MATHÉMATIQUES & APPLICATIONS 61

Huyên Pham

Optimisation et contrôle stochastique appliqués à la finance

SMAI

Springer

MATHÉMATIQUES
&
APPLICATIONS

Directeurs de la collection :
G. Allaire et M. Benaïm

61

Huyên Pham

Optimisation et contrôle stochastique appliqués à la finance

Huyên Pham

Université Paris 7
2, Place Jussieu
75251 Paris Cedex 05
France
e-mail: pham@math.jussieu.fr

et

Institut Universitaire de France
103, boulevard Saint-Michel
75005 Paris
France

Library Congress Control Number: 2007931193

Mathematics Subject Classification (2000): 93E20, 91B28, 49L20, 49L25, 60H30

ISSN 1154-483X
ISBN-10 3-540-73736-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73736-0 Springer Berlin Heidelberg New York

springer.com
WMXDesign GmbH

Imprimé sur papier non acide 3100/SPi - 5 4 3 2 1 0 -

A Châu, Hugo et Antoine.

Préface

Les problèmes d'optimisation stochastique ont de nombreuses applications dans des problèmes de gestion, d'économie et de finance. Ce sont des situations où l'on fait face à des systèmes dynamiques évoluant dans des conditions d'incertitude et où il s'agit de prendre des décisions à chaque date afin d'optimiser un critère économique.

Les approches traditionnelles pour résoudre les problèmes d'optimisation stochastique sont basées sur le principe de la programmation dynamique de Bellman et le principe du maximum de Pontryagin. Elles ont conduit à de nombreux développements mathématiques. Le principe de la programmation dynamique appliqué au contrôle de processus de Markov en temps continu se traduit en termes d'équations aux dérivées partielles non linéaires pour la fonction valeur du problème. Ce type d'équations appelé équations d'Hamilton-Jacobi-Bellman a trouvé un cadre théorique adéquat grâce au concept des solutions de viscosité. Le principe du maximum appliqué au cadre de processus d'Itô a conduit à l'étude des équations différentielles stochastiques rétrogrades et engendré une littérature considérable sur le sujet. Plus récemment et motivé par les problèmes d'optimisation de portefeuille en finance, une nouvelle approche, dite méthode martingale de dualité convexe, s'est développée. Elle repose sur des théorèmes récents d'analyse stochastique et sur des méthodes plus classiques d'analyse convexe en optimisation. Il existe plusieurs ouvrages traitant soit de l'approche mathématique par programmation dynamique ([FR75], [BL78], [Kry80], [FSo93], [YZ00]) soit des équations différentielles stochastiques rétrogrades [MY00]. Ils privilégient souvent l'aspect théorique et sont d'un niveau techniquement avancé et parfois difficiles à lire pour un non spécialiste du sujet. D'autre part, bien qu'il existe de nombreux articles sur la maximisation d'utilité par approche duale, cette méthode est encore peu abordée dans les ouvrages d'enseignement et de recherche.

L'objectif de ce livre est de combler cette lacune et de présenter les différents aspects et méthodes utilisés dans la résolution des problèmes d'optimisation stochastique avec en vue des applications plus spécifiques à la finance. Nous avons inclus certains développements récents sur le sujet sans chercher à priori la plus grande généralité. Nous avons voulu une exposition graduelle des méthodes mathématiques en présentant d'abord les idées intuitives puis en énonçant précisément les résultats avec des démonstrations complètes et détaillées. Nous avons aussi pris soin d'illustrer chacune des méthodes de résolution sur de nombreux exemples issus de la finance. Nous espérons ainsi que ce livre puisse être utile aussi bien pour des étudiants que pour des chercheurs du monde académique ou professionnel intéressés par l'optimisation et le contrôle stochastique appliqués à la finance.

Le plan du livre est le suivant. Nous commençons par énoncer au chapitre 1 quelques prérequis en calcul stochastique. Nous avons essentiellement collecté les notions et résultats dans la vaste littérature sur l'analyse stochastique qui sont utiles pour l'optimisation stochastique. Le chapitre 2 formule de façon générale la structure d'un problème de contrôle stochastique et présente de nombreux exemples d'applications réelles en finance. Ces exemples seront résolus dans les chapitres suivants selon les différentes approches abordées. Nous discutons aussi très brièvement dans ce chapitre d'autres modèles de contrôle apparaissant en finance que ceux étudiés dans ce livre. Le chapitre 3 expose l'approche du contrôle stochastique de processus de diffusion par la programmation dynamique. Nous avons d'abord suivi une démarche classique basée sur un théorème de vérification associé à l'équation aux dérivées partielles d'Hamilton-Jacobi-Bellman. Le chapitre 4 reprend le principe de la programmation dynamique mais en adoptant une démarche plus moderne basée sur la théorie des solutions de viscosité. Ce concept permet de résoudre des problèmes de contrôle lorsque la fonction valeur n'est pas régulière comme supposée au chapitre précédent. Ce cas est illustré sur le problème de la surréplication en finance pour des modèles à volatilité incertaine. Comme nous l'avons mentionné plus tôt, le principe du maximum de Pontryagine a conduit naturellement au concept d'équations différentielles stochastiques rétrogrades. Le chapitre 5 est une introduction à cette théorie en insistant plus particulièrement sur ses applications au contrôle stochastique. Dans le chapitre 6, nous exposons l'approche martingale par dualité convexe inspirée originellement par le problème de gestion de portefeuille. Nous y avons inclus des développements récents et le problème de couverture quadratique. La caractéristique de cette approche est de combiner des résultats puissants d'analyse stochastique et des méthodes de dualité en analyse convexe.

Ce livre est basé sur mes notes de cours rédigées pour un enseignement en troisième année à l'ENSAE depuis 1998, puis aux DEA (maintenant Master 2) de Statistique et Modèles Aléatoires à l'université Paris 7 et de Probabilités et Finance à l'université Paris 6. Je tiens en particulier à remercier Laure Elie qui m'a donné l'opportunité d'enseigner ce cours à Paris 7. C'est grâce aux cours et exposés de Nicole El Karoui et Pierre-Louis Lions, lorsque j'étais étudiant en DEA et en thèse, que je me suis intéressé au contrôle stochastique et leurs travaux ont un impact évident dans la rédaction de cette monographie.

Paris, Septembre 2005 *Huyên PHAM*

Table des matières

Notations XIII

1 Quelques éléments d'analyse stochastique 1
1.1 Processus stochastiques 1
1.1.1 Filtration et processus 1
1.1.2 Temps d'arrêt 3
1.1.3 Mouvement brownien 5
1.1.4 Martingales, semimartingales 6
1.2 Intégrale stochastique et applications 13
1.2.1 Intégrale stochastique par rapport à une semimartingale continue 13
1.2.2 Processus d'Itô 18
1.2.3 Formule d'Itô 19
1.2.4 Théorèmes de représentation de martingales 20
1.2.5 Théorème de Girsanov 21
1.3 Equations différentielles stochastiques 24
1.3.1 Solutions fortes d'EDS 24
1.3.2 Estimations sur les moments de solutions d'EDS 27
1.3.3 Formule de Feynman-Kac 28

2 Problèmes d'optimisation stochastique. Exemples en finance 31
2.1 Introduction 31
2.2 Exemples 32
2.2.1 Allocation de portefeuille 32
2.2.2 Modèle de production et consommation 34
2.2.3 Modèles d'investissement irréversible d'une firme 35
2.2.4 Couverture quadratique d'options 35
2.2.5 Coût de surréplication en volatilité incertaine 36
2.3 Autres modèles de contrôles en finance 36
2.3.1 Arrêt optimal 36

2.3.2 Contrôle ergodique 37
2.3.3 Surréplication sous contraintes gamma 38
2.3.4 Optimisation d'utilité robuste et mesures du risque 38
2.4 Commentaires bibliographiques 39

3 Approche EDP classique de la programmation dynamique . 41
3.1 Introduction 41
3.2 Contrôle de processus de diffusion 41
3.3 Principe de la programmation dynamique 45
3.4 Equation d'Hamilton-Jacobi-Bellman 47
3.4.1 Dérivation formelle de HJB 47
3.4.2 Remarques-Extensions 50
3.5 Théorème de vérification 52
3.6 Applications 57
3.6.1 Problème de choix de portefeuille de Merton en horizon fini 57
3.6.2 Un modèle de production et consommation en horizon infini 59
3.7 Exemple de problème de contrôle stochastique singulier 62
3.8 Commentaires bibliographiques 64

4 Approche des équations de Bellman par les solutions de viscosité 65
4.1 Introduction 65
4.2 Définition des solutions de viscosité 66
4.3 Principe de comparaison 68
4.4 De la programmation dynamique aux solutions de viscosité . . . 70
4.5 Un modèle d'investissement irréversible 75
4.5.1 Problème 75
4.5.2 Régularité et construction de la fonction valeur 76
4.5.3 Stratégie optimale 82
4.6 Calcul du coût de surréplication en volatilité incertaine 83
4.6.1 Volatilité bornée 84
4.6.2 Volatilité non bornée 86
4.7 Commentaires bibliographiques 92

5 Méthodes d'équations différentielles stochastiques rétrogrades 93
5.1 Introduction 93
5.2 Propriétés générales 93
5.2.1 Résultats d'existence et d'unicité 93
5.2.2 EDSR linéaires 96
5.2.3 Principe de comparaison 97
5.3 EDSR, EDP et formules de type Feynman-Kac 98

5.4 Contrôle et EDSR . . . 102
5.4.1 Optimisation d'une famille d'EDSR . . . 103
5.4.2 Principe du maximum stochastique . . . 105
5.5 Applications . . . 108
5.5.1 Maximisation d'utilité exponentielle avec actif contingent . . . 108
5.5.2 Critère moyenne-variance d'allocation de portefeuille . . . 111
5.6 Commentaires bibliographiques . . . 116

6 Méthodes martingales de dualité convexe . . . 117
6.1 Introduction . . . 117
6.2 Représentation duale du problème de surréplication . . . 119
6.2.1 Formulation du problème de surréplication . . . 119
6.2.2 Probabilités martingales et arbitrage . . . 120
6.2.3 Le théorème de décomposition optionnelle et la représentation duale du coût de surréplication . . . 120
6.2.4 Le cadre de processus d'Itô et de filtration Brownienne . . . 124
6.3 Dualité pour la maximisation d'utilité . . . 128
6.3.1 Formulation du problème d'optimisation de portefeuille . . . 128
6.3.2 Résultat général d'existence . . . 129
6.3.3 Résolution via la formulation duale . . . 131
6.3.4 Le cas des marchés complets . . . 145
6.3.5 Exemples en marché incomplet . . . 146
6.4 Problème de couverture quadratique . . . 149
6.4.1 Formulation du problème . . . 149
6.4.2 Le cas martingale . . . 150
6.4.3 Probabilité martingale variance-optimale et numéraire quadratique . . . 152
6.4.4 Résolution du problème par changement de numéraire . . . 158
6.4.5 Exemple . . . 163
6.5 Commentaires bibliographiques . . . 164

A Compléments d'intégration . . . 167
A.1 Uniforme intégrabilité . . . 167
A.2 Essentiel supremum d'une famille de variables aléatoires . . . 169
A.3 Quelques théorèmes de compacité en probabilité . . . 169

B Considérations d'analyse convexe . . . 171
B.1 Fonctions semicontinues, convexes . . . 171
B.2 Transformée de Fenchel-Legendre . . . 173
B.3 Exemple dans $\mathbb{R}$ . . . 174

Références . . . 177

Index . . . 185

Notations

I. Notations générales

Pour tous nombres réels x et y :
$x \wedge y = \min(x, y)$, $\quad x \vee y = \max(x, y)$
$x^+ = \max(x, 0)$, $\quad x^- = \max(-x, 0)$

Pour toute suite positive $(x_n)_{n\geq 1}$ croissante, sa limite croissante dans $[0, +\infty]$ est notée $\lim_{n\to+\infty} \uparrow x_n$.

Pour toute suite $(x_n)_{n\geq 1}$, $y_n \in \mathrm{conv}(x_k, k \geq n)$ signifie que $y_n = \sum_{k=n}^{N_n} \lambda_k x_k$ où $\lambda_k \in [0, 1]$, $n \leq k \leq N_n < +\infty$ et $\sum_{k=n}^{N_n} \lambda_k = 1$.

II. Ensembles

$\mathbb{N}$ est l'ensemble des entiers naturels

$\mathbb{R}^d$ est l'espace réel euclidien de dimension d. $\mathbb{R} = \mathbb{R}^1$, $\mathbb{R}_+$ est l'ensemble des réels positifs, $\mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\}$ et $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Pour tous $x = (x^1, \ldots, x^d)$, $y = (y^1, \ldots, y^d)$ dans $\mathbb{R}^d$, on note . le produit scalaire et $|.|$ la norme euclidienne :

$$x.y \;=\; \sum_{i=1}^{d} x_i y_i \quad \text{et} \quad |x| \;=\; \sqrt{x.x}$$

$\mathbb{R}^{n\times d}$ est l'ensemble des matrices réelles $n \times d$ ($\mathbb{R}^{n\times 1} = \mathbb{R}^n$). I_n est la matrice identité $n \times n$. Pour tout $\sigma = (\sigma^{ij})_{1\leq i\leq n, 1\leq j\leq d} \in \mathbb{R}^{n\times d}$, on note $\sigma' = (\sigma^{ji})_{1\leq j\leq d, 1\leq i\leq n}$ la matrice transposée dans $\mathbb{R}^{d\times n}$. On note $\mathrm{tr}(A) = \sum_{i=1}^n a^{ii}$ la trace d'une matrice carrée $A = (a^{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$. On choisit comme norme matricielle sur $\mathbb{R}^{n\times d}$:

$$|\sigma| = \left(\mathrm{tr}(\sigma\sigma')\right)^{\frac{1}{2}}.$$

$\mathcal{S}_n$ est l'ensemble des matrices symétriques $n \times n$ et $\mathcal{S}_n^+$ est le sous-ensemble de $\mathcal{S}_n$ des matrices définies non-négatives. On définit l'ordre sur $\mathcal{S}_n$ par :

$$A \leq B \Longleftrightarrow B - A \in \mathcal{S}_n^+.$$

L'intérieur, l'adhérence et la frontière d'une partie $\mathcal{O}$ de $\mathbb{R}^d$ sont notées respectivement $\mathrm{int}(\mathcal{O})$, $\bar{\mathcal{O}}$ et $\partial\mathcal{O}$.

III. Fonctions et espaces fonctionnels

Pour tout ensemble A, l'indicatrice de A est noté :

$$1_A(x) = \begin{cases} 1, \, x \in A \\ 0, \, x \notin A \end{cases}$$

$C^k(\mathcal{O})$ est l'ensemble des fonctions continues f de $\mathcal{O}$ dans $\mathbb{R}$ qui ont des dérivées continues de tout ordre jusqu'à k. Ici $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$.

$C^0(\mathbb{T} \times \mathcal{O})$ est l'ensemble des fonctions continues f de $\mathbb{T} \times \mathcal{O}$ dans $\mathbb{R}$. Ici $\mathbb{T} = [0, T]$, avec $0 < T < +\infty$, ou $\mathbb{T} = [0, +\infty[$.

$C^{1,2}([0,T[\times\mathcal{O})$ est l'ensemble des fonctions f de $[0,T[\times\mathcal{O}$ dans $\mathbb{R}$ dont les dérivées partielles $\dfrac{\partial f}{\partial t}$, $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$, $1 \leq i,j \leq n$, existent et sont continues sur $[0,T[$ (T pouvant prendre la valeur $+\infty$). Si ces dérivées partielles de $f \in C^{1,2}([0,T[\times\mathcal{O})$ se prolongent par continuité sur $[0,T] \times \mathcal{O}$ (dans le cas $T < +\infty$), on note $f \in C^{1,2}([0,T] \times \mathcal{O})$. On définit de manière analogue pour $k \geq 3$, l'espace $C^{1,k}([0,T] \times \mathcal{O})$.

Pour une fonction $f \in C^2(\mathcal{O})$, on note Df le vecteur gradient dans $\mathbb{R}^n$ de composantes $\dfrac{\partial f}{\partial x_i}$, $1 \leq i \leq n$, et $D^2 f$ la matrice hessienne dans $\mathcal{S}_n$ de composantes $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$, $1 \leq i,j \leq n$. Ceux-ci sont notés parfois f_x et f_{xx}. Lorsque $\mathcal{O}$ est un ouvert de $\mathbb{R}$, on écrit aussi simplement f' et f''. Le vecteur gradient et la matrice hessienne d'une fonction $x \rightarrow f(t,x) \in C^2(\mathcal{O})$ sont notées $D_x f$ et $D_x^2 f$.

IV. Intégration et Probabilité

$(\Omega, \mathcal{F}, P)$: espace de probabilité.

P p.s. est la notation presque sûrement pour la mesure de probabilité P (on omettra souvent la référence à P lorsqu'il n'y a pas d'ambiguité). μ p.p. est la notation presque partout pour la mesure μ.

$\mathcal{B}(U)$: tribu borélienne engendrée par les ouverts de U espace topologique.

$\sigma(\mathcal{G})$: la plus petite σ-algèbre contenant $\mathcal{G}$, collection de sous-ensembles de Ω.

$Q \ll P$: la mesure Q est absolument continue par rapport à la mesure P.

$Q \sim P$: la mesure Q est équivalente à P, i.e. $Q \ll P$ et $P \ll Q$.

$\frac{dQ}{dP}$: densité de Radon-Nikodym de $Q \ll P$.

$E^Q(X)$ est l'espérance de la variable aléatoire X par rapport à Q.

$E(X)$ est l'espérance de la variable aléatoire X par rapport à une probabilité P fixée initialement. $E[X|\mathcal{G})$ est l'espérance conditionnelle de X sachant $\mathcal{G}$. $\mathrm{Var}(X) = E[(X - E(X))(X - E(X))']$ est la variance de X.

