiper les variations futures d'un actif sous-jacent et peut donc servir à couvrir un portefeuille contre les fluctuations à venir du marché. Il permet aussi de dynamiser les performances de son portefeuille. Le future constitue un engagement ferme, il doit être exécuté à sa date d'échéance par ses contreparties : l'acheteur du future doit acheter l'actif sous-jacent au prix convenu et le vendeur doit livrer l'actif.

10.1.5 Les obligations convertibles

Une obligation convertible est un prêt qui, à son terme, peut être remboursé soit en actions de l'entreprise, soit en numéraire. Elle s'apparente donc à une option d'achat sur les actions. Ce mécanisme de financement permet à des investisseurs qui ne souhaitent pas prendre le risque d'entrer immédiatement au capital d'une entreprise de tout de même bénéficier du développement de celle-ci.

10.1.6 Les Equity Swaps

Les equity swaps sont des produits OTC. Ils permettent notamment

- ❖ **couvrir contre une baisse probable à court terme d'un portefeuille d'actions.** Le système permet ainsi d'éviter les impacts négatifs sur le court terme sans pour autant devoir vendre le portefeuille d'actions jugé rentable sur le long terme.
- ❖ à un investisseur qui n'a pas légalement la possibilité d'acquérir un stock de **bénéficier de son rendement tout en respectant les restrictions légales.**
- ❖ Bénéficier de la performance d'un actif sans forcément acquérir l'actif en question, ce qui peut parfois être pratique pour des titres assez chers.

Hedge d'un equity swap. Un investisseur peut, au lieu d'acheter des stocks directement sur le marché, passer par un **prime broker**. Dans ce cas, il va traiter un total return swap avec son prime broker qui lui paiera la performance des stocks et recevra en contrepartie les intérêts d'un taux flottant (floating rate). Afin de se couvrir, le prime broker va alors acheter les stocks et les détenir physiquement pour couvrir le risque de marché lié aux variations des prix du stock.

Exemple. Prenons l'exemple d'un prime broker qui va conclure un contrat price return swap pour payer la performance de N stocks A et recevoir les intérêts d'un nominal N .

Sur la période du swap, pour des raisons de simplicité, considérons l'absence de tombée de dividendes sur le stock A , et que le swap a une seule période entre t_0 et t_1 . Le taux associé à la floating leg à l'issue de la période étant r_1 , S_0 le spot du stock A à t_0 et S_1 le spot du stock A à t_1 .

À t_0 :

- ❖ Flux sur la floating leg: 0
- ❖ Pour se hedger, le prime broker achète n stocks A au prix unitaire S_0 , le flux associé à cette transaction : $-nS_0$

À t_1 :

- ❖ Flux sur la floating leg: Nr_1
- ❖ Flux sur la performance leg : le prime broker paie la performance des stocks sur la période $-n(S_1 - S_0)$
- ❖ Le prime broker solde sa position sur les stocks détenus et reçoit donc $n * S_1$

À la maturité du swap, le total T des flux financiers payés et reçus par le prime broker :

$$T = -nS_0 + Nr_1 - n(S_1 - S_0) + nS_1 = Nr_1 \quad (54)$$

Le prime broker est donc couvert contre les fluctuations du prix du stock et reçoit sans risque les intérêts du nominal N .

10.2 Réplication d'une option digitale

On peut répliquer une option digitale de strike K à l'aide de positions short et long d'un call option de strike K .

Pour cela il suffit de

- ❖ Long 1 call de maturité K
- ❖ Short δ call de maturité $K + \delta$ et de faire tendre δ vers 0 pour être le plus proche possible de l'option digitale

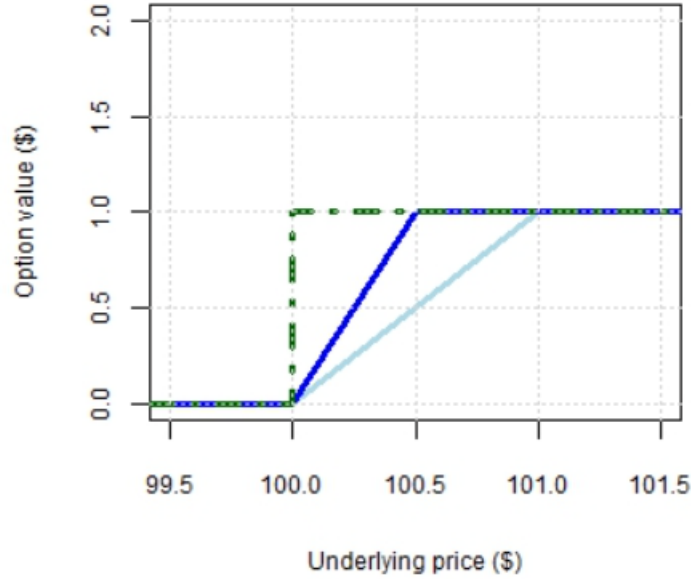


FIGURE 9 – replication d'une digital avec 1 call de maturité $100 + 1$ (bleu clair) et 0.5 call de maturité $100 + 0.5$ (bleu marine)

10.3 Evaluation d'un contrat forward et future

10.3.1 Prix d'un future – Arbitrage *cash and carry*

10.3.2 Prix d'un forward

10.3.3 Détermination sous Black–Scholes

10.4 Les Greeks

les grecques (ou "greek letters" en anglais) font référence à un ensemble de paramètres qui mesurent la sensibilité d'un instrument financier, tel qu'une option, à certains facteurs de risque, tels que la volatilité, les taux d'intérêt ou les mouvements du sous-jacent. Elles sont très utiles pour les traders et les investisseurs car ils leur permettent de comprendre les risques et les opportunités associés à un instrument financier donné, ainsi que de concevoir des stratégies de couverture et d'optimisation de leur portefeuille. En reprenant le modèle de Black-Scholes Et les paramètre d_1 et d_2 qui représentent

On utilise le développement en série de Taylor (à l'ordre 2) pour faire apparaître les *sensitivities* i.e.

$$\Delta V = \frac{\partial V}{\partial S} \Delta S + \frac{\partial V}{\partial T} \Delta T + \frac{\partial V}{\partial \sigma} \Delta \sigma + \frac{\partial V}{\partial r} \Delta r + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} (\Delta S)^2 \quad (55)$$

Greek	Expression	Exposition	Définition
Delta	$\Delta = \frac{\partial V}{\partial S}$	Actif	variation du SJ
Gamma	$\gamma = \frac{\partial^2 V}{\partial S^2}$	Convexité du payout	variation du Delta
theta	$\theta = \frac{\partial V}{\partial T}$	Temps	temps
rho	$\rho = \frac{\partial V}{\partial r}$	interest rate	RFR
vega	$\vartheta = \frac{\partial V}{\partial \sigma}$	volatilité	vol
volga	$\frac{\partial^2 V}{\partial \sigma^2}$	vol de vol	vol de vol
vanna	$\frac{\partial^2 V}{\partial \sigma \partial S}$	Skew	var de delta (vol) ou vega (SJ)

Table 1: Résumé des principales Greeks

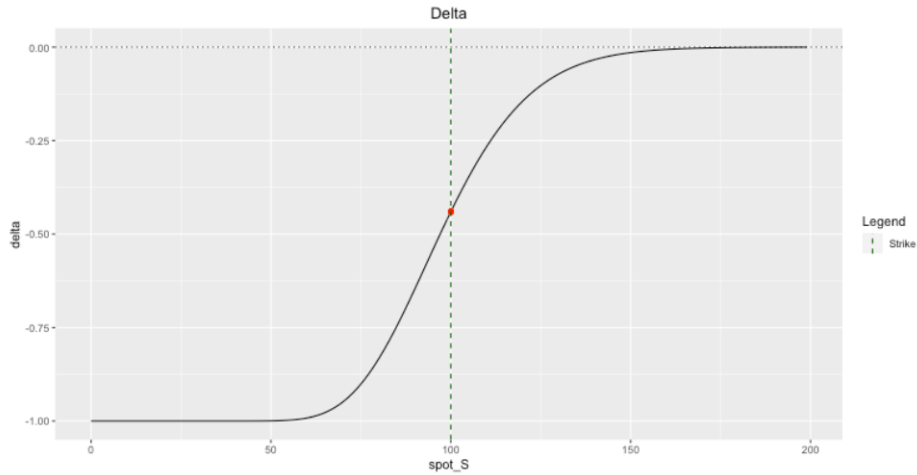
On peut identifier les principales Greeks dans le tableau, voici les détails de chacune d'entre elles

10.4.1 Le delta δ

Delta is the first-order sensitivity of the price to a movement in the spot price, S . Le delta d'une option mesure la sensibilité de son prix à une variation donnée du cours du sous-jacent.

$$\delta = \frac{\partial P}{\partial S}, \quad \delta_{call} = N(d_1), \quad \delta_{put} = \delta_{call} - 1 = N(d_1) - 1$$

La prime d'un *call* est une fonction croissante du prix du sous-jacent, $\delta_{call} \geq 0$, alors que celle d'un *put* en est une fonction décroissante, $\delta_{put} \leq 0$. En effet, plus le prix du sous-jacent est élevé, plus la probabilité que le call soit dans la monnaie est grande. Symétriquement, plus le prix du sous-jacent est bas, plus la probabilité que le put soit dans la monnaie est grande. Ainsi, lorsqu'une option a un delta égal (en valeur absolue) à ou proche de 0,5, on dit qu'elle est à la monnaie.

FIGURE 10 – Evolution du Δ en fonction du spot price S

Delta Hedging Le Delta Hedging permet de se couvrir d'une position

10.4.2 Le gamma γ

Gamma measures the change in delta due to the change in underlying price. Le gamma représente la convexité ou la termaxité du prix d'une option en fonction du cours du sous-jacent.

$$\gamma_{call} = \gamma_{put} = \frac{\partial^2 P}{\partial S^2} = \frac{N'(d_1)}{S\sigma\sqrt{T}}$$

Il indique si le prix de l'option a tendance à évoluer plus ou moins vite que le prix du sous-jacent.

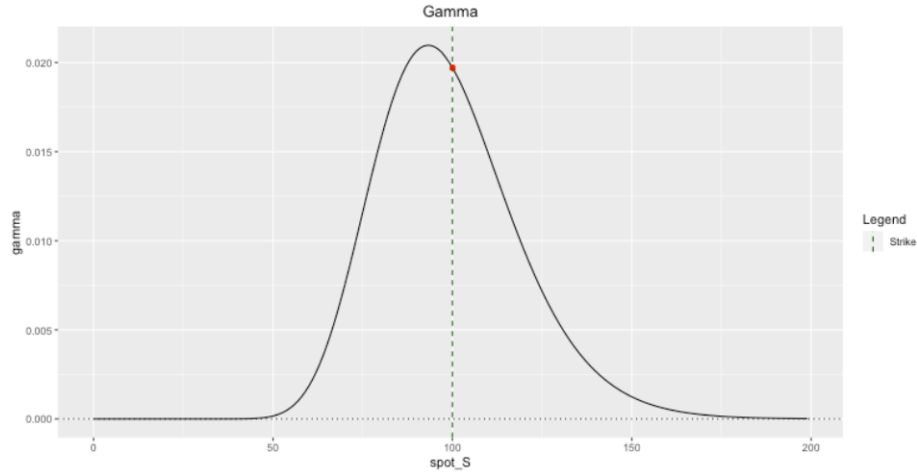


FIGURE 11 – Evolution du γ en fonction du spot price S

Par analogie, on peut comparer le delta à la vitesse et le gamma à l'accélération. Le gamma est une fonction décroissante de la maturité. Comme $\gamma_{call} \geq 0$ et $\gamma_{put} \geq 0$, on dit qu'un acheteur de call ou de put sera *long de gamma*, ou que son portefeuille sera gamma positif, et qu'un vendeur sera court (short) de gamma, ou gamma négatif. Toutes choses égales par ailleurs, le gamma est maximum lorsque l'option est à la monnaie (i.e. lorsque $\delta_{option} = 0.5$). Un portefeuille comportant des positions acheteuses (dites longues) et vendeuses (dites courtes) d'options à différents prix d'exercice (sur un même sous-jacent) verra donc la valeur de son gamma évoluer, voire changer de signe, en fonction des évolutions du prix du sous-jacent.