$L^0_+(\Omega, \mathcal{F}, P)$ est l'ensemble des variables aléatoires $\mathcal{F}$-mesurables qui sont positives p.s.

$L^p(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ est l'ensemble des variables aléatoires X, à valeurs dans $\mathbb{R}^n$, $\mathcal{F}$-mesurables tel que $E|X|^p < +\infty$, pour $p \in [1, +\infty[$. On omettra parfois certains arguments et on écrira L^p ou $L^p(P)$ lorsqu'il n'y a pas d'ambiguité.

$L^\infty(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ est l'ensemble des variables aléatoires, à valeurs dans $\mathbb{R}^n$, bornées, $\mathcal{F}$-mesurables. On écrira parfois L^∞.

V. Abréviations

EDS : Equations différentielles stochastiques.

EDSR : Equations différentielles stochastiques rétrogrades.

EDP : Equations aux dérivées partielles.

HJB : Hamilton-Jacobi-Bellman

1

Quelques éléments d'analyse stochastique

Dans ce chapitre, nous présentons les concepts et résultats d'analyse stochastique utiles pour ce cours. Il y a de nombreux livres détaillant la théorie classique exposée dans ce chapitre, parmi lesquels Dellacherie et Meyer [DM80], Jacod [Jac79], Karatzas et Shreve [KaSh88], Protter [Pro90] ou Revuz et Yor [ReY91], d'où sont tirés la plupart des résultats rappelés ici sans démonstration. Le lecteur est supposé familier avec les bases élémentaires de la théorie de l'intégration et des probabilités (voir par exemple Revuz [Rev94], [Rev97]). Dans la suite de ce chapitre, $(\Omega, \mathcal{F}, P)$ désigne un espace de probabilité. Pour $p \in [1, +\infty[$, on note $L^p = L^p(\Omega, \mathcal{F}, P)$ l'ensemble des variables aléatoires ξ (à valeurs dans $\mathbb{R}^d$) tel que $|\xi|^p$ soit intégrable, i.e. $E|\xi|^p < +\infty$.

1.1 Processus stochastiques

1.1.1 Filtration et processus

Un processus stochastique est une famille $X = (X_t)_{t\in\mathbb{T}}$ de variables aléatoires à valeurs dans un espace mesurable $\mathcal{X}$ et indéxées par le temps t. Dans ce chapitre et pour les objectifs de ce livre, on prendra $\mathcal{X} = \mathbb{R}^d$ muni de sa tribu borélienne. Le paramètre de temps t variant dans $\mathbb{T}$ peut être discret ou continu. Dans ce livre, on considère des processus stochastiques à temps continu et l'intervalle de variation du temps $\mathbb{T}$ est soit fini $\mathbb{T} = [0, T]$, $0 < T < +\infty$, soit infini $\mathbb{T} = [0, +\infty[$. On écrira souvent processus pour processus stochastique. Pour chaque $\omega \in \Omega$, l'application $X(\omega) : t \in \mathbb{T} \to X_t(\omega)$ est appelé une trajectoire du processus dans le scénario ω. Le processus stochastique X est dit càd-làg (resp. continu) si pour chaque $\omega \in \Omega$, la trajectoire $X(\omega)$ est continue à droite et admet une limite à gauche (resp. continue). Etant donné un processus stochastique $Y = (Y_t)_{t\in\mathbb{T}}$, on dit que Y est une modification de X si pour tout $t \in \mathbb{T}$, on a $X_t = Y_t$ p.s., i.e. $P[X_t = Y_t] = 1$. On dit que Y est indistinguable de X si leurs trajectoires coincident p.s. : $P[X_t = Y_t, \forall t \in \mathbb{T}] = 1$. Bien entendu, la notion d'indistinguabilité est plus forte que celle de modification, mais si on sait que les deux processus X et Y sont càd-làg, et si Y est une modification de X, alors X et Y sont indistinguables.

L'interprétation du paramètre t comme index de temps introduit un aspect dynamique : pour modéliser le fait que l'incertitude des évènements de Ω devient de moins en moins incertaine lorsque le temps s'écoule, i.e. on possède de plus en plus d'information, on introduit la notion de filtration.

Définition 1.1.1 *(Filtration)*
Une filtration sur $(\Omega, \mathcal{F}, P)$ est une famille croissante $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ de sous-tribus de $\mathcal{F}$: $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ pour tous $0 \leq s \leq t$ dans $\mathbb{T}$.

$\mathcal{F}_t$ s'interprète comme l'information connue à la date t et elle augmente avec le temps. On pose $\mathcal{F}_{\bar{T}} = \sigma(\cup_{t\in\mathbb{T}}\mathcal{F}_t)$ la plus petite σ-algèbre contenant tous les $\mathcal{F}_t$, $t \in \mathbb{T}$. Le quadruplet $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}, P)$ est appelé espace de probabilité filtré. L'exemple canonique de filtration est le suivant : si $X = (X_t)_{t\in\mathbb{T}}$ est un processus stochastique, la filtration naturelle (ou canonique) de X est

$$\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t), \quad t \in \mathbb{T},$$

la plus petite σ-algèbre par rapport à laquelle X_s est mesurable pour tous $0 \leq s \leq t$. $\mathcal{F}_t^X$ s'interprète comme toute l'information qu'on peut extraire de l'observation des trajectoires de X entre 0 et t.

On dit qu'une filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ satisfait les *conditions habituelles* si elle est continue à droite, i.e. :

$$\mathcal{F}_{t^+} \ := \ \cap_{s\geq t}\mathcal{F}_s = \mathcal{F}_t, \quad \forall t \in \mathbb{T},$$

et si elle est complète, i.e. $\mathcal{F}_0$ contient les ensembles négligeables de $\mathcal{F}_{\bar{T}}$. On dit alors que l'espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}, P)$ satisfait les conditions habituelles. La continuité à droite de $\mathcal{F}_t$ signifie intuitivement qu'ayant observé toute l'information disponible jusqu'en t inclus, on n'apprend rien de plus par une observation infinitésimale dans le futur. La complétion d'une filtration signifie que si un événement est impossible, cette impossibilité est déjà connue à la date 0. Partant d'une filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$ quelconque, on construit une filtration satisfaisant les conditions habituelles, en prenant pour tout $t \in \mathbb{T}$ la tribu $\mathcal{F}_{t^+}$ à laquelle on rajoute la classe des ensembles négligeables de $\mathcal{F}_{\bar{T}}$. La filtration ainsi construite est appelée *l'augmentation habituelle* de $(\mathcal{F}_t)_{t\in\mathbb{T}}$.

Dans la suite, on se donne une filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ sur $(\Omega, \mathcal{F}, P)$.

Définition 1.1.2 *(Processus adapté)*
Un processus $(X_t)_{t\in\mathbb{T}}$ est dit adapté (par rapport à $\mathbb{F}$) si pour tout $t \in \mathbb{T}$, X_t est $\mathcal{F}_t$-mesurable.

Lorsqu'on veut préciser par rapport à quelle filtration le processus est adapté, on écrira $\mathbb{F}$-adapté. Un processus adapté est donc un processus dont la valeur à toute date t est révélée par l'information $\mathcal{F}_t$. On dit parfois que le processus est non anticipatif. Il est clair que tout processus X est adapté par rapport à sa filtration naturelle $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\in\mathbb{T}}$.

Jusquà présent, le processus stochastique X est vu soit comme une fonction du temps t à ω fixé (lorsqu'on parle de trajectoire) ou comme une fonction de ω à t fixé (lorsqu'on considère la variable aléatoire comme dans la définition 1.1.2). On peut considérer les deux aspects en regardant le processus comme une fonction définie sur $\mathbb{T} \times \Omega$. Ceci conduit aux définitions suivantes :

Définition 1.1.3 *(Processus progressif, optionnel, prévisible)*
1) Un processus $(X_t)_{t\in\mathbb{T}}$ est dit progressif (par rapport à $\mathbb{F}$) si pour tout $t \in \mathbb{T}$, l'application $(s,\omega) \to X_s(\omega)$ est mesurable sur $[0,t] \times \Omega$ muni de la tribu produit $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$.

2) Un processus $(X_t)_{t\in\mathbb{T}}$ est dit optionnel (par rapport à $\mathbb{F}$) si l'application $(t,\omega) \to X_s(\omega)$ est mesurable sur $\mathbb{T} \times \Omega$ muni de la tribu engendrée par les processus $\mathbb{F}$-adaptés et càd-làg.

3) Un processus $(X_t)_{t\in\mathbb{T}}$ est dit prévisible (par rapport à $\mathbb{F}$) si l'application $(t,\omega) \to X_s(\omega)$ est mesurable sur $\mathbb{T} \times \Omega$ muni de la tribu engendrée par les processus $\mathbb{F}$-adaptés et continus.

Lorsqu'on veut préciser la filtration, on écrira $\mathbb{F}$-progressif (optionnel ou prévisible). Evidemment, tout processus progressif est adapté et mesurable sur $\mathbb{T} \times \Omega$ muni de la tribu produit $\mathcal{B}(\mathbb{T}) \otimes \mathcal{F}$. Il est aussi clair avec la définition que tout processus càd-làg adapté est optionnel (la réciproque n'étant pas vraie). De même, tout processus X continu et adapté est prévisible (la réciproque n'étant pas vraie) : puisque dans ce cas $X_t = \lim_{s\nearrow t} X_s$, ceci signifie que la valeur de X_t est annoncée par ses valeurs précédentes. Puisqu'un processus continu est càd-làg, il est évident que tout processus prévisible est optionnel. Le résultat suivant donne le lien entre processus optionnel et progressif :

Proposition 1.1.1 *Si le processus X est optionnel, il est progressif. En particulier, s'il est adapté et càd-làg alors il est progressif.*

Par abus de language, on écrit souvent dans la littérature processus adapté pour processus progressif.

1.1.2 Temps d'arrêt

Ayant à l'esprit l'interprétation de $\mathcal{F}_t$ comme l'information connue jusqu'à la date t, on s'intéresse à savoir si un évènement donné, caractérisé par sa première date $\tau(\omega)$ d'apparition, a eu lieu ou non avant la date t sachant l'observation de l'information $\mathcal{F}_t$. Ceci conduit à la notion de temps d'arrêt.

Définition 1.1.4 *(Temps d'arrêt)*
1) Une variable aléatoire $\tau : \Omega \to [0, +\infty]$, i.e. un temps aléatoire, est appelé temps d'arrêt (par rapport à la filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$) si pour tout $t \in \mathbb{T}$:

$$\{\tau \le t\} := \{\omega \in \Omega \ : \tau(\omega) \le t\} \ \in \ \mathcal{F}_t.$$

2) Un temps d'arrêt τ est dit prévisible s'il existe une suite de temps d'arrêt $(\tau_n)_{n\ge 1}$ telle que l'on ait p.s. :
(i) $\lim_n \tau_n = \tau$
(ii) $\tau_n < \tau$ pour tout n sur $\{\tau > 0\}$.
On dit alors que $(\tau_n)_{n\ge 1}$ annonce τ.

On vérifie aisément avec la définition que tout temps aléatoire égal à une constante positive t est un temps d'arrêt. On note aussi que si τ et σ sont deux temps d'arrêt alors $\tau \wedge \sigma$, $\tau \vee \sigma$ et $\tau + \sigma$ sont des temps d'arrêt.

Etant donné un temps d'arrêt τ, on mesure l'information accumulée jusqu'en τ par :

$$\mathcal{F}_\tau = \{B \in \mathcal{F}_{\bar{T}} \ : B \cap \{\tau \le t\} \in \mathcal{F}_t, \ \ \forall t \in \mathbb{T}\},$$

qui est une tribu de $\mathcal{F}$. Il est clair que τ est $\mathcal{F}_\tau$-mesurable. On a aussi immédiatement que si $\tau = t$ alors $\mathcal{F}_\tau = \mathcal{F}_t$. On énonce quelques autres propriétés élémentaires et utiles pour la suite sur les temps d'arrêt (voir par exemple les preuves dans le ch. I, sec. 1.2 de Karatzas et Shreve [KaSh88]).

Proposition 1.1.2 *Soient σ et τ des temps d'arrêt et ξ une variable aléatoire.*
(1) Pour tout $B \in \mathcal{F}_\sigma$, on a $B \cap \{\sigma \le \tau\} \in \mathcal{F}_\tau$. En particulier, si $\sigma \le \tau$ alors $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.
(2) Les évènements suivants

$$\{\sigma < \tau\}, \ \{\sigma \le \tau\}, \ \{\sigma = \tau\}$$

appartiennent à $\mathcal{F}_{\sigma\wedge\tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.
(3) ξ est $\mathcal{F}_\tau$-mesurable si et seulement si pour tout $t \in \mathbb{T}$, $\xi 1_{\tau\le t}$ est $\mathcal{F}_t$-mesurable.

Etant donné un processus $(X_t)_{t\in\mathbb{T}}$ et un temps d'arrêt τ, on définit la variable aléatoire X_τ sur $\{\tau \in \mathbb{T}\}$ par :

$$X_\tau(\omega) = X_{\tau(\omega)}(\omega).$$

On vérifie que si X est mesurable alors X_τ est une variable aléatoire sur $\{\tau \in \mathbb{T}\}$. On introduit alors le processus arrêté (en τ) X^τ défini par :

$$X_t^\tau = X_{\tau\wedge t}, \quad t \in \mathbb{T}.$$

Proposition 1.1.3 *Soit $(X_t)_{t\in\mathbb{T}}$ un processus progressif et τ un temps d'arrêt. Alors $X_\tau 1_{\tau\in\mathbb{T}}$ est $\mathcal{F}_\tau$-mesurable et le processus arrêté X^τ est progressif.*

On termine ce paragraphe par un résultat donnant une classe importante de temps d'arrêt.

Proposition 1.1.4 *Soit X une processus càd-làg adapté et Γ un sous-ensemble ouvert de $\mathcal{X} = \mathbb{R}^d$.*

(1) Si la filtration $\mathbb{F}$ satisfait les conditions habituelles, alors le temps d'atteinte de Γ défini par

$$\sigma_{\Gamma} = \inf\{t \geq 0\ : X_t \in \Gamma\}$$

(avec la convention $\inf \emptyset = +\infty$) est un temps d'arrêt.

(2) Si X est continu, alors le temps de sortie de Γ défini par

$$\tau_{\Gamma} = \inf\{t \geq 0\ : X_t \notin \Gamma\}$$

est un temps d'arrêt prévisible.

1.1.3 Mouvement brownien

L'exemple basique de processus est le mouvement brownien, nom donné par le botaniste Robert Brown en 1827 pour décrire le mouvement irrégulier de particules de pollen dans un fluide. Le cadre d'application du mouvement brownien a largement dépassé l'étude des particules microscopiques pour être utilisé en finance dans la modélisation des prix d'actions, historiquement depuis Bachelier en 1900.

Définition 1.1.5 *(Mouvement brownien standard)*
Un mouvement brownien standard vectoriel (d-dimensionnel) sur $\mathbb{T} = [0, T]$ ou $\mathbb{R}_+$ est un processus continu à valeurs dans $\mathbb{R}^d$, $(W_t)_{t\in\mathbb{T}} = (W_t^1, \ldots, W_t^d)_{t\in\mathbb{T}}$ tel que

(i) $W_0 = 0$

(ii) Pour tous $0 \leq s < t$ dans $\mathbb{T}$, l'accroissement $W_t - W_s$ est indépendant de $\sigma(W_u, u \leq s)$ et suit une loi gaussienne centrée de matrice de variance-covariance $(t-s)I_d$ où I_d est la matrice identité $d \times d$.

Une conséquence immédiate de la définition est que les coordonnées $(W_t^i)_{t\in\mathbb{T}}$, $i = 1, \ldots, d$, d'un mouvement brownien vectoriel sont des mouvements browniens réels et indépendants. Réciproquement des mouvements browniens réels indépendants engendrent un mouvement brownien vectoriel. Dans la définition d'un mouvement brownien standard, l'indépendance des accroissements est par rapport à la filtration naturelle $\mathcal{F}_s^W = \sigma(W_u, u \leq s)$ de W. La filtration naturelle de W est parfois appelée filtration brownienne. Il est souvent intéressant de travailler avec une filtration plus grande que la filtration naturelle. Ceci conduit à la définition plus générale.

Définition 1.1.6 *(Mouvement brownien par rapport à une filtration)*
Un mouvement brownien vectoriel (d-dimensionnel) sur $\mathbb{T} = [0, T]$ *ou* $\mathbb{R}_+$ *par rapport à une filtration* $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ *est un processus continu* $\mathbb{F}$*-adapté, à valeurs dans* $\mathbb{R}^d$, $(W_t)_{t \in \mathbb{T}} = (W_t^1, \ldots, W_t^d)_{t \in \mathbb{T}}$ *tel que*

(i) $W_0 = 0$

(ii) Pour tous $0 \leq s < t$ *dans* $\mathbb{T}$*, l'accroissement* $W_t - W_s$ *est indépendant de* $\mathcal{F}_s$ *et suit une loi gaussienne centrée de matrice de variance-covariance* $(t - s)I_d$ *où* I_d *est la matrice identité* $d \times d$.

Bien entendu, un mouvement brownien standard est un mouvement brownien par rapport à sa filtration naturelle.

Un problème majeur est celui de l'existence et de la construction et simulation d'un mouvement brownien. Nous ne discutons pas ici de ce problème et renvoyons le lecteur aux nombreux livres traitant du sujet (par exemple Hida [Hi80], Karatzas et Shreve [KaSh88], Le Gall [LeG89] ou Revuz et Yor [ReY91]). Nous énonçons seulement une propriété classique du mouvement brownien.