10.4.3 Le thêta θ

Le thêta est le coût (ou le gain) du temps qui passe sur un portefeuille d'options. Il évalue combien le passage du temps influe sur la valeur d'une option.

$$\theta = -\frac{\partial P}{\partial t}$$

En considérant des options européennes qui ne versent pas de dividendes

$$\theta_{call} = -\frac{SN'(d_1)\sigma}{2\sqrt{T}} - rKe^{-rT}N(d_2), \quad \theta_{put} = -\frac{SN'(d_1)\sigma}{2\sqrt{T}} + rKe^{-rT}N(-d_2)$$

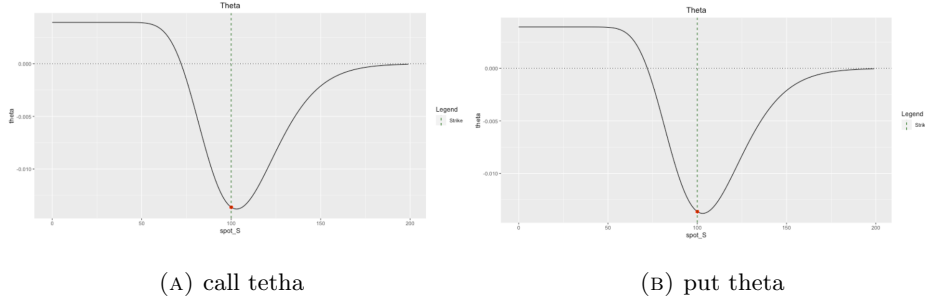


FIGURE 12 – Evolution du θ de l'option

Une position longue d'options (gamma positive) sera thêta négative. Le trader devra veiller tous les jours à payer son thêta journalier en profitant de sa position longue en gamma. On préférera donc être long d'une option qui soit suffisamment volatile, ainsi en rebalançant la position, on pourra payer le temps qui passe en tradant le gamma.

10.4.4 Le rhô ρ

Le rhô est le taux de variation de la valeur de la prime en fonction du taux sans risque.

$$\rho = \frac{\partial P}{\partial r}, \quad \rho_{call} = KTe^{-rT}N(d_2), \quad \rho_{put} = -KTe^{-rT}N(-d_2)$$

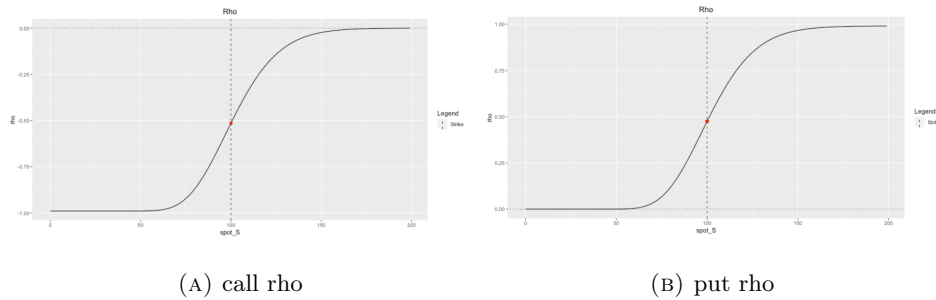


FIGURE 13 – Evolution du ρ de l'option

10.4.5 le véga ν

Le véga mesure de la sensibilité à la volatilité implicite (voir modèle Black-Scholes)

$$\nu_{call} = \nu_{put} = \frac{\partial P}{\partial \sigma} = S\sqrt{T}N'(d_1)$$

Comme $\nu_{call} \geq 0$ et $\nu_{put} \geq 0$, on dit qu'un acheteur de call ou de put sera long de véga, ou que son portefeuille sera véga positif, et qu'un vendeur sera court (short) de véga, ou véga négatif.

10.4.6 Exemple d'utilisation

an ATM call option price is \$7.97. The market parameters are the following:

- ❖ underlying price \$100 (the stock does not pay any dividend)
- ❖ volatility 20%
- ❖ rate of returns 0%
- ❖ delta 54%
- ❖ gamma 0.02
- ❖ vega 40
- ❖ theta -4

On peut alors poser les questions suivantes :

- Q1. Delta impact on price.** Quel sera le prix de l'option si celui du sous-jacent augmente de \$1 ?
- Q2. Gamma impact on Delta.** Si le gamma d'une option est de 0.04 et le delta est de 0.60, quel sera le nouveau delta si le prix sous-jacent augmente de \$1 ?
- Q3. Vega Impact on Price.** Si la vega d'une option est de 30 et le prix de l'option est actuellement de \$5, combien le prix de l'option augmentera-t-il si la volatilité implicite augmente de 1% ?
- Q4. Theta Impact on Time.** Si l'option perd 0.05 de valeur par jour en raison du theta, combien de valeur perdra-t-elle en une semaine ?
- Q5. Delta Hedging.** Vous détenez une position longue dans un call option avec un delta de 0.70 et un delta neutre à l'aide d'actions. Combien d'actions devrez-vous détenir pour neutraliser complètement la sensibilité au prix sous-jacent de l'option ?
- Q6. Vega Hedging.** Si vous détenez une option avec une vega de 25 et vous souhaitez vous couvrir contre une augmentation de la volatilité implicite de 5 points de pourcentage, combien de contrats d'options devrez-vous acheter ou vendre pour vous couvrir ?

- Q7. **Theta Decay Analysis.** Vous détenez une position longue dans une option dont le theta est -0.03. Combien de valeur votre option perdra-t-elle au bout d'un mois en raison de la décroissance temporelle ?
- Q8. **Impact of Volatility on Price.** Vous détenez un call option avec un vega de 40. Si la volatilité implicite augmente de 2 points de pourcentage, de combien le prix de l'option augmentera-t-il ?

11 Options exotiques

11.1 Options barrière

On distingue plusieurs types d'options barrière, on peut les classer selon quatre groupes

- ❖ **Knock-In & Up (UI):** Les options Up&In naissent lorsque la barrière est touchée, sachant que le spot au départ du sous-jacent se situe en dessous du niveau de barrière.
- ❖ **Knock-Out & Up (UO):** Les options Up&Out meurent lorsque la barrière est touchée, sachant que le spot au départ du sous-jacent se situe en dessous du niveau de barrière.
- ❖ **Knock-In & Down (DI):** Les options Down&In naissent lorsque la barrière est touchée, sachant que le spot au départ du sous-jacent se situe au dessus du niveau de barrière.
- ❖ **Knock-Out & Down (DO):** Les options Down&Out meurent lorsque la barrière est touchée, sachant que le spot au départ du sous-jacent se situe au dessus du niveau de barrière.

11.2 Baskets

11.2.1 Définition

Les baskets sont des options barrière sur plusieurs sous-jacents corrélés entre eux. On constate à chaque date d'observation la performance du basket. On considère plusieurs type de *performances*:

- ❖ **Worst-of/Best-of:** La performance du basket est donnée par l'actif présentant la pire (resp. meilleure) performance du basket.
- ❖ **Basket:** La performance est donné par la moyenne des performances de chaque actif, pondérée par le poids que représente l'actif dans le basket.
- ❖ **Memory:** La performance est donnée par chaque actif relativement à sa barrière par rapport à son spot à la date d'observation. Le basket est rappelé si chaque actif a tapé sa barrière au moins une fois pendant la durée d'existence du basket. Il n'y a pas besoin que tous les actifs tapent en même temps leurs barrière, un effet de *mémoire* est appliqué pendant toute la durée du basket.

11.2.2 Matrice de corrélation

Définition 10 On considère un basket de sous-jacents S_1, S_2, \dots, S_n répartis selon des poids w_1, w_2, \dots, w_n . La realized correlation du basket est la quantité

$$\rho_r = \frac{\sum_{1 \leq i < j \leq n} w_i w_j \rho_{i,j}}{\sum_{1 \leq i < j \leq n} w_i w_j}$$

Théorème 11.1 La relation entre la variance du basket et la corrélation est

$$\sigma_{ptf}^2 = \sum_{i=1}^n w_i^2 \sigma_i^2 + 2 \sum_{1 \leq i < j \leq n} w_i w_j \sigma_i \sigma_j \rho_{i,j} \quad (56)$$

11.2.3 Corrélation implicite

On considère deux options: une option (européenne) sur un index (par exemple sur le S&P500) et un basket option sur l'ensemble des actifs composants l'indice. On peut alors introduire une corrélation implicite dérivée de la volatilité réelle donné par le marché.

Définition 11 La corrélation implicite de l'index est une mesure de la dépendance entre un indice et ses composants

$$\rho_{index \text{ implied}} = \frac{\sigma_{index}^2 - \sum_{i=1}^n w_i^2 \sigma_i^2}{2 \sum_{1 \leq i < j \leq n} w_i w_j \sigma_i \sigma_j \rho_{i,j}}$$

11.3 Autocallables

11.3.1 Description

Les autocalls (automatically callable) sont des produits structurés qui ont la particularité de pouvoir être rappelé en cas de dépassement de barrière (AB) à des dates d'observation définies à l'émission. Ils donne aussi la possibilité à l'investisseur de recevoir des coupons à des dates d'observation si la barrière Coupon (CB) n'est pas tapée. Les barrières Coupon et Autocall (CB et AB) peuvent être les mêmes ou être différentes. Le pay-off d'un autocall contient deux composantes, le Coupon et le Rachat:

$$Coupon(t_i) = N * C\% * \mathbb{1}_{\{$$

11.4 Hedge d'une option barrière

Pour hedger une option barrière, le trader va acheter un *call spread*.

12 Etude de cas – Stochastic repo hybrid model

12.1 Repo rate dynamics

12.2 Linear Gaussian Markov model (LGM)

12.3 Exchanges rate model

12.4 Pricing dans une devise étrangère

12.5 Calibration des corrélations

12.5.1 Quanto options

12.5.2 Hybrid stock/funding exchange

12.5.3 Hybrid basket

13 Dérivés de Crédit

14 Modèles de regression

14.1 Décomposition biais–variance

Le cadre général des modèles statistiques est le suivant. On suppose que les données x_i sont liés aux variables y_i par la relation

$$y_i = f(x_i) + \epsilon_i \quad (57)$$

Où f est une fonction inconnue et ϵ_i suit une loi normale de moyenne nulle et de variance σ^2 . Le but de la regression est de trouver une fonction \hat{f} qui approxime le mieux f en tenant le moins possible compte de ϵ_i . Ainsi, on cherche à minimiser l'erreur au carré moyenne qui peut etre décomposée sous la forme

$$\mathbb{E} \left[(y - \hat{f}(x))^2 \right] = \underbrace{\mathbb{E} \left[f(x_i) - \hat{f}(x) \right]^2}_{\text{biais}} + \underbrace{\mathbb{E} \left[f(x_i)^2 \right] - \mathbb{E} \left[\hat{f}(x) \right]^2}_{\text{variance}} + \sigma^2 \quad (58)$$

On peut interpréter les différents termes comme

❖ Le biais $\mathbb{E}[f(x_i) - \hat{f}(x)]^2$ représente à quel point le modèle \hat{f} approxime la fonction f

❖ la variance

$$\mathbb{E}[f(x_i)^2] - \mathbb{E}[\hat{f}(x)]^2$$

représente le niveau de variabilité de \hat{f} sans tenir compte de f

❖ σ^2 représente le niveau de bruit de (y_i, x_i)

Les méthodes de régression sont largement utilisées en finance de marché pour modéliser et analyser les relations entre différentes variables. Ces méthodes permettent aux professionnels de la finance d'analyser les données historiques, de prévoir les tendances futures, de mesurer les risques et de prendre des décisions éclairées.