Proposition 1.1.5 *Soit* $(W_t)_{t \in \mathbb{T}}$ *est un mouvement brownien par rapport à* $(\mathcal{F}_t)_{t \in \mathbb{T}}$.
(1) Symétrie : $(-W_t)_{t \in \mathbb{T}}$ *est aussi un mouvement brownien.*

(2) Echelle : Pour tout $\lambda > 0$*, le processus* $((1/\lambda)W_{\lambda^2 t})_{t \in \mathbb{T}}$ *est aussi un mouvement brownien.*

(3) Invariance par translation : Pour tout $s > 0$*, le processus* $(W_{t+s} - W_s)_{t \in \mathbb{T}}$ *est un mouvement brownien standard indépendant de* $\mathcal{F}_s$.

Nous rappelons aussi que l'augmentation habituelle de la filtration naturelle $(\mathcal{F}_t^W)_t$ d'un mouvement brownien W est $(\sigma(\mathcal{F}_t^W \cup \mathcal{N}))_t$ où $\mathcal{N}$ est l'ensemble des évènements négligeables de $(\Omega, \mathcal{F}_{\bar{T}}, P)$. De plus, W reste un mouvement brownien par rapport à sa filtration augmentée. Par abus de language, l'augmentation de la filtration naturelle de W est encore appelée filtration naturelle de W ou filtration brownienne.

1.1.4 Martingales, semimartingales

Dans cette section, on considère des processus à valeurs réelles. Les preuves des résultats énoncés dans ce paragraphe sont établies par exemple dans Dellacherie et Meyer [DM80].

Définition 1.1.7 *(Martingale)*
Un processus $(X_t)_{t \in \mathbb{T}}$ *adapté est appelé surmartingale si* $E[X_t^-] < +\infty$ *pour tout* $t \in \mathbb{T}$ *et*

$$E[X_t | \mathcal{F}_s] \leq X_s, \quad p.s. \quad \text{pour tous } \ 0 \leq s \leq t, \ s, t \in \mathbb{T}. \tag{1.1}$$

X *est une sous-martingale si* $-X$ *est une surmartingale. On dit que* X *est une martingale si elle est à la fois une surmartingale et une sous-martingale.*

La définition d'une sur(sous)martingale dépend crucialement de la probabilité P et de la filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ spécifiées sur l'espace mesurable $(\Omega, \mathcal{F})$. Dans ce cours, la filtration sera fixée et lorsque ce n'est pas précisé, la propriété de sur(sous)martingale sera toujours relative à la filtration. On sera amené, par contre, à considérer différentes probabilités Q sur $(\Omega, \mathcal{F})$ et pour souligner ce fait, on précisera alors Q-sur(sous)martingale.

Un exemple important de martingale est le mouvement brownien décrit dans le paragraphe précédent. D'autre part, une construction typique de martingales est la suivante : on se donne une variable aléatoire ξ sur $(\Omega, \mathcal{F})$ intégrable, i.e. $E|\xi| < +\infty$. Alors, le processus défini par

$$X_t = E[\xi|\mathcal{F}_t], \quad t \in \mathbb{T},$$

est clairement une martingale. On dit que X est fermée à droite par ξ. Réciproquement, lorsque $\mathbb{T} = [0, T]$, $T < +\infty$, toute martingale $(X_t)_{t\in[0,T]}$ est fermée à droite par $\xi = X_T$. Lorsque $\mathbb{T} = [0, +\infty[$, la fermeture à droite d'une martingale est donnée par le résultat de convergence suivant :

Théorème 1.1.1 *(Convergence des martingales)*
1) Soit $X = (X_t)_{t\geq 0}$ une surmartingale càd-làg bornée dans L^1 (en particulier si elle est positive). Alors X_t converge p.s. quand $t \to +\infty$.

2) Soit $X = (X_t)_{t\geq 0}$ une martingale càd-làg. Alors $(X_t)_{t\geq 0}$ est uniformément intégrable si et seulement si X_t converge p.s. et dans L^1 quand $t \to +\infty$ vers une variable aléatoire $X_{+\infty}$. Dans ce cas, $X_{+\infty}$ ferme X à droite, i.e. $X_t = E[X_{+\infty}|\mathcal{F}_t]$ pour tout $t \geq 0$.

Dans la suite, on notera $\bar{\mathbb{T}}$ l'intervalle égal à $[0, T]$ si $\mathbb{T} = [0, T]$ et égal à $[0, +\infty]$ si $\mathbb{T} = [0, +\infty[$. On note aussi $\bar{T}$ le bord à droite de $\mathbb{T}$. Avec cette convention, si $(X_t)_{t\in\mathbb{T}}$ est une martingale càd-làg uniformément intégrable alors $X_{\bar{T}}$ est la limite p.s. et dans L^1 de X_t quand t tend vers $\bar{T}$. De plus, $X_{\bar{T}}$ ferme à droite X : $X_t = E[X_{\bar{T}}|\mathcal{F}_t]$ pour tout $t \in \mathbb{T}$.

Le résultat suivant est une propriété très importante des martingales : il étend la propriété (1.1) pour des dates t et s remplacées par des temps d'arrêt.

Théorème 1.1.2 *(Théorème d'arrêt des martingales)*
Soit $M = (M_t)_{t\in\mathbb{T}}$ une martingale càd-làg et σ, τ deux temps d'arrêt bornés à valeurs dans $\mathbb{T}$ tel que $\sigma \leq \tau$. Alors

$$E[M_\tau|\mathcal{F}_\sigma] = M_\sigma, \quad p.s.$$

Une application utile de ce théorème d'arrêt est donnée dans le corollaire suivant :

Corollaire 1.1.1 *Soit $X = (X_t)_{t\in\mathbb{T}}$ un processus càd-làg adapté.*

(1) X est une martingale si et seulement si pour tout temps d'arrêt τ borné à valeurs dans $\mathbb{T}$, on a $X_\tau \in L^1$ et

$$E[X_\tau] = X_0.$$

(2) Si X est une martingale et τ est un temps d'arrêt alors le processus arrêté X^τ est une martingale.

Nous énonçons une première inégalité fondamentale pour les martingales.

Théorème 1.1.3 *(Inégalité de Doob)*
Soit $X = (X_t)_{t\in\mathbb{T}}$ une sousmartingale positive ou une martingale, càd-làg. Alors pour tout temps d'arrêt τ à valeurs dans $\mathbb{T}$, on a :

$$P\left[\sup_{0\le t\le\tau}|X_t| \ge \lambda\right] \le \frac{E|X_\tau|}{\lambda}, \quad \forall\lambda > 0,$$

$$E\left[\sup_{0\le t\le\tau}|X_t|\right]^p \le \left(\frac{p}{p-1}\right)^p E\big[|X_\tau|^p\big], \quad \forall p > 1.$$

Notons que la première inégalité ci-dessus et le théorème de convergence des martingales impliquent que si $(X_t)_{t\in\mathbb{T}}$ est une martingale càd-làg uniformément intégrable, alors $\sup_{t\in\mathbb{T}}|X_t| < +\infty$ p.s.

Dans toute la suite, on se place sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}, P)$ satisfaisant les conditions habituelles.

En théorie des processus, le concept de localisation est très utile. De manière générale, on dit qu'un processus progressif X est localement "truc" (ou possède la propriété "truc" locale) s'il existe une suite croissante de temps d'arrêt $(\tau_n)_{n\ge1}$ (appelée suite localisante) telle que τ_n tend p.s. vers l'infini et pour tout n, le processus arrêté X^{τ_n} est "truc" (ou possède la propriété "truc"). On introduit ainsi la notion de processus localement borné et on note que tout processus adapté X continu est localement borné : prendre pour suite de temps d'arrêt localisante $\tau_n = \inf\{t \ge 0 \ : |X_t| \ge n\}$ de telle sorte que X^{τ_n} est bornée par n. Remarquons que lorsque X n'est pas continu avec des sauts non bornés, X n'est pas localement borné. Un exemple important de localisation est donné dans la définition suivante.

Définition 1.1.8 *(Martingale locale)*
Soit X un processus càd-làg adapté. On dit que X est une martingale locale s'il existe une suite de temps d'arrêt $(\tau_n)_{n\ge1}$ telle que $\lim_{n\to+\infty}\tau_n = +\infty$ p.s. et le processus arrêté X^{τ_n} est une martingale pour tout n.

Toute martingale càd-làg est une martingale locale mais la réciproque n'est pas vraie : les martingales locales sont beaucoup plus générales que les martingales et on en rencontrera en particulier avec les intégrales stochastiques (voir section 1.2). Il est intéressant d'avoir des conditions assurant qu'une martingale locale est une martingale. Le critère suivant est très utile en pratique.

Proposition 1.1.6 *Soit $M = (M_t)_{t\in\mathbb{T}}$ une martingale locale. Supposons que*

$$E\left[\sup_{0\le s\le t}|M_s|\right] < +\infty, \quad \forall t \in \mathbb{T}. \tag{1.2}$$

Alors M est une martingale.

En fait, on a une condition nécessaire et suffisante pour qu'une martingale locale M soit une "vraie" martingale : c'est la condition dite (DL) que la famille $(M_\tau)_\tau$ où τ parcourt l'ensemble des temps d'arrêt bornés dans $\mathbb{T}$, est uniformément intégrable. La condition suffisante (1.2) est souvent utilisée en pratique pour assurer la condition (DL). Citons aussi le résultat utile suivant conséquence directe du lemme de Fatou.

Proposition 1.1.7 *Soit M une martingale locale positive telle que $M_0 \in L^1$. Alors M est une surmartingale.*

Nous définissons à présent le concept important de variation quadratique d'une martingale locale (continue). On dit qu'un processus $A = (A_t)_{t\in\mathbb{T}}$ est à variation finie si toutes ses trajectoires sont càd-làg et à variation finie, i.e. pour tout $\omega \in \Omega$, $t \in \mathbb{T}$,

$$\sup \sum_{i=1}^{n} \left|A_{t_i}(\omega) - A_{t_{i-1}}(\omega)\right| < +\infty, \tag{1.3}$$

où le supremum porte sur toutes les subdivisions $0 = t_0 < t_1 < \ldots < t_n = t$ de $[0,t]$. Le processus A est croissant si toutes ses trajectoires sont càd-làg et croissantes. Tout processus A à variation finie peut s'écrire comme $A = A^+ - A^-$ où A^+ et A^- sont deux processus croissants. Il y a unicité d'une telle décomposition si on impose que les mesures positives associées $A^+([0,t]) = A^+_t$ et $A^-([0,t]) = A^-_t$ soient portées par des Boréliens disjoints. On note A la mesure signée différence des deux mesures positives finies A^+ et A^-. La mesure positive (aléatoire) associée au processus croissant $A^+ + A^-$ est notée $|A|$: $|A|([0,t]) = A^+_t + A^-_t$ égale à (1.3), et est appelée variation de A. Pour tout processus α tel que

$$\int_0^t |\alpha_s(\omega)| d|A|_s(\omega) < +\infty, \quad \forall t \in \mathbb{T},\ \forall \omega \in \Omega,$$

le processus $\int \alpha dA$ défini par l'intégrale de Stieltjes $\int_0^t \alpha_s(\omega) dA_s(\omega)$, pour tout $t \in \mathbb{T}$ et $\omega \in \Omega$, est à variation finie. De plus si A est adapté et α est progressif, alors $\int \alpha dA$ est adapté. Pour tout processus A à variation finie, on note A^c :

$$A^c_t = A_t - \sum_{0\le s\le t} \Delta A_s, \quad \text{où } \Delta A_s = A_s - A_{s-} \quad (\Delta A_0 = A_0).$$

A^c est un processus continu à variation finie et est appelé partie continue de A.

Proposition 1.1.8 *Soit M une martingale locale continue, $M_0 = 0$. Alors si M est à variation finie, M est indistinguable de* 0.

Théorème 1.1.4 *(Variation quadratique, crochet)*
(1) Soient $M = (M_t)_{t\in\mathbb{T}}$ et $N = (N_t)_{t\in\mathbb{T}}$ deux martingales locales dont l'une des deux est localement bornée (par exemple continue). Alors il existe un unique processus prévisible à variation finie, noté $< M,N >$, nul en 0, tel que $MN- < M,N >$ soit une martingale locale. Cette martingale locale est continue si M et N le sont. De plus, pour tout $t \in \mathbb{T}$, si $0 = t_0^n < t_1^n < \dots t_{k_n}^n = t$ est une subdivision de $[0,t]$ de pas tendant vers 0, on a :

$$< M,N >_t = \lim_{n\to+\infty} \sum_{i=1}^{k_n} \left(M_{t_i^n} - M_{t_{i-1}^n}\right)\left(N_{t_i^n} - N_{t_{i-1}^n}\right),$$

au sens de la convergence en probabilité. Le processus $< M,N >$ est appelé crochet (oblique) de M et N. On dira de plus que M et N sont orthogonales si $< M,N > = 0$ ce qui signifie que le produit MN est une martingale locale.
(2) Lorsque $M = N$, le processus $< M,M >$, noté parfois $< M >$ et appelé la variation quadratique de M ou le processus croissant de M, est croissant. De plus, on a la relation de "polarisation" :

$$< M,N > = \frac{1}{2}\left(< M+N, M+N > - < M,M > - < N,N >\right).$$

Exemple
Si $W = (W^1,\dots,W^d)$ est un mouvement brownien d-dimensionnel, on a :

$$< W^i, W^j >_t = \delta_{ij} t$$

où $\delta_{ij} = 1$ si $i = j$ et 0 sinon. De plus les processus $W_t^i W_t^j - \delta_{ij} t$ sont des martingales.

L'inégalité suivante est utile pour définir plus bas l'intégrale stochastique.

Proposition 1.1.9 *(Inégalité de Kunita-Watanabe)*
Soient M et N deux martingales locales continues et α, β deux processus mesurables sur $\mathbb{T}\times\Omega$ muni de la tribu produit $\mathcal{B}(\mathbb{T})\otimes\mathcal{F}$. Alors on a pour tout $t \in \mathbb{T}$:

$$\begin{aligned}&\int_0^t |\alpha_s||\beta_s| d| < M,N > |_s \\ &\quad \le \left(\int_0^t \alpha_s^2 d < M,M >_s\right)^{\frac{1}{2}} \left(\int_0^t \beta_s^2 d < N,N >_s\right)^{\frac{1}{2}} \quad p.s.\end{aligned}$$

Ceci montre en particulier que p.s. la mesure signée $d < M,N >$ est absolument continue par rapport à la mesure positive $d < M >$.

L'inégalité fondamentale suivante pour les martingales (locales) sera très utile lorsqu'on s'intéressera aux martingales locales définies par des intégrales stochastiques pour lesquelles on arrive souvent en pratique à calculer la variation quadratique.

Théorème 1.1.5 *(Inégalité de Burkholder-Davis-Gundy)*
Pour tout $p > 0$, il existe des constantes positives c_p et C_p telles que pour toute martingale locale continue $M = (M_t)_{t\in\mathbb{T}}$ et tout temps d'arrêt τ à valeurs dans $\bar{\mathbb{T}}$, on ait :

$$c_p E[< M >_\tau^{p/2}] \leq \ E\left[\sup_{0\leq t<\tau} |M_t|\right]^p \ \leq C_p E\left[< M >_\tau^{p/2}\right].$$

En combinant l'inégalité de Burkholder-Davis-Gundy avec la condition (1.2), on voit en particulier avec $p = 1$ que si la martingale locale continue M vérifie $E[\sqrt{< M >_t}] < +\infty$ pour tout $t \in \mathbb{T}$, alors M est une martingale.

On dit qu'une martingale càd-làg $M = (M_t)_{t\in\mathbb{T}}$ est de carré intégrable si $E[|M_t|^2] < +\infty$ pour tout $t \in \mathbb{T}$. On introduit la distinction supplémentaire (dans le cas $\mathbb{T} = [0, +\infty[$) pour dire que M est bornée dans L^2 si $\sup_{t\in\mathbb{T}} E[|M_t|^2] < +\infty$. En particulier, une martingale bornée dans L^2 est uniformément intégrable et admet une limite p.s. $M_{\bar{T}}$ quand t tend vers $\bar{T}$. On note $\mathbb{H}^2_c$ l'ensemble des martingales continues bornées dans L^2. Le résultat suivant est une conséquence des inégalités de Doob et Burkholder-Davis-Gundy.

Proposition 1.1.10 *(Martingale de carré intégrable)*
Soit $M = (M_t)_{t\in\mathbb{T}}$ une martingale locale continue. Alors M est une martingale de carré intégrable si et seulement si $E[< M >_t] < +\infty$ pour tout $t \in \mathbb{T}$. Dans ce cas, $M^2 - < M >$ est une martingale continue et si $M_0 = 0$, on a :

$$E[M_t^2] = E[< M >_t], \quad \forall t \in \mathbb{T}.$$

De plus, M est bornée dans L^2 si et seulement si $E[< M >_{\bar{T}}] < +\infty$ et dans ce cas :

$$E[M_{\bar{T}}^2] = E[< M >_{\bar{T}}].$$

L'espace $\mathbb{H}^2_c$ muni du produit scalaire

$$(M, N)_{\mathbb{H}^2} = E[< M, N >_{\bar{T}}]$$

est un espace de Hilbert.

Le théorème suivant est connu sous le nom de théorème de décomposition de Doob-Meyer des surmartingales.

Théorème 1.1.6 *(Décomposition de Doob-Meyer)*
Soit X une surmartingale càd-làg. Alors X admet une décomposition unique de la forme

$$X = X_0 + M - A$$

où M est une martingale locale càd-làg nulle en 0 et A est un processus prévisible, croissant et nul en 0. Si X est positif, alors A est intégrable, i.e. $E[A_{\bar{T}}] < +\infty$ où $A_{\bar{T}} = \lim_{t\to\bar{T}} A_t$ p.s.

On introduit finalement une classe fondamentale de processus à variation quadratique finie, étendant les sur(sous)martingales (locales), et très largement utilisé dans la modélisation en finance, en particulier dans le cadre de ce livre.