14.2 Moindres carrés (OLS)

Cadre théorique. On considère un échantillon $(Y_i, X_{i1}, \dots, X_{in})$. On cherche à estimer la *variable endogène* Y_i à partir des *variables explicatives* (X_{i1}, \dots, X_{in})

$$Y_i = a_{i1}X_{i1} + \dots + a_{in}X_{in} + \epsilon_i \quad (59)$$

Ce qui donne sous forme matricielle, avec \mathbf{Y} de dimension $(n, 1)$, \mathbf{X} de dimension $(n, p+1)$, \mathbf{a} de dimension $(p+1, 1)$, ϵ de dimension $(n, 1)$, n le nombre d'observations et p le nombre de variables explicatives,

$$\mathbf{Y} = \mathbf{X}\mathbf{a} + \epsilon \quad (60)$$

On va estimer les paramètres $(a_i)_{0 \leq i \leq p+1}$ afin d'obtenir

$$\hat{y}_i = \hat{a}_0 + \hat{a}_1 x_{i,1} + \dots + \hat{a}_p x_{i,p}$$

La méthode des moindres carrés consiste à minimiser la somme des carrés des résidus, à savoir

$$\min \sum_{i=1}^n \hat{\epsilon}_i^2 = \min_{\hat{a}_0, \dots, \hat{a}_p} \sum_{i=1}^n (\hat{y}_i - \hat{a}_0 - \hat{a}_1 x_{i,1} - \dots - \hat{a}_p x_{i,p})^2 \quad (61)$$

Théorème. La solution de ce problème est

$$\hat{\mathbf{a}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (62)$$

C'est un estimateur sans biais i.e. $\mathbb{E}[\hat{a}] = a$.

14.3 Regression Ridge (L^2)

On rappelle le système linéaire de notre modèle.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(q)} \\ 1 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(q)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(q)} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Le regression ridge impose une contrainte L^2 sur les coefficients pour éviter qu'ils partent dans tous les sens. On cherche donc à trouver

$$\min_{\beta_1, \dots, \beta_p} \sum_{i=1}^n \left(y_i - \sum_{k=1}^p \beta_k z_{i,k} \right)^2 \quad (63)$$

Sous la contrainte

$$\sum_{i=1}^n \beta_i^2 \leq \tau$$

Avec τ un paramètre à fixer. Les variables de base X_i on été centrées réduites en Z_i pour éviter que les variables à trop haute variance n'aient trop d'influence.

On peut montrer que le problème est équivalent à ajouter une *pénalité* à la fonction à minimiser. On cherche alors un paramètre $\lambda \geq 0$ et on cherche

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|Y - X\beta\|^2 + \lambda \|\beta\|^2 \quad (64)$$

Théorème. L'estimateur ridge s'écrit alors:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \mathbf{X}^T \mathbf{Y} \quad (65)$$

Démonstration.

Démonstration 6 Posons notre fonction de coût qu'on cherche à minimiser

$$J_\lambda(\beta) = \|Y - X\beta\|^2 + \lambda \|\beta\|^2 = (Y - X\beta)^T (Y - X\beta) + \lambda \beta^T \beta \quad (66)$$

En développant on obtient

$$J_\lambda(\beta) = Y^T Y - 2X^T Y + \beta^T (X^T X + \lambda I_p) \beta$$

Comme la matrice $X^T X + \lambda I_p$ est symétrique on utilise la propriété

$$\frac{\partial x^T A x}{\partial x} = 2Ax, \text{ pour } A \text{ symétrique.} \quad (67)$$

On obtient alors l'estimateur $\hat{\beta}$ tel que

$$\partial_\beta J_\lambda(\hat{\beta}) = 0 \Leftrightarrow (X^T X + \lambda I_p) \hat{\beta} = X^T Y$$

Et on obtient le résultat.

Note. Avec un raisonnement similaire on peut obtenir l'estimateur MCO (LSE) et on pourra introduire la matrice pseudo inverse de X

$$X^\dagger = (X^T X)^{-1} X^T \quad (68)$$

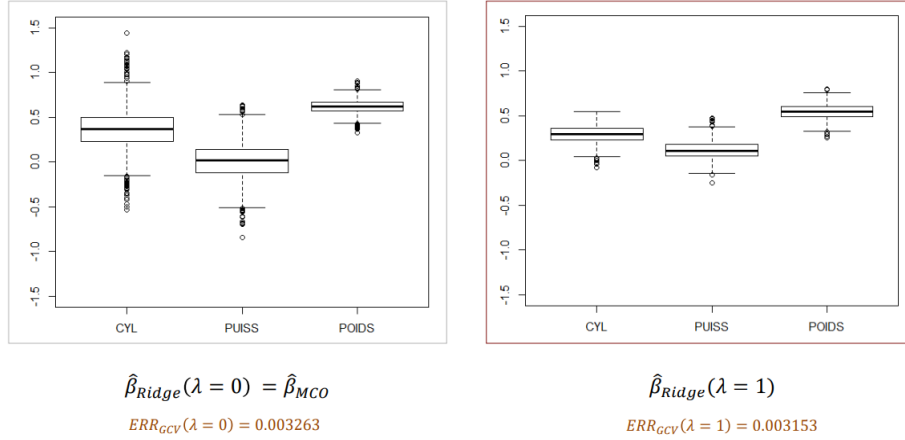


FIGURE 14 – Stabilisation des coefficients avec Ridge

C'est un petit exemple ($n = 27, p = 3$), le gap de performances n'est pas très flamboyant. En revanche, la réduction de la variance des coefficients estimés saute aux yeux.

On peut déterminer la valeur de λ de manière empirique ou par le calcul. On trouve alors les estimateurs LSO et on a deux calculs *Hoerl, Kennard, Baldwin (1975)* et *Lawless, Wang (1976)*

$$\lambda_{HKB} = \frac{(p-2)\hat{\sigma}_\epsilon^2}{\sum \hat{\beta}_j^2} \text{ et } \lambda_{LW} = \frac{n(p-2)\hat{\sigma}_\epsilon^2}{\sum \hat{y}_j^2}$$

14.4 Regression Lasso (L^1)

14.4.1 Enoncé du problème

Tout comme la regression Ridge, on impose ici une pénalité mais cette fois-ci de type L^1 sur les coefficients c'est-à-dire

$$\min_{\beta_1, \dots, \beta_p} \sum_{i=1}^n \left(y_i - \sum_{k=1}^p \beta_k z_{i,k} \right)^2 \quad (69)$$

Sous la contrainte

$$\sum_{i=1}^n |\beta_i| \leq \tau$$

Qu'on peut écrire de manière équivalente

$$\min_{\beta_1, \dots, \beta_p} \sum_{i=1}^n \left(y_i - \sum_{k=1}^n \beta_k z_{i,k} \right)^2 + \lambda \sum_{i=1}^p |\beta_i| \quad (70)$$

Quel intérêt par rapport à Ridge ? La regression Lasso peut faire office de dispositif de sélection de variables en annulant certains coefficients β_j : les variables associées à ($\beta_j = 0$) sont de facto exclues du modèle prédictif.

14.4.2 Forward Stagewise Algorithm

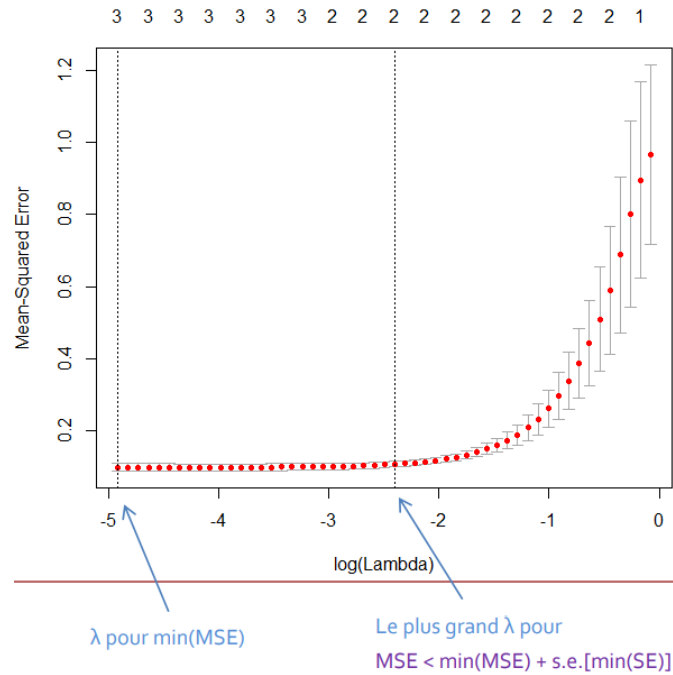
Il n'y a pas de calcul direct pour la regression Lasso. On résoud alors par un algorithme (LARS ou FSA). C'est une démarche itérative où l'on commence avec tous les ($\beta_j = 0$). On fait évoluer les coefficients sélectivement

Algorithm 1 *Incremental Forward Stagewise Regression: FS_ϵ*

1. Start with $\mathbf{r} = \mathbf{y} - \bar{\mathbf{y}}$, $\beta_1, \beta_2, \dots, \beta_p = 0$.
 2. Find the predictor \mathbf{x}_j most correlated with \mathbf{r} .
 3. Update $\beta_j \leftarrow \beta_j + \delta_j$, where $\delta_j = \epsilon \cdot \text{sign}[\text{corr}(\mathbf{r}, \mathbf{x}_j)]$;
 4. Update $\mathbf{r} \leftarrow \mathbf{r} - \delta_j \mathbf{x}_j$, and repeat steps 2 and 3 until no predictor has any correlation with \mathbf{r} .
-

FIGURE 15 – Forward Stagewise Algorithm

Dans l'optique prédictive, la minimisation de l'erreur de prédiction en balayant les valeurs de λ reste la méthode la plus efficace. On procède souvent par validation croisée

FIGURE 16 – Choix du paramètre λ

En résumé:

- ❖ (+) Capacité à sélectionner les variables en acceptant les coefficients nuls
- ❖ (–) Dans les problèmes à très grandes dimensions ($p \gg n$), LASSO ne sélectionne que n variables prédictives au maximum, mécaniquement. C'est une vraie limitation de l'algorithme.

- ❖ (–) Parmi un groupe de variables corrélées, LASSO en choisit une, **celle qui est la plus liée à a cible souvent**, masquant l'influence des autres. Cet inconvénient est inhérent aux techniques intégrant un mécanisme de sélection de variables

15 Méthodes numériques

15.1 Monte Carlo

15.1.1 Algorithme de base

15.1.2 Réduction de variance

15.1.3 Markov Chain Monte Carlo

15.2 Résolutions de PDE

15.2.1 Adjoint Algorithmic Differentiation

<https://quant.opengamma.io/Adjoint-Algorithmic-Differentiation-OpenGamma.pdf>

L'adjoint algorithmique de différentiation est une technique utilisée pour calculer les gradients de fonctions numériques efficacement et précisément. Il est souvent utilisé en finance pour le calcul des sensibilités (ou des "Greeks") de produits financiers complexes tels que les options et les dérivés.