Définition 1.1.9 *(Semimartingale)*
Une semimartingale est un processus càd-làg adapté X admettant une décomposition de la forme :

$$X = X_0 + M + A \tag{1.4}$$

où M est une martingale locale càd-làg nulle en 0 et A est un processus adapté à variation finie et nul en 0. Une semimartingale continue est une semimartingale telle que dans la décomposition (1.4), M et A sont continus. Une telle décomposition où M et A sont continus, est unique.

On définit le crochet droit d'une semimartingale continue $X = X_0 + M + A$ par : $< X, X > = < M, M >$ et on a la propriété que pour tout $t \in \mathbb{T}$, si $0 = t_0^n < t_1^n < \ldots t_{k_n}^n = t$ est une subdivision de $[0, t]$ de pas tendant vers 0, on a la convergence en probabilité :

$$< X, X >_t = \lim_{n\to+\infty} \sum_{i=1}^{k_n} \left(X_{t_i^n} - X_{t_{i-1}^n} \right)^2 .$$

Cette propriété est très importante car elle montre que le crochet droit ne change pas lorsqu'on change P par une probabilité Q absolument continue sous laquelle X est encore une Q-semimartingale.

Les principaux théorèmes énoncés ci-dessus pour les sur(sous)martingales supposent des trajectoires càd-làg du processus. Le théorème suivant donnent les conditions pour lesquelles c'est le cas.

Théorème 1.1.7 *Soit $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ une filtration satisfaisant les conditions habituelles et $X = (X_t)_{t\in\mathbb{T}}$ une surmartingale. Alors X admet une modification càd-làg si et seulement si l'application $t \in \mathbb{T} \to E[X_t]$ est continue à droite (c'est le cas en particulier si X est une martingale). De plus, dans ce cas, cette modification càd-làg reste une surmartingale par rapport à $\mathbb{F}$.*

Pour une preuve de ce résultat, voir Karatzas et Shreve [KaSh88], Théorème 3.13 du Ch. 1.

Dans la suite, par définition, on dit qu'un processus vectoriel $X = (X^1, \ldots, X^d)$ est une martingale (resp. surmartingale, resp. semimartingale) (locale) si chacune de ses composantes réelles X^i, $i = 1, \ldots, d$ est une martingale (resp. surmartingale, resp. semimartingale) (locale). On définit aussi le crochet matriciel d'une semimartingale continue vectorielle $X = (X^1, \ldots, X^d)$: $< X > = < X, X' >$, par ses composantes $< X^i, X^j >$, $i, j = 1, \ldots, d$.

1.2 Intégrale stochastique et applications

1.2.1 Intégrale stochastique par rapport à une semimartingale continue

Dans cette section, on définit l'intégrale stochastique par rapport à une semimartingale continue X. On considère dans un premier temps le cas où X est unidimensionnel. Avec la décomposition (1.4), on définit d'abord l'intégrale par rapport à X comme la somme de deux intégrales, l'une par rapport à la partie à variation finie A et l'autre par rapport à la partie martingale locale continue M. L'intégrale par rapport à A est définie trajectoriellement (pour presque tout ω) comme une intégrale de Stieljes. Par contre, si la martingale M n'est pas nulle, elle n'est pas à variation finie, et on ne peut plus définir l'intégrale par rapport à M de manière trajectorielle comme des intégrales de Stieljes. La notion d'intégrale stochastique par rapport à M, due à Itô dans le cas où M est un mouvement brownien, est basée sur l'existence d'une variation quadratique $< M >$, qui permet de définir l'intégrale comme limite de suites simples de type Riemann dans L^2.

Un processus simple (ou élémentaire) est un processus de la forme

$$\alpha_t = \sum_{k=1}^{n} \alpha_{(k)} 1_{]t_k, t_{k+1}]}(t), \tag{1.5}$$

où $0 \leq t_0 < t_1 < \ldots < t_n$ est une suite de temps d'arrêt dans $\mathbb{T}$ et $\alpha_{(k)}$ est une variable $\mathcal{F}_{t_k}$-mesurable et bornée pour chaque k. On note $\mathcal{E}$ l'ensemble des processus simples. Lorsque M est bornée dans L^2, i.e. $M \in \mathbb{H}^2_c$, on définit $L^2(M)$ comme l'ensemble des processus progressifs α tels que $E[\int_0^{\bar{T}} |\alpha_t|^2 d < M >_t] < +\infty$. C'est un espace de Hilbert pour le produit scalaire :

$$(\alpha, \beta)_{L^2(M)} = E\left[\int_0^{\bar{T}} \alpha_t \beta_t d < M >_t\right],$$

et l'espace des processus simples est dense dans $L^2(M)$. L'intégrale stochastique d'un processus simple (1.5) est définie par :

$$\int_0^t \alpha_s dM_s = \sum_{k=1}^{n} \alpha_{(k)}.(M_{t_{k+1} \wedge t} - M_{t_k \wedge t}), \quad t \in \mathbb{T},$$

et appartient à $\mathbb{H}_c^2$. De plus, on a la relation d'isométrie :

$$E\left[\left(\int_0^{\bar{T}} \alpha_t \, dM_t\right)^2\right] = E\left[\int_0^{\bar{T}} |\alpha_t|^2 d < M >_t\right]$$

Par densité de $\mathcal{E}$ dans $L^2(M)$, on peut donc étendre l'application $\alpha \to \int \alpha dM$ à une isométrie de $L^2(M)$ dans $\mathbb{H}_c^2$. Pour tout $\alpha \in L^2(M)$, l'intégrale stochastique $\int \alpha dM$ est caractérisée par la relation :

$$< \int \alpha dM, N > = \int \alpha \, d < M, N >, \quad \forall N \in \mathbb{H}_c^2.$$

De plus si τ est un temps d'arrêt, on a

$$\int_0^{t\wedge\tau} \alpha_s dM_s = \int_0^t \alpha_s 1_{[0,\tau]} dM_s \ = \int_0^t \alpha_s dM_{s\wedge\tau}, \quad t \in \mathbb{T}.$$

Cette identité permet d'étendre par localisation la définition de l'intégrale stochastique à une martingale locale continue et pour une classe plus large d'intégrands.

Soit M une martingale locale continue. On note $L^2_{loc}(M)$ l'ensemble des processus progressifs α tels que pour tout $t \in \mathbb{T}$:

$$\int_0^t |\alpha_s|^2 d < M >_s < +\infty, \quad p.s.$$

Pour tout $\alpha \in L^2_{loc}(M)$, il existe une unique martingale locale continue nulle en 0, appelé intégrale stochastique de α par rapport à M et notée $\int \alpha dM$, telle que :

$$< \int \alpha dM, N > = \int \alpha \, d < M, N >$$

pour toute martingale locale continue N. Cette définition étend celle vue ci-dessus lorsque $M \in \mathbb{H}_c^2$ et $\alpha \in L^2(M)$.

Lorsqu'on considère l'intégrale stochastique par rapport à une martingale locale continue M vectorielle, une première idée est de prendre la somme des intégrales stochastiques par rapport à chacune des composantes. Cependant, pour obtenir des “bonnes” propriétés de représentation et de décomposition des martingales (voir paragraphe 1.2.4), on doit construire l'intégrale stochastique en utilisant une relation d'isométrie comme dans le cas unidimensionnel afin d'avoir les propriétés de fermeture adéquates. On doit donc élargir la classe des intégrands à une classe plus large que celle des intégrands prise composante par composante. On parle d'intégration stochastique vectorielle. Des conditions suffisantes assurant que les deux notions d'intégration stochastique vectorielle et composante par composante sont identiques, sont étudiées dans [CS93]. C'est par exemple le cas si M est un mouvement brownien d-dimensionnel. D'autre part, lorsqu'on définit l'intégration stochastique par rapport à une semimartingale, s'il est naturel et suffisant en pratique dans les applications de considérer dans un premier temps les intégrands par rapport à la variation finie et par rapport à la partie martingale, cette classe d'intégrands n'est pas assez large en théorie lorsqu'on s'intéresse aux propriétés de fermeture et de stabilité des intégrales stochastiques. La topologie des semimartingales a été introduite par Emery [Em79] et est très appropriée pour l'étude des intégrales stochastiques.

Les trois paragraphes suivants, marqués d'un astérique et concernant l'intégration stochastique vectorielle et la topologie des semimartingales peuvent être omis en première lecture. Nous commençons par définir l'intégrale par rapport à un processus vectoriel à variation finie, puis par rapport à une martingale locale continue vectorielle. La présentation est inspirée de Jacod [Jac79], ch. IV.

Intégrale de Stieltjes par rapport à un processus vectoriel à variation finie*

Soit $A = (A^1, \ldots, A^d)$ un processus vectoriel dont les composantes sont continues à variation finie. Il existe alors un processus croissant Γ (par exemple $d\Gamma = \sum_{i=1}^d d|A^i|$) et un processus vectoriel $\gamma = (\gamma^1, \ldots, \gamma^d)$ prévisible (obtenu par le théorème de Radon-Nikodym) tels que :

$$A^i = \int \gamma^i d\Gamma, \quad i = 1, \ldots, d. \tag{1.6}$$

On note $L_S(A)$ l'ensemble des processus progressifs $\alpha = (\alpha^1, \ldots, \alpha^d)$ tels que :

$$\int_0^t \left| \sum_{i=1}^d \alpha_u^i \gamma_u^i \right| d\Gamma_u < +\infty, \quad p.s.\ \forall t \in \mathbb{T}.$$

On pose alors

$$\int \alpha dA = \int \sum_{i=1}^d \alpha^i \gamma^i d\Gamma,$$

avec la convention que $\int_0^t \alpha_u(\omega) dA_u(\omega) = 0$ si ω est dans l'ensemble négligeable où $\int_0^t \left| \sum_{i=1}^d \alpha_u^i \gamma_u^i(\omega) \right| d\Gamma_u(\omega) = +\infty$. Le processus $\int \alpha dA$ ne dépend pas du choix de (γ, Γ) vérifiant (1.6) et est aussi continu à variation finie. Notons que $L_S(A)$ contient (et en général strictement) l'ensemble des processus progressifs $\alpha = (\alpha^1, \ldots, \alpha^d)$ tels que pour tout $i = 1, \ldots, d$, $\alpha^i \in L_S(A^i)$, i.e. $\int_0^t |\alpha_u^i| d|A^i|_u < +\infty$ pour tout t dans $\mathbb{T}$ pour lesquels on a : $\int \alpha dA = \sum_{i=1}^d \int \alpha^i dA^i$ qu'on notera souvent $\int \alpha' dA$.

Intégrale stochastique par rapport à une martingale locale continue vectorielle*

Soit $M = (M^1, \ldots, M^d)$ une martingale locale continue à valeurs dans $\mathbb{R}^d$. On note $< M >$ le crochet matriciel de composantes $< M^i, M^j >$. Il existe alors un processus prévisible croissant C tel que $d < M^i, M^j >$ soit absolument continue par rapport à la mesure positive dC : par exemple $C = \sum_{i=1}^d < M^i >$. Par le théorème de Radon-Nikodym, il existe un processus prévisible c à valeurs dans l'espace des matrices symétriques définies non-négatives $d \times d$ tel que :

$$< M > = \int c\, dC, \quad i.e. \;< M^i, M^j > = \int c^{ij} dC, \;\; i, j = 1, \ldots, d. \tag{1.7}$$

On définit $L^2_{loc}(M)$ comme l'ensemble de tous les processus progressifs $\alpha = (\alpha^1, \ldots, \alpha^d)$ à valeurs dans $\mathbb{R}^d$ tels que pour tout $t \in \mathbb{T}$:

$$\int_0^t \alpha'_u d < M >_u \alpha_u := \int_0^t \alpha'_u c_u \alpha_u dC_u \;<\; +\infty, \quad p.s. \tag{1.8}$$

Il est à noter que l'expression $\int_0^t \alpha'_u d < M >_u \alpha_u$ ne dépend pas du choix de c et C vérifiant (1.7). Pour tout $\alpha \in L^2_{loc}(M)$, il existe une unique martingale locale continue nulle en 0, appelé intégrale stochastique de α par rapport à M et notée $\int \alpha dM$, telle que :

$$< \int \alpha dM, N > = \int \alpha \; d < M, N >$$

pour toute martingale locale continue N. L'intégrale dans le terme à droite de l'égalité ci-dessus est l'intégrale de Stieltjes par rapport au processus vectoriel à variation finie $< M, N > = (< M^1, N >, \ldots, < M^d, N >)$. Cette intégrale de Stieltjes est bien fini d'après l'inégalité de Kunita-Watanabe. En particulier, on a :

$$< \int \alpha dM, \int \alpha dM > = \int \alpha' d < M > \alpha.$$

Notons que si pour tout $i = 1, \ldots, d$, $\alpha^i \in L^2_{loc}(M^i)$, i.e. $\int_0^t |\alpha^i_u|^2 d < M^i >_u$ $< +\infty$ pour tout t dans $\mathbb{T}$, alors $\alpha = (\alpha^1, \ldots, \alpha^d) \in L^2_{loc}(M)$ et $\int \alpha dM$ $= \sum_{i=1}^d \int \alpha^i dM^i$. La réciproque n'est pas vraie en général : l'ensemble $L^2_{loc}(M)$ est strictement plus grand que l'ensemble $\{\alpha = (\alpha^1, \ldots, \alpha^d) : \alpha^i \in L^2_{loc}(M^i), i = 1, \ldots, d\}$. Cependant si les M^i sont orthogonaux deux à deux, i.e. $< M^i, M^j > = 0$ pour $i \neq j$, on a égalité de ces deux ensembles et on écrira souvent à la place de $\int \alpha dM$:

$$\int \alpha' dM := \sum_{i=1}^d \int \alpha^i dM^i.$$

C'est typiquement le cas lorsque M est un mouvement brownien d-dimensionnel $W = (W^1, \ldots, W^d)$, la condition (1.8) sur $L^2_{loc}(W)$ s'écrivant alors :

$$\int_0^t |\alpha_u|^2 du \; = \; \sum_{i=1}^d \int_0^t |\alpha_u^i|^2 du < +\infty, \quad p.s., \; \forall t \in \mathbb{T}.$$

De plus, $\int \alpha dM$ est une martingale bornée dans L^2 si et seulement si $\alpha \in L^2(M)$ défini comme l'ensemble des processus progressifs α tels que $E[\int_0^{\bar{T}} \alpha_t' d < M >_t \alpha_t] < +\infty$. Dans ce cas, on a :

$$E\left[\left(\int_0^{\bar{T}} \alpha_t dM_t\right)^2\right] = E\left[\int_0^{\bar{T}} \alpha_t' d < M >_t \alpha_t\right].$$

Dans le cas général, pour $\alpha \in L^2_{loc}(M)$, une suite de temps d'arrêt localisante pour la martingale locale continue $\int \alpha dM$ est par exemple

$$\tau_n = \inf\left\{ t \in \mathbb{T} \;:\; \int_0^t \alpha_u' d < M >_u \alpha_u \; \geq \; n \right\}$$

pour laquelle l'intégrale stochastique arrêtée $\int^{.\wedge \tau_n} \alpha dM$ est une martingale bornée dans L^2 avec $E[< \int^{.\wedge\tau_n} \alpha dM >_{\bar{T}}] \leq n$. Lorsque M est un mouvement brownien d-dimensionnel W et α est un processus continu, un autre exemple de suite de temps d'arrêt localisante pour la martingale locale continue $\int \alpha dW$ est

$$\tau_n = \inf\left\{ t \in \mathbb{T} \;: |\alpha_t| \; \geq \; n \right\}.$$

Dans ce cas l'intégrale stochastique arrêtée $\int^{.\wedge\tau_n} \alpha dW$ est une martingale de carré intégrable avec $E[< \int^{.\wedge\tau_n} \alpha dW >_t] = E[\int_0^{t\wedge\tau_n} |\alpha_u|^2 du] \leq n^2 t$, pour tout $t \in \mathbb{T}$.

Intégrale stochastique par rapport à une semimartingale continue vectorielle*

Soit X une semimartingale continue vectorielle décomposée sous la forme $X = M + A$ où M est une martingale locale continue vectorielle et A un processus continu vectoriel à variation finie. Naturellement, pour tout $\alpha \in L^2_{loc}(M) \cap L_S(A)$, on peut définir l'intégrale stochastique de α par rapport à X en posant :

$$\int \alpha dX = \int \alpha dM + \int \alpha dA.$$

Notons que les processus progressifs α localement bornés, i.e. $\sup_{0 \leq s \leq t} |\alpha_s| < +\infty$ p.s. pour tout t dans $\mathbb{T}$, sont dans $L^2_{loc}(M) \cap L_S(A)$. Cette classe d'intégrands, suffisante en pratique dans les applications, n'est pas assez large en théorie lorsqu'on s'intéresse aux propriétés de fermeture et de stabilité des intégrales stochastiques. La topologie des semimartingales est très appropriée pour l'étude des intégrales stochastiques. Cette topologie est définie par la distance entre deux semimartingales (ici continues) $X = (X_t)_{t\in\mathbb{T}}$ et $Y = (Y_t)_{t\in\mathbb{T}}$:

$$D_E(X,Y) = \sup_{|\alpha|\leq 1} \left(\sum_{n\geq 1} 2^{-n} E\left[\left| \int_0^{\bar{T}\wedge n} \alpha_t dX_t - \int_0^{\bar{T}\wedge n} \alpha_t dY_t \right| \wedge 1 \right] \right),$$

où le supremum est pris sur tous les processus progressifs α bornés par 1.

Définition 1.2.10 *Soit X une semimartingale continue. Soit α un processus progressif et $\alpha^{(n)}$ le processus tronqué borné $\alpha 1_{|\alpha|\leq n}$. On dit que α est intégrable par rapport à X et on note $\alpha \in L(X)$, si la suite de semimartingales $\int \alpha^{(n)} dX$ converge pour la topologie des semimartingales vers une semimartingale Y et on pose alors $\int \alpha dX = Y$.*

On a les propriétés suivantes pour cette classe générale d'intégrands :

– Si X est une martingale locale, $L^2_{loc}(X) \subset L(X)$.
– Si X est à variation finie, $L_S(X) \subset L(X)$.
– $L(X) \cap L(Y) \subset L(X+Y)$ et $\int \alpha dX + \int \alpha dY = \int \alpha d(X+Y)$.
– $L(X)$ est un espace vectoriel et $\int \alpha dX + \int \beta dX = \int (\alpha+\beta) dX$.

De plus, l'espace $\{\int \alpha dX \ : \alpha \in L(X)\}$ est fermé dans l'espace des semimartingales pour la topologie des semimartingales. Enfin, comme la topologie des semimartingales est invariante par changement de probabilité équivalente P sur $(\Omega, \mathcal{F})$, il en est de même de $L(X)$. Attention, si X est une martingale locale continue et α est dans $L(X)$, l'intégrale stochastique $\int \alpha dX$ n'est pas toujours une martingale locale. En fait, $\int \alpha dS$ est une martingale locale si et seulement si α appartient à $L^2_{loc}(M)$. On sait aussi que lorsque le processus $\int \alpha dX$ est borné inférieurement alors c'est une martingale locale et donc aussi une surmartingale.

1.2.2 Processus d'Itô

En finance, on utilise souvent des processus d'Itô comme semimartingales continues pour modéliser la dynamique des prix d'actifs risqués.

Définition 1.2.11 *Soit $W = (W^1, \ldots, W^d)$ un mouvement brownien d-dimensionnel sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F}, P)$. On appelle processus d'Itô, un processus $X = (X^1, \ldots, X^n)$ à valeurs dans $\mathbb{R}^n$ tel que l'on ait p.s. :*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \quad t \in \mathbb{T}, \tag{1.9}$$

$$i.e. \quad X_t^i = X_0^i + \int_0^t b_s^i ds + \sum_{j=1}^d \int_0^t \sigma_s^{ij} dW_s^j, \quad t \in \mathbb{T},\ 1 \leq i \leq n,$$

où X_0 *est* $\mathcal{F}_0$*-mesurable,* $b = (b^1, \ldots, b^n)$ *et* $\sigma = (\sigma^1, \ldots, \sigma^n) = (\sigma^{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ *sont des processus progressifs à valeurs respectivement dans* $\mathbb{R}^n$ *et* $\mathbb{R}^{n \times d}$ *tels que* $b^i \in L_S(dt)$ *et* $\sigma^i \in L^2_{loc}(W)$, $i = 1, \ldots, n$, *i.e.*

$$\int_0^t |b_s| ds + \int_0^t |\sigma_s|^2 ds < +\infty, \quad p.s., \quad \forall t \in \mathbb{T}.$$

On écrira souvent (1.9) sous forme différentielle

$$dX_t = b_t dt + \sigma_t dW_t.$$

1.2.3 Formule d'Itô

1. Soit $X = (X^1, \ldots, X^d)$ une semimartingale continue à valeurs dans $\mathbb{R}^d$ et f une fonction de classe $C^{1,2}$ sur $\mathbb{T} \times \mathbb{R}^d$. Alors $(f(t, X_t))_{t \in \mathbb{T}}$ est une semimartingale et on a pour tout $t \in \mathbb{T}$:

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(u, X_u) du \ + \ \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(u, X_u) dX_u^i$$
$$+ \ \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(u, X_u) d < X^i, X^j >_u .$$

Dans cette expression, les divers intégrands sont continus donc localement bornés et les intégrales stochastiques ou de Stieltjes sont bien définies.

2. Dans ce livre, on utilisera aussi dans un exemple (voir section 4.5) la formule d'Itô pour une semimartingale $X = (X^1, \ldots, X^d)$ de la forme

$$X^i = M^i + A^i, \quad i = 1, \ldots, d,$$

où M^i est une martingale continue et A^i est un processus adapté à variation finie. On note $A^{i,c}$ la partie continue de A^i. Alors si f une fonction de classe $C^{1,2}$ sur $\mathbb{T} \times \mathbb{R}^d$, $(f(t, X_t))_{t \in \mathbb{T}}$ est une semimartingale et on a pour tout $t \in \mathbb{T}$:

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(u, X_u) du \ + \ \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(u, X_u) dM_u^i$$
$$+ \ \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(u, X_u) d < M^i, M^j >_u$$
$$+ \ \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(u, X_u) dA_u^{i,c} \ + \ \sum_{0 < s \leq t} \left[f(s, X_s) - f(s, X_{s^-}) \right],$$

1.2.4 Théorèmes de représentation de martingales

Dans ce paragraphe, nous énonçons quelques théorèmes de représentation de martingales utilisant l'intégrale stochastique.

Le premier résultat est le théorème de représentation des martingales browniennes, appelé aussi théorème de représentation d'Itô.

Théorème 1.2.8 *(Représentation des martingales browniennes)*
On suppose que $\mathbb{F}$ est la filtration naturelle (augmentée) d'un mouvement brownien standard d-dimensionnel $W = (W^1, \ldots, W^d)$. Soit $M = (M_t)_{t \in \mathbb{T}}$ une martingale locale càd-làg. Alors il existe $\alpha = (\alpha^1, \ldots, \alpha^d) \in L^2_{loc}(W)$ tel que :

$$M_t = M_0 + \int_0^t \alpha'_u dW_u \ = \ M_0 + \sum_{i=1}^d \int_0^t \alpha^i_t dW^i_t, \quad t \in \mathbb{T}, \ \ p.s.$$

De plus si M est bornée dans L^2 alors $\alpha \in L^2(W)$, i.e. $E[\int_0^{\bar{T}} |\alpha_t|^2 dt] < +\infty$.

Pour une preuve, on peut consulter le ch. 3, sec. 3.4 de Karatzas et Shreve [KaSh88] ou le ch. V de Revuz et Yor [ReY91]. Ce résultat montre en particulier que toute martingale par rapport à une filtration brownienne est continue (à un processus indistinguable près).

Le second résultat est un théorème de projection dans l'espace des intégrales stochastiques par rapport à une martingale locale continue. Il est originellement apparu dans Kunita-Watanabe [KW67], puis dans Galtchouk [Ga76]. On trouvera aussi une preuve dans Jacod [Jac79], ch. IV, sec. 2.

Théorème 1.2.9 *(Décomposition de Galtchouk-Kunita-Watanabe)*
Soit $M = (M^1, \ldots, M^d)$ une martingale locale continue à valeurs dans $\mathbb{R}^d$ et N une martingale càd-làg locale réelle. Alors il existe $\alpha \in L^2_{loc}(M)$ et R martingale càd-làg locale orthogonale à M ($< R, M^i > = 0$, $i = 1, \ldots, d$) nulle en 0 tels que :

$$N_t = N_0 + \int_0^t \alpha_u dM_u + R_t, \quad t \in \mathbb{T}, \ \ p.s.$$

Si de plus N est bornée dans L^2 alors $\alpha \in L^2(M)$ et R est aussi bornée dans L^2.

Le résultat suivant, dû à Yor [Yo78], examine le problème de la limite dans L^1 d'intégrands par rapport une martingale.

Théorème 1.2.10 *Soit M une martingale locale continue et $(\alpha^{(n)})_{n \geq 1}$ une suite de processus dans $L(M)$ telle que pour tout n, $\int \alpha^{(n)} dM$ est une martingale uniformément intégrable et la suite $(\int_0^{\bar{T}} \alpha^{(n)}_t dM_t)_{n \geq 1}$ converge dans L^1 vers une variable aléatoire $\xi \in L^1$. Alors il existe $\alpha \in L(M)$ tel que $\int \alpha dM$ soit une martingale uniformément intégrable et $\xi = \int_0^{\bar{T}} \alpha_t dM_t$.*

1.2.5 Théorème de Girsanov

Dans ce paragraphe, on s'intéresse à l'effet d'un changement de probabilité sur les notions de semimartingales et martingales. La présentation est inspirée du ch. 4, sec. 3.5 de Karatzas et Shreve [KaSh88] et du ch. VIII de Revuz et Yor [ReY91]. On se donne un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}, P)$ satisfaisant les conditions habituelles. Pour simplifier les notations, on suppose $\mathcal{F} = \mathcal{F}_{\bar{T}}$.

Soit Q une mesure de probabilité sur $(\Omega, \mathcal{F})$ telle que Q soit absolument continue par rapport à P, noté $Q \ll P$. Soit $Z_{\bar{T}}$ la densité de Radon-Nikodym correspondante, notée souvent

$$Z_{\bar{T}} \;=\; \frac{dQ}{dP} \quad \text{ou} \;\; Q \;=\; Z_{\bar{T}}.P$$

Pour tout $t \in \mathbb{T}$, la restriction de Q à $\mathcal{F}_t$ est évidemment absolument continue par rapport à la restriction de P à $\mathcal{F}_t$. Soit Z_t la densité de Radon-Nikodym correspondante, notée souvent

$$Z_t = \left.\frac{dQ}{dP}\right|_{\mathcal{F}_t}$$

Alors le processus $(Z_t)_{t\in\mathbb{T}}$ est une martingale (sous P par rapport à $(\mathcal{F}_t)_{t\in\mathbb{T}}$) positive, fermée à droite par $Z_{\bar{T}}$:

$$Z_t = E[Z_{\bar{T}}|\mathcal{F}_t], \quad p.s. \;\; \forall\, t \in \mathbb{T}.$$

Quitte à remplacer Z par une modification, on peut supposer que ses trajectoires sont càd-làg. Z est appelé processus de densité martingale de Q (par rapport à P). Z est positive, mais en fait on a

$$Z_t > 0, \quad \forall\, t \in \mathbb{T}, \;\; Q\; p.s,$$

ce qui signifie que $Q[\tau < \infty] = 0$, où $\tau = \inf\{t \;:\; Z_t = 0 \text{ ou } Z_{t^-} = 0\}$: ceci découle du fait que la martingale Z s'annule sur $[\tau, +\infty[$. Lorsque Q est équivalente à P, noté $Q \sim P$, on écrit sans ambiguité $Z_t > 0$, pour tout t dans $\mathbb{T}$, p.s. Dans la suite, on note E^Q l'opérateur d'espérance sous Q. Lorsqu'une propriété est relative à Q, on précisera la référence à Q. Lorsque cela n'est pas précisé, la propriété est implicitement relative à P la probabilité initiale sur $(\Omega, \mathcal{F})$.

Proposition 1.2.11 *(Formule de Bayes)*
Soit $Q \ll P$ et Z son processus de densité martingale. Pour tous temps d'arrêt $\sigma \leq \tau$ à valeurs dans $\mathbb{T}$, ξ variable aléatoire $\mathcal{F}_\tau$-mesurable tel que $E^Q[|\xi|] < +\infty$, on a :

$$E^Q\left[\xi|\,\mathcal{F}_\sigma\right] = \frac{E\left[Z_\tau \xi|\,\mathcal{F}_\sigma\right]}{Z_\sigma}, \qquad Q \quad p.s.$$

La formule de Bayes montre en particulier qu'un processus X est une Q-(sur)martingale (locale) si et seulement si ZX est une (sur)martingale (locale).

Théorème 1.2.11 *(Girsanov)*
Soit $Q \ll P$ et Z son processus de densité martingale. On suppose que Z est continu. Soit M une martingale locale continue. Alors le processus

$$M^Q = M - \int \frac{1}{Z} d < M, Z >$$

est une Q-martingale locale continue. De plus, si N est une martingale locale continue, on a

$$< M^Q, N^Q > = < M, N > .$$

Si on a de plus $Q \sim P$, i.e. Z est strictement positif p.s., alors il existe une unique martingale locale continue L nulle en $t = 0$, telle que :

$$Z_t = \exp\left(L_t - \frac{1}{2} < L, L >_t\right) =: \mathcal{E}_t(L), \quad t \in \mathbb{T}, \ p.s.$$

et L est donnée par

$$L_t = \int_0^t \frac{1}{Z_s} dZ_s, \quad t \in \mathbb{T}, \ p.s.$$

La Q-martingale locale M^Q s'écrit alors aussi :

$$M^Q = M - < M, L > .$$

Dans le cas d'un mouvement brownien, on a le résultat important suivant.

Théorème 1.2.12 *(Cameron-Martin)*
Soit W un mouvement brownien. Soit $Q \sim P$ de processus de densité martingale

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \mathcal{E}_t(L),$$

où L est une martingale locale continue. Alors le processus

$$W^Q = W - < W, L >$$

est un Q-mouvement brownien.

Dans les applications du théorème de Girsanov-Cameron-Martin, on part d'une martingale locale continue L nulle en 0 et on construit $Q \sim P$ en posant $Q = \mathcal{E}_t(L).P$ sur $\mathcal{F}_t$. Ceci requiert bien sûr que le processus $\mathcal{E}(L)$ soit une martingale. De façon générale, on sait que $\mathcal{E}(L)$ est une martingale locale, appelée exponentielle de Doléans-Dade. C'est aussi un processus (strictement) positif et donc une surmartingale ce qui assure l'existence de la limite p.s. $\mathcal{E}_{\bar{T}}(L)$ de $\mathcal{E}_t(L)$ quand t tend vers $\bar{T}$. L'exponentielle de Doléans-Dade $\mathcal{E}(L)$ est donc une martingale si et seulement si $E[\mathcal{E}_{\bar{T}}(L)] = 1$. Dans ce cas, on peut effectivement définir une probabilité $Q \sim P$ de densité de Radon-Nikodym $\mathcal{E}_{\bar{T}}(L) = dQ/dP$ et de processus de densité martingale $\mathcal{E}_t(L) = dQ/dP|_{\mathcal{F}_t}$. On peut alors ensuite appliquer le théorème de Girsanov-Cameron-Martin. Il est donc important d'avoir des conditions assurant que $\mathcal{E}(L)$ soit une martingale.

Proposition 1.2.12 *(Condition de Novikov)*
Soit L une martingale locale continue avec $L_0 = 0$ telle que

$$E\left[\exp\left(\frac{1}{2} < L, L >_{\bar{T}}\right)\right] < +\infty.$$

Alors L est une martingale uniformément intégrable avec $E[\exp(L_{\bar{T}}/2)] < +\infty$ et $\mathcal{E}(L)$ est une martingale uniformément intégrable.

Cas du mouvement brownien
Considérons un mouvement brownien d-dimensionnel $W = (W^1, \ldots, W^d)$ par rapport à $(\mathcal{F}_t)_{t\in\mathbb{T}}$. Un choix courant de martingale locale L est

$$L_t = \int_0^t \lambda'_s dW_s \ = \ \sum_{i=1}^d \int_0^t \lambda^i_s dW^i_s,$$

où $\lambda = (\lambda^1, \ldots, \lambda^d)$ est dans $L^2_{loc}(W)$. Le processus $\mathcal{E}(L)$ associé est

$$\mathcal{E}_t(L) = \exp\left(\int_0^t \lambda'_u dW_u - \frac{1}{2}\int_0^t |\lambda_u|^2 du\right), \quad t \in \mathbb{T}. \tag{1.10}$$

Lorsque $\mathcal{E}(L)$ est une martingale permettant de définir une probabilité $Q \sim P$ de processus de densité martingale $\mathcal{E}(L)$, le théorème de Girsanov-Cameron-Martin stipule que $W^Q = (W^{Q,1}, \ldots, W^{Q,d})$ avec

$$W^{Q,i} = W^i - \int \lambda^i dt, \quad i = 1, \ldots, d \tag{1.11}$$

est un Q-mouvement brownien. Notons que dans le cas où $\mathbb{F}$ est la filtration brownienne et d'après le théorème de représentation d'Itô 1.2.8, toute probabilité $Q \sim P$ a un processus de densité martingale de la forme (1.10). Finalement, la condition de Novikov assurant que $\mathcal{E}(L)$ est une martingale s'écrit ici :

$$E\left[\exp\left(\frac{1}{2}\int_0^{\bar{T}} |\lambda_t|^2 dt\right)\right] < +\infty.$$

On termine ce paragraphe en donnant un théorème de représentation des martingales browniennes sous un changement de probabilité.

Théorème 1.2.13 *(Représentation d'Itô sous un changement de probabilité) On suppose que $\mathbb{F}$ est la filtration naturelle (augmentée) d'un mouvement brownien standard d-dimensionnel $W = (W^1, \ldots, W^d)$. Soit $Q \sim P$ de processus de densité martingale Z donné par* (1.10) *et W^Q le mouvement brownien sous Q donné par* (1.11). *Soit $M = (M_t)_{t \in \mathbb{T}}$ une Q-martingale locale càd-làg. Alors il existe $\alpha = (\alpha^1, \ldots, \alpha^d) \in L^2_{loc}(W^Q)$ tel que :*

$$M_t = M_0 + \int_0^t \alpha'_u dW^Q_u \;=\; M_0 + \sum_{i=1}^d \int_0^t \alpha^i_t dW^{Q,i}_t, \quad t \in \mathbb{T}, \;\; p.s.$$

Ce résultat est à la base du problème de la réplication en marché complet en finance. Il sera aussi utilisé dans la suite de ce livre. Notons qu'il ne peut être déduit comme conséquence directe du théorème de représentation d'Itô 1.2.8 appliqué à la Q-martingale M par rapport à la filtration $\mathbb{F}$. En effet, la filtration $\mathbb{F}$ est la filtration naturelle de W et non de W^Q. Pour montrer le théorème 1.2.13, il faut d'abord se ramener sous P par la formule de Bayes en écrivant que $N = ZM$ est une martingale locale. On peut alors appliquer le théorème de représentation d'Itô 1.2.8 (sous P) à N et revenir ensuite à $M = N/Z$ par la formule d'Itô (voir les détails dans la preuve de la Proposition 5.8.6 de Karatzas et Shreve [KaSh88]).

1.3 Equations différentielles stochastiques

On rappelle dans cette section quelques résultats sur les équations différentielles stochastiques (EDS) à coefficients aléatoires par rapport à un mouvement brownien.