L'idée principale de l'adjoint algorithmique de différentiation est de calculer les gradients en résolvant un système d'équations différentielles, appelé équations adjointes, qui est dérivé du calcul original. Ce système d'équations adjointes peut être résolu de manière très efficace par rapport aux techniques de différenciation numérique classiques telles que la différence finie.

En utilisant l'adjoint algorithmique de différentiation, il est possible de calculer les **sensibilités des portefeuilles de produits financiers complexes de manière plus rapide et plus précise, ce qui peut être crucial pour les stratégies de trading et de gestion des risques.**

15.2.2 Finite Differences

Les différences finies sont une méthode numérique utilisée en finance pour calculer les prix et les sensibilités de produits financiers tels que les options, les futures et les obligations. La méthode des différences finies consiste à discrétiser les équations différentielles qui sous-tendent les modèles financiers continus, en les transformant en une série d'équations discrètes qui peuvent être résolues numériquement.

En utilisant les différences finies, les traders et les gestionnaires de risques peuvent calculer les prix et les sensibilités de produits financiers complexes de manière efficace et précise, en fonction des paramètres tels que la volatilité du marché, les taux d'intérêt et les dividendes.

```
S, K, T, r, sigma = 100, 105, 1, 0.05, 0.2
```

```
# Inputs pour le maillage
```

```

N, M = 100, 100 # Nombre de pas de temps et d'espace
dt = T/N # Pas de temps
dx = np.sqrt(dt)*sigma # Pas d'espace
x_max = 5*sigma*np.sqrt(T) # Valeur maximale de x
x_min = -x_max # Valeur minimale de x
x_values = np.linspace(x_min, x_max, M+1)

# Initialisation des tableaux
u_old = np.zeros(M+1)
u_new = np.zeros(M+1)

# Conditions aux limites
u_old[0] = 0
u_old[M] = np.maximum(S - K*np.exp(x_max), 0)

# Calcul de u_new
for i in range(1, N+1):
    # Conditions aux limites pour chaque pas de temps
    u_new[0] = 0
    u_new[M] = np.maximum(S - K*np.exp(x_max), 0)

    # Boucle de calcul pour chaque point d'espace
    for j in range(1, M):
        a = 0.5*sigma**2*(x_values[j]**2) - r*x_values[j]
        b = -sigma**2*(x_values[j]**2) - r
        c = 0.5*sigma**2*(x_values[j]**2) + r*x_values[j]
        u_new[j] = (a*u_old[j-1] + b*u_old[j] +
                    c*u_old[j+1])/(1+r*dt)

    # Mise a jour de u_old pour le prochain pas de temps
    u_old = u_new.copy()

# Calcul du prix de l'option
price = np.interp(np.log(S/K), x_values, u_new)

```

16 Neural Networks

16.1 Artificial Neural Network

16.1.1 Motivations

Les régressions linéaires assument une relation linéaire entre les variables indépendantes et dépendantes. Cependant, de nombreuses relations dans les données réelles peuvent être non linéaires. Les réseaux de neurones, avec leurs fonctions d'activation non linéaires, sont capables de capturer des relations complexes et flexibles entre les entrées et les sorties.

Exemple. XOR On cherche ici à modéliser la porte XOR (non-linéaire). On dispose alors d'un dataset d'entrée

$$\mathcal{D} = \left\{ ((0,0)^T, 0), ((1,0)^T, 1), ((0,1)^T, 1), ((1,1)^T, 0) \right\}$$

En utilisant une regression de type LSE, on cherche alors à minimiser la fonction cout

$$J(\beta) = \|Y - W^T X - w_0\|^2$$

Et on a la fonction estimée $Y(X) = W^T X + w_0$. Cependant, à cause de la non-linéarité, on trouve comme solution optimale $W = 0$ et $w_0 = 0.5$ i.e. $Y(X) = 0.5$ partout.

On choisit alors de construire un 2-layer perceptron:

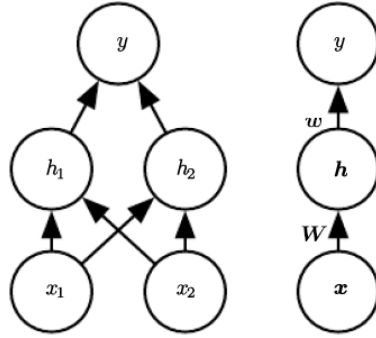


FIGURE 17 – Two-layer perceptron

Les caractéristiques sont les suivantes:

- ❖ **Hidden unit.** $h = g(W^T x + c)$ Avec $g(\alpha) = \max(0, \alpha)$
- ❖ **Sortie.** La sortie est $y = w^T h + b$ Soit $y = w^T \max(0, W^T x + c) + b$
- ❖ **Loss function.** La fonction de perte à minimiser est l'erreur LSE

$$J(\beta) = \frac{1}{N} \sum_{n=1}^N (t_n - y(x_n))^2$$

La solution optimale est alors

$$W = \text{mat}(1, 1, 1, 1), c = (0, -1)^T, w = (1, -2)^T, b = 0$$

16.1.2 Fonctionnement

Un réseau de neurones artificiels (*Artificial Neural Network*, *Multilayer Perceptron*, *Feedforward Neural Network*) est un modèle informatique qui est inspiré par le fonctionnement du cerveau humain. Il est composé de plusieurs unités appelées neurones, organisées en couches, et ces neurones sont connectés les uns aux autres. Il se compose de plusieurs caractéristiques clés:

- ❖ **Entrées** : Le processus commence avec la couche d'entrée, qui représente **les données d'entrée du modèle**. Chaque neurone dans cette couche correspond à une caractéristique ou à une dimension des données d'entrée. Par exemple, dans le cas de la reconnaissance d'image, chaque neurone de la couche d'entrée pourrait représenter **la valeur d'un pixel**.
- ❖ **Poids et biais** : Chaque connexion entre les neurones de deux couches consécutives est associée à un **poids**. Les poids sont des paramètres ajustables qui modifient l'influence d'une entrée sur la sortie d'un neurone. Un **biais** est également ajouté à chaque neurone pour introduire un certain niveau de flexibilité dans le modèle.
- ❖ **Somme pondérée** : Pour chaque neurone dans la couche suivante (*hidden layer*), la somme pondérée des entrées (multipliées par les poids correspondants) est calculée, et le biais est ajouté. **Cette somme pondérée est ensuite passée à travers une fonction d'activation.**
- ❖ **Fonction d'activation** : La fonction d'activation introduit de la non-linéarité dans le modèle. Elle détermine si et dans quelle mesure un neurone doit être activé. Différentes fonctions d'activation, telles que la sigmoïde, la tangente hyperbolique, ou la ReLU (Rectified Linear Unit), peuvent être utilisées.
- ❖ **Propagation avant** : Ce processus de calcul de la somme pondérée, d'ajout du biais, d'application de la fonction d'activation **est répété pour chaque couche** jusqu'à ce que les sorties soient générées par la couche de sortie. C'est ce qu'on appelle la propagation avant (*forward propagation*).
- ❖ **Fonction de perte** : Les sorties du réseau sont comparées aux valeurs attendues à l'aide d'une fonction de perte (*loss function*). La fonction de perte mesure l'écart entre les prédictions du réseau et les valeurs réelles. L'objectif lors de l'entraînement du réseau est de minimiser cette fonction de perte.
- ❖ **Rétropropagation** : Pour ajuster les poids et les biais afin de minimiser la fonction de perte, on utilise un algorithme d'optimisation. La rétropropagation du gradient (*backpropagation*) est l'algorithme le plus couramment utilisé. Il calcule les gradients de la fonction de perte par rapport aux poids du réseau, puis ajuste les poids en conséquence pour réduire l'erreur.
- ❖ **Apprentissage** : La propagation avant, suivi de la rétropropagation, est répété sur plusieurs itérations (époques) jusqu'à ce que le réseau apprenne à effectuer la tâche souhaitée de manière satisfaisante. Les poids et les biais sont ajustés itérativement pendant cette phase d'apprentissage.

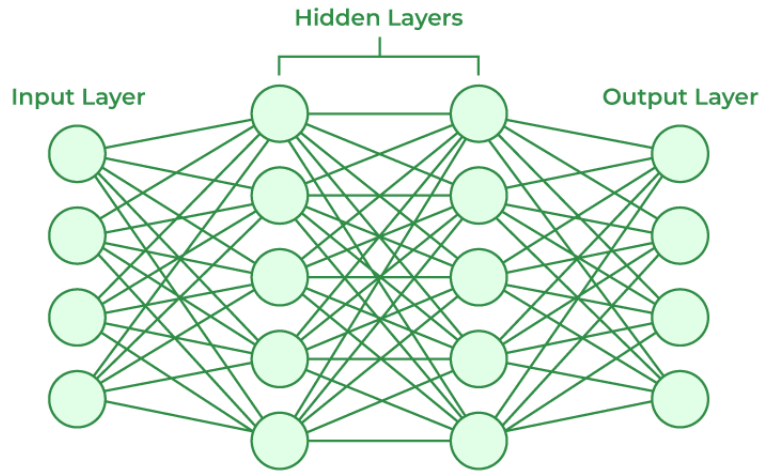


FIGURE 18 – Artificial Neural Network

16.1.3 Output units

On considère la sortie d'un *hidden layer* $h = f(x, \theta^{n-1})$. On cherche alors la sortie $y = f(h, \theta^n)$ ainsi que la fonction de perte $J(\theta)$. On trouve alors

- ❖ Regression
- ❖ Binary classification
- ❖ Multi-class classification

16.1.4 Fonctions d'activation

La fonction d'activation est une fonction appliquée à un signal en sortie d'un neurone artificiel. Le terme de "fonction d'activation" vient de l'équivalent biologique "potentiel d'activation", seuil de stimulation qui, une fois atteint entraîne une réponse du neurone.

Propriétés. On utilise souvent des fonctions qui ont les propriétés suivantes:

- (i) **Non-linéarité.** permet au réseau de neurones de modéliser des relations complexes et non linéaires entre les variables d'entrée et de sortie.
- (ii) **Différentiabilité.** Partout différentiable, cela facilite l'optimisation du réseau à l'aide de techniques telles que la rétropropagation du gradient.
- (iii) **Etendue.** Quand l'étendue est finie, les méthodes d'apprentissage basées sur les gradients sont plus stables (impact sur un nombre de poids limités). Quand l'étendue est infinie, l'apprentissage est généralement plus efficace (impact sur davantage de poids).

- (iv) **Monotonie de la dérivée** Les fonctions à dérivée monotone ont été montrées comme ayant une meilleure capacité à généraliser dans certains cas. ★

Exemples de fonctions d'activation


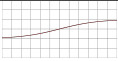

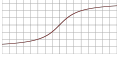
Nom	Graphe	Expression	Dérivée	Étendue
Heaviside		$H(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1, & \text{si } x \geq 0 \end{cases}$	$\delta(x)$	$\{0, 1\}$
Sigmoïde		$\sigma(x) = \frac{1}{1+e^{-x}}$	$\sigma'(x) = \sigma(x)(1 - \sigma(x))$	$[0, 1]$
ReLU		$\text{ReLU}(x) = \max(0, x)$	$\text{ReLU}'(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1, & \text{si } x \geq 0 \end{cases}$	\mathbb{R}
Tanh		$\tanh(x) = \frac{e^{2x}-1}{e^{2x}+1}$	$\tanh'(x) = 1 - \tanh^2(x)$	$[-1, 1]$

Table 2: Tableau des fonctions d'activation

17 Brainteasers