1.3.1 Solutions fortes d'EDS

On introduira ici seulement le concept de solutions fortes d'EDS. On part donc d'un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}, P)$ satisfaisant les conditions habituelles et d'un mouvement brownien d-dimensionnel $W = (W^1, \ldots, W^d)$ par rapport à $\mathbb{F}$. On se donne des fonctions $b(t, x, \omega) = (b_i(t, x, \omega))_{1 \leq i \leq d}$, $\sigma(t, x, \omega) = (\sigma_{ij}(t, x, \omega))_{1 \leq i \leq n, 1 \leq j \leq d}$ définies sur $\mathbb{T} \times \mathbb{R}^n \times \Omega$ à valeurs respectivement dans $\mathbb{R}^n$ et $\mathbb{R}^{n \times d}$. On suppose que pour tout ω, les fonctions $b(.,.,\omega)$ et $\sigma(.,.,\omega)$ sont boréliennes sur $\mathbb{T} \times \mathbb{R}^n$ et pour tout $x \in \mathbb{R}^n$, les processus $b(.,x,.)$ et $\sigma(.,x,.)$, notés parfois pour simplifier $b(.,x)$ et $\sigma(.,x)$ sont progressifs. On considère alors l'EDS à valeurs dans $\mathbb{R}^n$:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{1.12}$$

qui s'écrit aussi composante par composante

$$dX^i_t = b_i(t, X_t)dt + \sum_{j=1}^d \sigma_{ij}(t, X_t)dW^j_t, \quad 1 \leq i \leq d. \tag{1.13}$$

Dans les applications de ce cours, on considèrera deux types de situations :

(1) b et σ sont des fonctions déterministes Boréliennes $b(t,x)$, $\sigma(t,x)$ de t et x et on parle alors de diffusion pour l'EDS (1.12).

(2) Les coefficients aléatoires $b(t,x,\omega)$ et $\sigma(t,x,\omega)$ sont de la forme $\tilde{b}(t,x,\alpha_t(\omega))$, $\tilde{\sigma}(t,x,\alpha_t(\omega))$ où $\tilde{b}$, $\tilde{\sigma}$ sont des fonctions déterministes Boréliennes sur $\mathbb{T}\times\mathbb{R}^n\times A$, A ensemble de $\mathbb{R}^m$, et $\alpha = (\alpha_t)_{t\in\mathbb{T}}$ est un processus progressif à valeurs dans A. Ce cas intervient dans les problèmes de contrôle stochastique étudiés aux chapitres 3 et 4 et on dit parfois que l'EDS (1.12) est une diffusion contrôlée par α.

Définition 1.3.12 *(Solution forte d'EDS)*
Une solution forte de l'EDS (1.12) *partant à l'instant* t *est un processus vectoriel* $X = (X^1,\ldots,X^n)$ *progressif tel que l'on ait*

$$\int_t^s |b(u,X_u)|du + \int_t^s |\sigma(u,X_u)|^2 du < +\infty, \quad p.s., \ \forall t \le s \in \mathbb{T}$$

et que les relations

$$X_s = X_t + \int_t^s b(u,X_u)du + \int_t^s \sigma(u,X_u)dW_u, \quad t \le s \in \mathbb{T},$$

i.e.

$$X_s^i = X_t^i + \int_t^s b_i(u,X_u)du + \sum_{j=1}^d \int_t^s \sigma_{ij}(u,X_u)dW_u^j, \quad t \le s \in \mathbb{T},\ 1 \le i \le d,$$

soient vraies p.s.

Notons qu'une solution forte d'EDS (1.12) est un processus continu (à un processus indistinguable près) : ses trajectoires sont donc p.s. continues.

Nous mentionnons qu'il existe un autre concept de solution, dite faible, à l'EDS (1.12), où l'espace de probabilité filtré $(\Omega,\mathcal{F},\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}},P)$ et le mouvement brownien W font partie des inconnues de l'EDS, en plus de X.

L'existence (et l'unicité) d'une solution forte à l'EDS (1.12) est assurée par les conditions de Lipschitz et croissance linéaire suivantes : il existe une constante (déterministe) finie K et un processus réel κ tels que pour tous $t \in \mathbb{T}$, $\omega \in \Omega$, $x,y \in \mathbb{R}^n$

$$|b(t,x,\omega) - b(t,y,\omega)| + |\sigma(t,x,\omega) - \sigma(t,y,\omega)| \le K|x-y|, \tag{1.14}$$

$$|b(t,x,\omega)| + |\sigma(t,x,\omega)| \le \kappa_t(\omega) + K|x|, \tag{1.15}$$

avec

$$E\left[\int_0^t |\kappa_u|^2 du\right] < +\infty, \quad \forall t \in \mathbb{T}. \tag{1.16}$$

Sous (1.14), un choix naturel pour κ est $\kappa_t = |b(t,0)| + |\sigma(t,0)|$ dès lors que celui-ci vérifie la condition (1.16).

Théorème 1.3.14 *Sous les conditions* (1.14)-(1.15)-(1.16)*, il existe pour tout $t \in \mathbb{T}$, une solution forte à l'EDS* (1.12) *partant à l'instant t. De plus, pour tout ξ $\mathcal{F}_t$-mesurable à valeurs dans $\mathbb{R}^n$, tel que $E[|\xi|^2] < +\infty$, il y a unicité d'une solution forte X partant à l'instant t de ξ, i.e. $X_t = \xi$. L'unicité est trajectorielle (ou indistinguable) et signifie que si X et Y sont deux telles solutions fortes, on a $P[X_s = Y_s, \forall t \leq s \in \mathbb{T}] = 1$. Cette solution est de carré intégrable : pour tout $T > t$, il existe une constante C_T tel que :*

$$E\left[\sup_{t\leq s\leq T} |X_s|^2\right] \leq C_T(1 + E[|\xi|^2]).$$

Ce résultat est standard et on peut en trouver une preuve dans les livres de Gihman et Skorohod [GS72], Ikeda et Watanabe [IW81], Krylov [Kry80] ou Protter [Pro90]. On note usuellement par $X^{t,\xi} = \{X_s^{t,\xi}, t \leq s \in \mathbb{T}\}$ la solution forte de l'EDS (1.12) partant à l'instant t de ξ. Lorsque $t = 0$, on note simplement $X^\xi = X^{0,\xi}$. Par unicité trajectorielle, on note que pour tout $t \leq \theta$ dans $\mathbb{T}$, et $x \in \mathbb{R}^n$, on a p.s. :

$$X_s^{t,x} = X_s^{\theta, X_\theta^{t,x}}, \quad \forall\, \theta \leq s \in \mathbb{T}.$$

Lorsque b et σ sont des fonctions déterministes de t et x, on sait de plus que la solution forte de l'EDS (1.12) est adaptée par rapport à la filtration naturelle de W. On a aussi dans ce cas la propriété de Markov de toute solution forte X à l'EDS (1.12) : pour toute fonction borélienne bornée g sur $\mathbb{R}^d$, pour tous $t \leq \theta$ dans $\mathbb{T}$, on a

$$E\left[g(X_\theta)|\,\mathcal{F}_t\right] = \varphi_\theta(t, X_t)$$

où φ_θ est la fonction définie sur $\mathbb{T} \times \mathbb{R}^d$ par :

$$\varphi_\theta(t,x) = E\left[g(X_\theta^{t,x})\right].$$

Etant donnée une solution forte X de l'EDS (1.12) partant à l'instant t, et une fonction f de classe $C^{1,2}$ sur $\mathbb{T} \times \mathbb{R}^n$, la formule d'Itô pour $f(s, X_s)$, $t \leq s$ dans $\mathbb{T}$, s'écrit :

$$\begin{aligned} f(s, X_s) = f(t, X_t) &+ \int_t^s \frac{\partial f}{\partial t}(u, X_u) + \mathcal{L}_u f(u, X_u) du \\ &+ \int_t^s D_x f(u, X_u)' \sigma(u, X_u) dW_u \end{aligned}$$

où

$$\begin{aligned} \mathcal{L}_t(\omega) f(t,x) &= b(t,x,\omega).D_x f(t,x) + \frac{1}{2}\mathrm{tr}(\sigma(t,x,\omega)\sigma'(t,x,\omega)D_x^2 f(t,x)) \\ &= \sum_{i=1}^n b_i(t,x,\omega))\frac{\partial f}{\partial x_i}(t,x) + \sum_{i,k=1}^n \gamma_{ik}(t,x,\omega)\frac{\partial^2 f}{\partial x_i \partial x_k}(t,x), \end{aligned}$$

et $\gamma(t,x) = \sigma(t,x)\sigma'(t,x)$ est à valeurs matricielles $n \times n$ de composantes

$$\gamma_{ik}(t,x) = \sum_{j=1}^{d} \sigma_{ij}(t,x)\sigma_{kj}(t,x).$$

1.3.2 Estimations sur les moments de solutions d'EDS

Dans ce paragraphe, nous donnons des estimations sur les moments de solutions d'EDS (1.12). Notons que sous la condition (1.14), il existe une constante finie positive β_0 telle que pour tous $t \in \mathbb{T}$, $\omega \in \Omega$, $x, y \in \mathbb{R}^n$:

$$(b(t,x,\omega) - b(t,y,\omega)).(x-y) + \frac{1}{2}|\sigma(t,x,\omega) - \sigma(t,y,\omega)|^2 \leq \beta_0|x-y|^2.$$

Les estimations qui suivent sont basées essentiellement sur les inégalités de Burkholder-Davis-Gundy, Doob, la formule d'Itô et le lemme de Gronwall que nous rappelons ici.

Lemme 1.3.1 *(Gronwall)*
Soit g une fonction positive continue sur $\mathbb{R}_+$ telle que

$$g(t) \leq h(t) + C\int_0^t g(s)ds, \quad 0 \leq t \leq T,$$

où C est une constante positive et h est une fonction intégrable sur $[0,T]$, $T > 0$. Alors

$$g(t) \leq h(t) + C\int_0^t h(s)e^{C(t-s)}ds, \quad 0 \leq t \leq T.$$

Théorème 1.3.15 *Supposons les conditions* (1.14)-(1.15)-(1.16) *satisfaites.*
1) Il existe une constante C (dépendante de K) telle que pour tous $t \leq \theta$ dans $\mathbb{T}$ et $x \in \mathbb{R}^n$:

$$E\left[\sup_{t\leq s\leq\theta}\left|X_s^{t,x}\right|^2\right] \leq C|x|^2 + Ce^{C(\theta-t)}E\left[\int_t^\theta |x|^2 + |\kappa_u|^2 du\right] \quad (1.17)$$

$$E\left[\sup_{t\leq s\leq\theta}\left|X_s^{t,x} - x\right|^2\right] \leq Ce^{C(\theta-t)}E\left[\int_t^\theta |x|^2 + |\kappa_u|^2 du\right]. \quad (1.18)$$

2) Pour tous $0 \leq t \leq s$ dans $\mathbb{T}$ et $x, y \in \mathbb{R}^n$

$$E\left|X_s^{t,x} - X_s^{t,y}\right|^2 \leq e^{2\beta_0(s-t)}|x-y|^2. \quad (1.19)$$

On trouvera une preuve de ces estimations dans Krylov [Kry80], ch. 2. Notons que l'inégalité (1.18) implique en particulier que

$$\lim_{h\downarrow 0} E\left[\sup_{s\in[t,t+h]}\left|X_s^{t,x} - x\right|^2\right] = 0.$$

1.3.3 Formule de Feynman-Kac

Dans ce paragraphe, on considère l'EDS (1.12) avec des coefficients déterministes $b(t,x)$ et $\sigma(t,x)$. Pour tout $t \in \mathbb{T}$, on introduit l'opérateur (déterministe) différentiel du second-ordre :

$$(\mathcal{L}_t\varphi)(x) = b(t,x).D_x\varphi(x) + \frac{1}{2}\mathrm{tr}(\sigma(t,x)\sigma'(t,x)D_x^2\varphi(x)), \quad \varphi \in C^2(\mathbb{R}^n).$$

$\mathcal{L}_t$ est appelé générateur infinitésimal de la diffusion (1.12). Si X est une solution de l'EDS (1.12), $v(t,x)$ une fonction (réelle) de classe $C^{1,2}$ sur $\mathbb{T}\times\mathbb{R}^n$ et $r(t,x)$ une fonction continue sur $\mathbb{T}\times\mathbb{R}^d$, on a d'après la formule d'Itô :

$$\begin{aligned} M_t := e^{-\int_0^t r(s,X_s)ds}v(t,X_t) - \int_0^t e^{-\int_0^s r(u,X_u)du}\left(\frac{\partial v}{\partial t} + \mathcal{L}_s v - rv\right)(s,X_s)ds \\ = v(0,X_0) + \int_0^t e^{-\int_0^s r(u,X_u)du} D_x v(s,X_s)'\sigma(s,X_s)dW_s. \end{aligned} \quad (1.20)$$

Le processus M est donc une martingale locale continue.

On se place sur un horizon fini $\mathbb{T} = [0,T]$ et on considère le problème d'équation aux dérivées partielles (EDP) linéaire parabolique de Cauchy :

$$rv - \frac{\partial v}{\partial t} - \mathcal{L}_t v = f, \qquad \text{sur } [0,T[\times\mathbb{R}^n \quad (1.21)$$

$$v(T,.) = g, \qquad \text{sur } \mathbb{R}^n. \quad (1.22)$$

où f (resp. g) est une fonction continue de $[0,T]\times\mathbb{R}^n$ (resp. $\mathbb{R}^n$) dans $\mathbb{R}$. On suppose aussi que la fonction r est positive. On donne ici une version simple du théorème de représentation de Feynman-Kac. On rappelle que $X^{t,x}$ désigne la solution de la diffusion (1.12) partant à l'instant t de x.

Théorème 1.3.16 *(Représentation de Feynman-Kac)*
Soit v une fonction $C^{1,2}([0,T[\times\mathbb{R}^d) \cap C^0([0,T]\times\mathbb{R}^d)$ à dérivée en x bornée et solution du problème de Cauchy (1.21)-(1.22). Alors v admet la représentation

$$v(t,x) = E\left[\int_t^T e^{-\int_t^s r(u,X_u^{t,x})du} f(s,X_s^{t,x})ds + e^{-\int_t^T r(u,X_u^{t,x})du} g(X_T^{t,x})\right], \quad (1.23)$$

pour $(t,x) \in [0,T]\times\mathbb{R}^d$.

La preuve de ce résultat découle du fait que lorsque v est à derivée en x bornée, l'intégrand de l'intégrale stochastique dans (1.20) est dans $L^2(W)$ d'après la croissance linéaire en x de σ et l'estimation (1.17). Ainsi M est une martingale (de carré intégrable) et la représentation (1.23) s'obtient simplement en écrivant que $E[M_T] = E[M_t]$. On peut également obtenir cette représentation de Feynman-Kac sous d'autres conditions sur v, par exemple v à croissance quadratique. On montrera ce type de résultat au chapitre 3 dans le cas plus général de diffusion contrôlée (voir théorème 3.5.2 et remarque 3.5.5). On verra aussi une version du théorème de Feynman-Kac en horizon infini pour des problèmes d'EDP elliptiques (voir théorème 3.5.3 et remarque 3.5.6).

L'application du théorème précédent requiert l'existence d'une solution régulière v au problème de Cauchy (1.21)-(1.22). Ce genre de résultats s'obtient typiquement sous une hypothèse d'uniforme ellipticité de l'opérateur $\mathcal{L}_t$:

$$\exists \varepsilon > 0, \; \forall \; (t,x) \in [0,T] \times \mathbb{R}^n, \; \forall y \in \mathbb{R}^n, \; y'\sigma\sigma'(t,x)y \; \geq \; \varepsilon |y|^2 \tag{1.24}$$

et des hypothèses de bornitude sur b, σ et de croissance polynomiale sur f et g (voir Friedman [Fr75] p. 147). Il existe aussi d'autres types de conditions suffisantes qui relachent l'hypothèse d'uniforme ellipticité (1.24) mais imposent des conditions de régularité plus fortes sur les coefficients (voir Krylov [Kry80] p. 118).

Dans le cas général où il n'y a pas forcément de solution régulière au problème de Cauchy (1.21)-(1.22), on peut donner un sens à cette EDP avec un concept de solution faible appelée solution de viscosité. Cette notion sera développée au chapitre 4 dans le cadre plus général de diffusion contrôlée conduisant alors à une EDP non linéaire appelée équation d'Hamilton-Jacobi-Bellman.

2

Problèmes d'optimisation stochastique. Exemples en finance

2.1 Introduction

Dans ce chapitre, nous exposons la structure de base d'un problème de contrôle que nous illustrons à travers plusieurs exemples issus des mathématiques financières. L'analyse et la résolution de ces problèmes seront détaillées plus tard.

De façon générale, un problème de contrôle se formule selon les caractéristiques suivantes :

• **Etat du système** : On considère un système dynamique caractérisé par son état à tout instant. Le temps peut être discret ou continu. Nous considérons ici qu'il varie de façon continue et dans des conditions d'incertitude. L'horizon (intervalle de variation du temps) peut être fini ou infini.

L'état du système représente l'ensemble des variables quantitatives constituant une description 'exhaustive' du système. Les variables d'état sont supposées en nombre fini à valeurs réelles. On notera $X_t(\omega)$ l'état du système à l'instant t dans un scénario du monde $\omega \in \Omega$ espace mesurable muni d'une probabilité P.

Une fois défini l'état, il s'agit de décrire les lois d'évolution de cet état en fonction du temps. L'application $t \mapsto X_t$ décrit l'évolution du système. Cette évolution est fournie par un modèle probabiliste.

• **Contrôle** : La dynamique X_t de l'état du système est influencée par un contrôle que nous modélisons comme un processus $(\alpha_t)_t$ dont la valeur peut être décidée à tout instant t en fonction des informations disponibles à cet instant, c'est à dire que α est adapté par rapport à une certaine filtration, et prend ses valeurs dans un espace de contrôle A.

• **Critère de coût/performance** : L'objectif est de minimiser (ou maximiser) sur les contrôles une fonctionnelle $J(X, \alpha)$. Dans ce cours, on considérera des fonctionnelles de la forme

$$E\left[\int_0^T f(X_t,\omega,\alpha_t)dt + g(X_T,\omega)\right], \qquad \text{en horizon fini } T < +\infty$$

et

$$E\left[\int_0^{+\infty} e^{-\beta t} f(X_t,\omega,\alpha_t)dt\right], \qquad \text{en horizon infini.}$$

La fonction f est la fonction de coût intégral, g est le coût final et $\beta > 0$ est un coefficient d'actualisation. On définit alors la fonction valeur

$$v = \inf_\alpha J.$$

Les objectifs seront de déterminer d'une part la fonction valeur, et d'autre part les infima pour ces critères et les contrôles optimaux, s'ils existent, qui les réalisent.

2.2 Exemples

Nous présentons dans cette section plusieurs exemples pratiques de contrôle optimal intervenant en finance. Deux classes d'exemples seront considérées : la première concerne les problèmes d'allocation de portefeuille et d'investissement de firme, et la seconde les problèmes de couverture et valorisation d'options.

2.2.1 Allocation de portefeuille

On considère un marché financier avec un actif sans risque de processus de prix S^0 strictement positif représentant le compte d'épargne et n actifs risqués de processus de prix S représentant des actions.

Soit un investisseur ou agent qui investit à toute date t dans ces actifs, avec une quantité α_t dans les n actifs risqués. En notant par X_t sa richesse, la quantité investie dans l'actif sans risque à la date t est $(X_t - \alpha_t.S_t)/S_t^0$. Son processus de richesse autofinançante évolue selon :

$$dX_t = (X_t - \alpha_t.S_t)\frac{dS_t^0}{S_t^0} + \alpha_t dS_t.$$

Le contrôle est le processus α à valeurs dans A sous-ensemble de $\mathbb{R}^n$. Le problème d'allocation de portefeuille est de choisir quel est le meilleur investissement dans ces actifs dans un contexte incertain. Deux modélisations sont couramment utilisées pour représenter le comportement des individus et investisseurs : le critère d'espérance d'utilité et le critère moyenne-variance.

Dans le premier critère reposant sur une théorie du choix en univers incertain, l'individu compare des revenus aléatoires dont il connait les lois de probabilité. Sous certains conditions sur ses préférences, Von Neumann et Morgenstern montrent qu'elles peuvent se représenter par l'espérance d'une fonction, dite d'utilité. En notant U la fonction d'utilité de l'individu, cela signifie qu'un revenu aléatoire X sera préféré à un revenu aléatoire X' si $E[U(X)] \geq E[U(X')]$. Cette fonction d'utilité est croissante ce qui exprime l'amour de la richesse de l'individu. Elle est aussi supposée usuellement concave pour formaliser l'aversion pour le risque de l'individu. En effet, si l'individu n'aime pas le risque, à un revenu aléatoire X, il préfère obtenir avec certitude l'espérance $E[X]$ de ce revenu. Autrement dit, sa fonction d'utilité U vérifie

$$U(E[X]) \geq E[U(X)].$$

En particulier, si le revenu X vaut x avec probabilité λ et x' avec probabilité $1-\lambda$, $\lambda \in [0,1]$, on a

$$U(\lambda x + (1-\lambda)x') \geq \lambda U(x) + (1-\lambda)U(x'),$$

ce qui traduit la propriété de concavité de U. Dans notre contexte, ce critère économique consiste à maximiser l'espérance de l'utilité de la richesse terminale à un horizon fini $T < +\infty$:

$$\sup_\alpha E\left[U(X_T)\right], \tag{2.1}$$

où $U(x)$ est une fonction croissante et concave de $\mathbb{R}$ dans $[-\infty, +\infty[$.

Le problème initialement étudié par Merton [Mer69] est un cas particulier du modèle ci-dessus avec un modèle de Black-Scholes pour l'actif risqué :

$$\begin{aligned} dS_t^0 &= rS_t^0 dt, \\ dS_t &= \mu S_t dt + \sigma S_t dW_t. \end{aligned}$$

où μ, r, σ sont des constantes positives, W est un mouvement brownien réel sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}, P)$, et une fonction d'utilité de la forme :

$$U(x) = \begin{cases} \frac{x^p-1}{p}, & x \geq 0 \\ -\infty & x < 0, \end{cases}$$

où $0 < p < 1$, le cas limite $p = 0$ correspondant à une fonction d'utilité Logarithmique : $U(x) = \ln x$, $x > 0$. Ces fonctions d'utilité sont appelées isoélastiques ou CRRA (Constant Relative Risk Aversion) puisque l'indice relatif d'aversion pour le risque, défini par $\eta = -xU"(x)/U'(x)$, est constant dans ce cas.

Le critère moyenne-variance, initié par Markowitz [Ma52], repose sur l'hypothèse que les préférences de l'individu ne dépendent que de la moyenne et de la variance de ses revenus aléatoires. Pour exprimer le fait que l'individu aime la richesse et est averse au risque, le critère moyenne-variance s'intéressera aux portefeuilles MV-efficaces, c'est à dire minimisant la variance à espérance donnée. Dans notre contexte, le problème d'optimisation s'écrit :

$$\inf_{\alpha} \left\{ \mathrm{Var}(X_T) \ : E(X_T) = m \right\}.$$

Nous verrons qu'on peut se ramener à la résolution d'un problème de la forme (2.1) pour une fonction d'utilité de la forme :

$$U(x) = \lambda - x^2, \quad \lambda \in \mathbb{R}.$$

2.2.2 Modèle de production et consommation

On considère le modèle suivant d'une unité de production. La valeur K_t de son capital à la date t varie selon le taux d'investissement I_t en capital et le prix S_t par unité de capital :

$$dK_t = K_t \frac{dS_t}{S_t} + I_t dt.$$

La dette L_t de cette unité de production évolue selon le taux d'intérêt r, la consommation C_t et le taux de productivité P_t du capital :

$$dL_t = rL_t dt - \frac{K_t}{S_t} dP_t + (I_t + C_t)dt.$$

On choisit un modèle d'évolution de (S_t, P_t) selon :

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_1 S_t dW_t^1 \\ dP_t &= b dt + \sigma_2 dW_t^2, \end{aligned}$$

où (W^1, W^2) est un mouvement brownien de dimension 2 sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_t, P)$ et μ, b, σ_1, σ_2 sont des constantes, $\sigma_1, \sigma_2 > 0$. La valeur nette de l'unité de production est

$$X_t = K_t - L_t.$$

On impose les contraintes

$$K_t \geq 0, \ C_t \geq 0, \ X_t > 0, \ t \geq 0.$$

On note par $k_t = K_t/X_t$ et $c_t = C_t/X_t$ les variables de contrôle d'investissement et de consommation. La dynamique du système contrôlé est donc gouvernée par :

$$\begin{aligned} dX_t &= X_t \left[k_t \left(\mu - r + \frac{b}{S_t} \right) + (r - c_t) \right] dt \\ &\qquad + k_t X_t \sigma_1 dW_t^1 + k_t \frac{X_t}{S_t} \sigma_2 dW_t^2 \\ dS_t &= \mu S_t dt + \sigma_1 S_t dW_t^1 \end{aligned}$$

Etant donné un facteur d'actualisation $\beta > 0$ et une fonction d'utilité U, l'objectif est de déterminer l'investissement et la consommation optimale de l'unité de production :

$$\sup_{(k,c)} E\left[\int_0^{+\infty} e^{-\beta t} U(c_t X_t) dt\right].$$

2.2.3 Modèles d'investissement irréversible d'une firme

On considère une firme produisant un bien (électicité, pétrole, ...). Elle peut augmenter sa capacité de production X_t en transférant du capital à partir d'un autre secteur d'activité. L'évolution contrôlée de sa capacité de production évolue alors selon :

$$dX_t = X_t(-\delta dt + \sigma dW_t) + \alpha_t dt.$$

$\delta \geq 0$ est le taux de dépréciation de la production, $\sigma > 0$ sa volatilité, $\alpha_t dt$ est le nombre d'unités de capital acquis par la firme à un coût $\lambda \alpha_t dt$, $\lambda > 0$ s'interprétant comme un facteur de conversion d'un secteur d'activité à l'autre. Le contrôle α est à valeurs dans $\mathbb{R}_+$. C'est un modèle d'irréversibilité de l'expansion du capital de la firme. La fonction de profit de la firme est une fonction Π de $\mathbb{R}_+$ dans $\mathbb{R}$, concave, croissante. Le problème d'optimisation de la firme est donc :

$$\sup_{\alpha} E\left[\int_0^{\infty} e^{-\beta t} \left(\Pi(X_t) dt \ - \ \lambda \alpha_t dt\right)\right].$$

2.2.4 Couverture quadratique d'options

On considère des actifs risqués de processus de prix S et un actif sans risque de processus de prix S^0 strictement positif. Un produit dérivé d'actif sous-jacent S et de maturité T est un contrat financier dont la valeur à la date d'échéance T (on dit aussi le flux) est déterminée explicitement par le prix du sous-jacent S en (ou jusqu'en) T. Les options sont parmi les produits dérivés les plus traités sur les marchés financiers : ce sont des instruments financiers donnant le droit d'effectuer une transaction sur un actif sous-jacent à une date et un prix fixés à l'avance. Les options standard sur le sous-jacent S sont les options européennes d'achat (call) et de vente (put) qui peuvent être exercées uniquement à l'échéance T, à un prix d'exercice K, et dont le flux à cette date sont respectivement $g(S_T) = (S_T - K)_+$ et $(K - S_T)_+$. Plus généralement, on définira un actif contingent comme un contrat financier caractérisé par son flux H à la maturité T, où H est une variable aléatoire $\mathcal{F}_T$-mesurable. Le principe de couverture et valorisation par arbitrage d'une option ou actif contingent, consiste à construire une stratégie de portefeuille autofinançante sur les actifs de base S^0 et S du marché dont la richesse à maturité T soit égale au flux de l'option (on parle de réplication parfaite).

C'est toujours possible dans le cadre d'un marché complet, typiquement le modèle de Black-Scholes, mais en général, on est dans un contexte dit de marché incomplet où la réplication parfaite de l'option n'est pas réalisable, et divers critères sont utilisés pour couvrir et valoriser une option dans ce cadre. Nous considérons dans cet exemple le critère de couverture quadratique, et dans le suivant le critère de surréplication, qui conduisent chacun à des problèmes d'optimisation stochastique.

Un agent investit une quantité α_t dans les actifs risqués à la date t, en partant d'un capital initial x. L'évolution de sa richesse autofinancée est alors gouvernée par :

$$dX_t = \alpha_t dS_t + (X_t - \alpha_t S_t)\frac{dS_t^0}{S_t^0}.$$

On se donne un actif contingent de maturité T, représenté par une variable aléatoire H $\mathcal{F}_T$-mesurable. Le critère de couverture quadratique consiste à déterminer la stratégie de portefeuille α^* qui minimise l'écart pour la norme quadratique entre l'actif contingent et la valeur de la richesse à maturité :

$$\inf_\alpha E\left[H - X_T\right]^2.$$

2.2.5 Coût de surréplication en volatilité incertaine

On considère le prix d'un actif financier dont la volatilité α_t est inconnue et dont on sait à priori qu'elle est à valeurs dans un intervalle $A = [\underline{\sigma}, \bar{\sigma}]$. La dynamique du prix de l'actif risqué sous une probabilité martingale est :

$$dX_t = \alpha_t X_t dW_t.$$

Etant donné une option de flux $g(X_T)$ à la maturité T, on veut calculer son coût de surréplication donné par :

$$\sup_\alpha E[g(X_T)].$$

On peut montrer que ce coût est le plus petit capital initial permettant de construire une stratégie de portefeuille autofinançante dont la richesse à la maturité T domine le flux de l'option $g(S_T)$ pour toutes les réalisations possibles de la volatilité, voir par exemple Denis et Martini [DeMa03].

2.3 Autres modèles de contrôles en finance

2.3.1 Arrêt optimal

Dans les modèles présentés ci-dessus, l'horizon du problème est fixé, soit fini, soit infini. Il existe de nombreuses applications où le "contrôleur" a aussi la possibilité de décider l'horizon de son objectif. La décision de stopper le processus est modélisée par un temps d'arrêt et le problème d'optimisation associé est appelé problème d'arrêt optimal. Dans la formulation générale de tels problèmes, le contrôle est mixte, constitué du couple contrôle/temps d'arrêt (α, τ) et la fonctionnelle à optimiser s'écrit :

$$E\left[\int_0^\tau f(X_t, \alpha_t)dt + g(X_\tau)\right].$$

Ces problèmes interviennent en finance typiquement dans la valorisation des options américaines où, en plus par rapport aux options européennes, le détenteur de l'option peut exercer son droit et donc recevoir le flux asssocié à tout instant avant l'échéance. Dans cet exemple, il n'y a pas de contrôle α, X est le prix d'un actif risqué, $f = 0$ et $g(X_t)$ est le flux de l'option à la date t. Un autre exemple de problème d'arrêt optimal en finance concerne les modèles de production d'une firme qui peut décider de stopper à tout instant sa production. Dans ce cas, X s'interprète comme le prix du bien de production, f est la fonction de profit instantané et g le profit reçu au moment de la liquidation de la production.

2.3.2 Contrôle ergodique

Certains systèmes stochastiques peuvent exhiber sur le long terme un comportement stationnaire caractérisé par une mesure invariante. Cette mesure, si elle existe, est obtenue en calculant la moyenne de l'état du système sur le long terme. Un problème de contrôle ergodique consiste alors à optimiser sur le long terme un certain critère prenant en compte cette mesure invariante.

Une formulation standard résultant des critères étudiés précédemment consiste à optimiser sur les contrôles α une fonctionnelle de la forme

$$\limsup_{T\to+\infty} \frac{1}{T} E\left[\int_0^T f(X_t, \alpha_t)dt\right],$$

ou encore

$$\limsup_{T\to+\infty} \frac{1}{T} \ln E\left[\exp\left(\int_0^T f(X_t, \alpha_t)dt\right)\right].$$

Cette dernière formulation est appelée contrôle risk-sensible dans la littérature et a récemment été utilisée dans plusieurs travaux en mathématiques financières comme un critère de gestion de portefeuille à long terme.

Un autre critère basé sur le comportement de type grandes déviations du système : $P[X_T/T \geq c] \simeq e^{-I(c)T}$, quand T tend vers l'infini, consiste à maximiser sur les contrôles α une fonctionnelle de la forme

$$\limsup_{T\to+\infty} \frac{1}{T} \ln P\left[\frac{X_T}{T} \geq c\right].$$

Ce problème de contrôle de grandes déviations s'interprète en finance comme la version asymptotique (ergodique) du critère quantile consistant à maximiser la probabilité que la valeur terminale X_T du portefeuille soit au-dessus d'un certain index.

2.3.3 Surréplication sous contraintes gamma

On considère le modèle de Black Scholes pour un marché financier avec un actif risqué de prix S et un actif sans risque supposé constant pour simplifier :

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

Le processus de richesse X d'un agent qui dispose d'un capital initial x et investit une quantité α_t à la date t dans S_t est donné par

$$dX_t = \alpha_t dS_t, \qquad X_0 = x,$$

Le processus α est une semimartingale de la forme

$$d\alpha_t = dB_t + \gamma_t dS_t$$

où B est un processus à variation bornée et γ un processus adapté contraint à prendre ses valeurs dans un intervalle $[\underline{\gamma}, \bar{\gamma}]$.

Etant donné une option de flux $g(S_T)$, le problème de la surréplication est :

$$v = \inf \{ x \ : \exists\ (\alpha, B, \gamma), \ \ \text{tel que} \ \ X_T \geq g(S_T) \ \ p.s. \} .$$

La caractéristique nouvelle est que la contrainte porte non pas directement sur α mais sur la "dérivée" de α par rapport aux variations dS du prix.

2.3.4 Optimisation d'utilité robuste et mesures du risque

Dans les exemples décrits aux paragraphes 2.2.1 et 2.2.4, l'utilité du flux du portefeuille et/ou de l'option à optimiser, est mesurée selon le critère de préférences de Von Neumann-Morgenstern. Le risque d'un flux incertain $-X$ est évalué en terme d'une mesure du risque $\rho(-X)$ s'écrivant comme une espérance $-EU(-X)$ sous une certaine probabilité et pour une fonction d'utilité U. D'autres mesures du risque liées à l'incertitude des flux ont été proposées récemment telles que :

$$\rho(-X) = \sup_{Q \in \mathcal{Q}} E^Q[X] \tag{2.2}$$

ou encore

$$\rho(-X) = - \inf_{Q \in \mathcal{Q}} E^Q[U(-X)], \tag{2.3}$$

où $\mathcal{Q}$ est un ensemble de probabilités. La représentation (2.2) est motivée par la théorie des mesures cohérentes du risque initiée par Artzner et al [ADEH99]. On peut retrouver en particulier les mesures bien connues sur les marchés financiers de type "Value at Risk" pour certains choix de $\mathcal{Q}$. Les fonctionnelles d'utilité robustes de la forme (2.3) ont été suggérées par Gilboa et Schmeidler [GS89].

2.4 Commentaires bibliographiques

Les premiers travaux sur le contrôle optimal stochastique ont débuté dans les années 60 avec une dynamique d'état linéaire et un coût quadratique. Ce problème dit du régulateur linéaire stochastique intervient dans les applications en ingénierie, voir par exemple Davis [Da77] pour une introduction à ce sujet.

Les applications du contrôle optimal stochastique à des problèmes de gestion et de finance ont été développées à partir des années 70 avec notamment l'article pionnier de Merton [Mer69] sur l'allocation de portefeuille. Le modèle et les résultats de Merton ont été ensuite étendus par de nombreux auteurs, parmi lesquels Zariphopoulou [Zar88], Davis et Norman [DN90], Oksendal et Sulem [OS02]. Ces problèmes sont aussi étudiés dans la monographie de Karatzas et Shreve [KaSh98]. Le modèle avec production et consommation dans le paragraphe 2.2.2 est formulé et étudié dans Fleming et Pang [FP04]. Les modèles d'investissement pour des firmes sont largement développés dans le livre de Dixit et Pindick [DP94]. L'exemple du problème d'investissement irréversible présenté au paragraphe 2.2.3 est une version simplifiée d'un modèle étudié dans Guo et Pham [GP05]. La gestion de portefeuille par critère moyenne-variance a été formulée initialement par Markowitz [Ma52] dans un cadre statique à une période. Le critère de couverture d'options par minimisation quadratique présenté au paragraphe 2.2.4 a été introduit par Föllmer et Sondermann [FoS86]. Le problème du coût de surréplication en volatilité incertaine a été étudié initialement par Avellaneda, Levy et Paras [ALP95].

Les problèmes d'arrêt optimaux ont été très largement étudiés dans la littérature, initialement dans les années 60 puis avec un regain d'intérêt grâce au problème des options américaines et des options réelles. On peut citer les travaux de Dynkin [Dyn63], Van Morbecke [Van76], Shiryaev [Sh78], El Karoui [Elk81], ou plus récemment le livre d'Oksendal [Oks00]. Il existe de nombreux papiers sur la valorisation d'options américaines et on citera l'article de Lamberton [L98] pour une revue sur le sujet. Les modèles d'arrêt optimaux intervenant dans la production de firme sont étudiés par exemple dans Duckworth et Zervos [DZ00].

Les problèmes de contrôle ergodique ont été étudiés par Lasry [La74], Karatzas [Kar80], Bensoussan et Nagai [BN91], Fleming et McEneaney [FM95]. Les applications du contrôle risk-sensible en finance sont développés dans les articles de Bielecki et Pliska [BP99], Fleming et Sheu [FS00].

Le problème de contrôle de grandes déviations a été introduit et développé récemment par Pham [Pha03a], [Pha03b]. Une extension dans le cas avec observation partielle est étudiée par Hata [Ha04].

Le problème de la surréplication sous contraintes gamma a été développé par Soner-Touzi [ST00] et Cheridito, Soner et Touzi [CST03].

Les mesures du risque cohérentes introduites par Artzner et al [ADEH99] ont ensuite été développées par de nombreux autres auteurs. En particulier, nous mentionnons les mesures convexes du risque introduites par Föllmer et Schied [FoS02] et Frittelli et Rosazza Gianin [FG04]. Les problèmes d'optimisation en finance avec utilisation de ces mesures du risque sont étudiés récemment dans les travaux de Schied [Schi03], Gundel [Gu04], Barrieu et El Karoui [BElk04].

La liste des modèles de contrôle exposée dans ce livre n'est bien entendu pas exhaustive. Il existe de nombreux autres problèmes intéressants de contrôle en finance : citons les problèmes avec observation partielle (voir le livre de Bensoussan [Ben92]) ou encore le contrôle impulsionnel (voir Jeanblanc-Picqué et Shiryaev [JS95] et Oksendal et Sulem [OS04]). Pour d'autres exemples de modèles de contrôle en finance, on pourra se reporter aux livres de Kamien et Schwartz [KS81], Seierstad et Sydsaeter [SS87], Sethi et Zhang [SZ94] ou Korn [Kor97].

3

Approche EDP classique de la programmation dynamique

3.1 Introduction

Dans ce chapitre, nous étudions la méthode de la programmation dynamique pour résoudre un problème de contrôle stochastique. Le cadre adopté dans la section 3.2 est celui des processus de diffusion contrôlés à valeurs dans $\mathbb{R}^n$ et le problème formulé est en horizon fini ou en horizon infini.

L'idée basique de la méthode est de considérer une famille de problèmes de contrôle à différents états initiaux et d'établir des relations entre les fonctions valeurs associées. Ce principe, appelé principe de la programmation dynamique et initié dans les année 50 par Bellman, est énoncé précisément dans la section 3.3. L'équation de la programmation dynamique (en section 3.4) conduit à une équations aux dérivées partielles (EDP) nonlinéaire du second ordre, appelée Hamilton-Jacobi-Bellman (HJB). Lorsque cette EDP peut être résolue par l'obtention explicite ou théorique d'une solution régulière, le théorème de vérification, démontré en section 3.5, valide l'optimalité de ce candidat solution de HJB, et permet aussi de caractériser un contrôle optimal. Cette approche classique de la programmation dynamique est appelée étape de vérification. Nous illustrons cette méthode en section 3.6 sur deux exemples de modèles en finance. L'inconvénient majeur de cette approche est de supposer l'existence d'une solution régulière à l'EDP d'HJB. Ce n'est pas toujours le cas et nous illustrons ce fait sur un exemple simple en section 3.7 inspiré de la finance.

3.2 Contrôle de processus de diffusion

On considère un modèle de contrôle où l'état du système est gouverné par l'équation différentielle stochastique (EDS) à valeurs dans $\mathbb{R}^n$

$$dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s, \tag{3.1}$$

où W est un mouvement brownien d-dimensionnel sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, P)$ satisfaisant les conditions habituelles. Plus généralement, on peut aussi considérer des coefficients $b(t,x,a)$ et $\sigma(t,x,a)$ dépendants du temps t. Mais dans le cas des problèmes à horizon infini décrits ci-dessous, il est important de supposer que ces coefficients ne dépendent pas du temps afin d'avoir la stationnarité du problème et une fonction valeur indépendante du temps.

Le contrôle $\alpha = (\alpha_s)$ est un processus progressif (par rapport à $\mathbb{F}$) et à valeurs dans A, sous espace de $\mathbb{R}^m$.

Les fonctions mesurables $b : \mathbb{R}^n \times A \to \mathbb{R}^n$ et $\sigma : \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$ satisfont une condition de Lipschitz uniforme en A : $\exists\ K \geq 0$, $\forall\ x, y \in \mathbb{R}^n$, $\forall\ a \in A$,

$$|b(x,a) - b(y,a)| + |\sigma(x,a) - \sigma(y,a)| \leq K|x-y|. \tag{3.2}$$

Dans la suite pour $0 \leq t \leq T \leq +\infty$, on note $\mathcal{T}_{t,T}$ l'ensemble des temps d'arrêts à valeurs dans $[t,T]$. Lorsque $t = 0$ et $T = +\infty$, on note simplement $\mathcal{T} = \mathcal{T}_{0,+\infty}$.

Problème à horizon fini.
On fixe un horizon fini $0 < T < +\infty$. On note par $\mathcal{A}$ l'ensemble des processus de contrôle α tel que :

$$E\left[\int_0^T |b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2 dt\right] < +\infty. \tag{3.3}$$

Le point $x = 0$ est une valeur arbitraire de la diffusion et si ce point n'est pas dans le support de la diffusion, on peut choisir n'importe quelle autre valeur dans ce support. D'après la section 1.3 du chapitre 1, les conditions (3.2) et (3.3) assurent pour tout $\alpha \in \mathcal{A}$ et pour toute condition initiale (t,x) $\in [0,T]\times\mathbb{R}^n$, l'existence et l'unicité d'une solution forte à l'EDS (à coefficients aléatoires) (3.1) partant de x en $s = t$. On note alors par $\{X_s^{t,x}, t \leq s \leq T\}$ cette solution qui est p.s. à trajectoires continues. On rappelle aussi, que sous ces conditions sur b, σ et α, on a

$$E\left[\sup_{t\leq s\leq T} |X_s^{t,x}|^2\right] < +\infty. \tag{3.4}$$

$$\lim_{h\downarrow 0^+} E\left[\sup_{s\in[t,t+h]} \left|X_s^{t,x} - x\right|^2\right] = 0. \tag{3.5}$$

Critère de minimisation.

Soient $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ et $g : \mathbb{R}^n \to \mathbb{R}$ des fonctions mesurables. On suppose que

(**Hg**) (i) g est borné inférieurement
ou (ii) g est à croissance quadratique : $|g(x)| \leq C(1+|x|^2)$, $\forall x \in \mathbb{R}^n$, pour une constante C indépendante de x.

Pour $(t,x) \in [0,T] \times \mathbb{R}^n$, on note par $\mathcal{A}(t,x)$ le sous-ensemble des contrôles α de $\mathcal{A}$ tel que :

$$E\left[\int_t^T \left|f(s, X_s^{t,x}, \alpha_s)\right| ds\right] < +\infty. \tag{3.6}$$

On peut alors définir sous **(Hg)** la fonction de coût :

$$J(t,x,\alpha) = E\left[\int_t^T f(s, X_s^{t,x}, \alpha_s) ds + g(X_T^{t,x})\right],$$

pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$ et $\alpha \in \mathcal{A}(t,x)$. L'objectif étant de minimiser cette fonction coût, on introduit la fonction valeur :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} J(t,x,\alpha). \tag{3.7}$$

- Pour un état initial $(t,x) \in [0,T[\times\mathbb{R}^n$, on dit que $\hat{\alpha} \in \mathcal{A}(t,x)$ est un contrôle optimal si $v(t,x) = J(t,x,\hat{\alpha})$.

- Un processus de contrôle α de la forme $\alpha_s = a(s, X_s^{t,x})$ pour une fonction mesurable a de $[0,T] \times \mathbb{R}^n$ dans A, est appelé contrôle markovien.

Dans la suite, on supposera implicitement que la fonction valeur v est mesurable en ses arguments. Ce point n'est pas trivial à priori et on se reportera à des théorèmes de section (voir Appendice au chapitre III dans Dellacherie et Meyer [DM75]) pour des conditions assurant cette mesurabilité.

Remarque 3.2.1 Lorsque f est à croissance quadratique en x, i.e. il existe une constante positive C et une fonction positive $\kappa : A \to \mathbb{R}_+$ telles que :

$$|f(t,x,a)| \leq C(1+|x|^2) + \kappa(a), \quad \forall (t,x,a) \in [0,T] \times \mathbb{R}^n \times A, \tag{3.8}$$

alors l'estimation (3.4) montre que pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$, pour tout contrôle constant $\alpha = a$ dans A :

$$E\left[\int_t^T \left|f(s, X_s^{t,x}, a)\right| ds\right] < +\infty.$$

Ainsi, les contrôles constants dans A sont dans $\mathcal{A}(t,x)$. De plus, si il existe une constante positive C telle que $\kappa(a) \leq C(1 + |b(0,a)|^2 + |\sigma(0,a)|^2)$, pour tout a dans A, alors les conditions (3.3) et (3.4) montrent que pour tout (t,x) $\in [0,T] \times \mathbb{R}^n$, pour tout contrôle $\alpha \in \mathcal{A}$:

$$E\left[\int_t^T \left|f(s, X_s^{t,x}, \alpha_s)\right| ds\right] < +\infty.$$

Autrement dit, dans ce cas, $\mathcal{A}(t,x) = \mathcal{A}$.

Problème à horizon infini.
On note par $\mathcal{A}_0$ l'ensemble des processus de contrôle α tel que :

$$E\left[\int_0^T |b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2 dt\right] < +\infty, \quad \forall T > 0. \tag{3.9}$$

Etant donnés une condition initiale $t = 0$ et $x \in \mathbb{R}^n$, et un contrôle $\alpha \in \mathcal{A}_0$, il existe alors une unique solution forte, noté $\{X_s^x, s \geq 0\}$, de (3.1) partant de x en $t = 0$. On rappelle aussi qu'on a l'estimation suivante (voir théorème 1.3.15) :

$$E\left[|X_s^x|^2\right] \leq C|x|^2 + Ce^{Cs}E\left[\int_0^s |x|^2 + |b(0,\alpha_u)|^2 + |\sigma(0,\alpha_u)|^2 du\right], \tag{3.10}$$

pour une constante C indépendante de s, x et α.

Critère de minimisation.

Soit $\beta > 0$ et $f : \mathbb{R}^n \times A \to \mathbb{R}$ une fonction mesurable. Pour $x \in \mathbb{R}^n$, on note par $\mathcal{A}(x)$ le sous-ensemble des contrôles α de $\mathcal{A}_0$ tel que :

$$E\left[\int_0^{+\infty} e^{-\beta s} |f(X_s^x, \alpha_s)| \, ds\right] < +\infty. \tag{3.11}$$

On définit alors la fonction coût :

$$J(x,\alpha) = E\left[\int_0^{+\infty} e^{-\beta s} f(X_s^x, \alpha_s) ds\right],$$

pour $x \in \mathbb{R}^n$ et $\alpha \in \mathcal{A}(x)$, et la fonction valeur :

$$v(x) = \inf_{\alpha \in \mathcal{A}(x)} J(x,\alpha). \tag{3.12}$$

- Pour un état initial $x \in \mathbb{R}^n$, on dit que $\hat{\alpha} \in \mathcal{A}(x)$ est un contrôle optimal si $v(t,x) = J(t,x,\hat{\alpha})$.

- Un processus de contrôle α de la forme $\alpha_s = a(s, X_s^{t,x})$ pour une fonction mesurable a de $\mathbb{R}_+ \times \mathbb{R}^n$ dans A, est appelé contrôle markovien.

Il est important de supposer ici que la fonction $f(x,a)$ ne dépende pas du temps pour avoir la stationnarité du problème, i.e. la fonction valeur ne dépend pas de la date initiale à laquelle on considère le problème d'optimisation.

Remarque 3.2.2 Lorsque f est à croissance quadratique en x, i.e. il existe une constante positive C et une fonction positive $\kappa : A \to \mathbb{R}_+$ telles que :

$$|f(x,a)| \leq C(1 + |x|^2) + \kappa(a), \quad \forall (x,a) \in \mathbb{R}^n \times A, \tag{3.13}$$

alors l'estimation (3.10) montre que pour $\beta > 0$ assez grand, pour tout $x \in \mathbb{R}^n$, $a \in A$:

$$E\left[\int_0^{+\infty} e^{-\beta s} |f(X_s^x, a)| \, ds\right] < +\infty.$$

Autrement dit, les contrôles constants dans A appartiennent à $\mathcal{A}(x)$.

3.3 Principe de la programmation dynamique

Le principe de la programmation dynamique (PPD) est un principe fondamental pour la théorie du contrôle stochastique. Dans le contexte de contrôle de processus de diffusion décrit au paragraphe précédent, et même plus généralement pour des contrôles de processus de Markov, il s'énonce ainsi :

Théorème 3.3.1 *(Principe de la programmation dynamique)*
1) Horizon fini : Soit $(t,x) \in [0,T] \times \mathbb{R}^n$. Alors on a

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right] \quad (3.14)$$

$$= \inf_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (3.15)$$

2) Horizon infini : Soit $x \in \mathbb{R}^n$. Alors on a

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(x)} \inf_{\theta \in \mathcal{T}} E\left[\int_0^\theta e^{-\beta s} f(X_s^x, \alpha_s)ds + e^{-\beta\theta} v(X_\theta^x)\right] \quad (3.16)$$

$$= \inf_{\alpha \in \mathcal{A}(x)} \sup_{\theta \in \mathcal{T}} E\left[\int_0^\theta e^{-\beta s} f(X_s^x, \alpha_s)ds + e^{-\beta\theta} v(X_\theta^x)\right], \quad (3.17)$$

avec la convention que $e^{-\beta\theta(\omega)} = 0$ lorsque $\theta(\omega) = +\infty$.

Remarque 3.3.3 Le principe de la programmation dynamique énoncé ci-dessus peut se formuler de manière équivalente, dans le cas à horizon fini :
(i) Pour tout $\alpha \in \mathcal{A}(t,x)$ et $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) \leq E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (3.18)$$

(ii) Pour tout $\delta > 0$, il existe $\alpha \in \mathcal{A}(t,x)$ tel que pour tout $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) + \delta \geq E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (3.19)$$

C'est une version plus forte que la version traditionnelle du principe de la programmation dynamique, qui s'écrit en horizon fini :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right], \quad (3.20)$$

pour tout temps d'arrêt $\theta \in \mathcal{T}_{t,T}$. On a une remarque analogue dans le cas à horizon infini.

L'idée intuitive de ce principe est qu'un contrôle optimal $\hat{\alpha}$ sur $[t,T]$ peut être recollé en deux contrôles optimaux, l'un sur $[t,\theta]$ et l'autre sur $[\theta,T]$, et ceci quel que soit le temps d'arrêt θ. La preuve rigoureuse de ce résultat dans ce contexte stochastique est très technique. Nous donnons ici une preuve formelle de ce principe.

Preuve formelle du PPD.
On considère le cas de problème à horizon fini.
1. Etant donné un contrôle $\alpha \in \mathcal{A}(t,x)$, on a par unicite du flot de l'EDS gouvernant X, la structure markovienne :

$$X_s^{t,x} = X_s^{\theta, X_\theta^{t,x}}, \quad s \geq \theta,$$

où θ est un temps d'arrêt à valeurs dans $[t,T]$. Par la loi des espérances conditionnelles itérées, on a alors :

$$J(t,x,\alpha) = E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + J(\theta, X_\theta^{t,x},\alpha)\right],$$

d'où puisque $J(.,.,\alpha) \geq v$ et comme θ est quelconque dans $\mathcal{T}_{t,T}$:

$$\begin{aligned} J(t,x,\alpha) &\geq \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta, X_\theta^{t,x})\right] \\ &\geq \inf_{\alpha\in\mathcal{A}(t,x)} \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \end{aligned}$$

En passant à l'infimum sur α dans le terme de gauche, on obtient l'inégalité :

$$v(t,x) \geq \inf_{\alpha\in\mathcal{A}(t,x)} \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (3.21)$$

2. On considère pour tout $\varepsilon > 0$ et $\theta \in \mathcal{T}_{t,T}$, un contrôle ε-optimal α^ε de $v(\theta, X_\theta^{t,x})$:

$$J(\theta, X_\theta^{t,x}) \leq v(\theta, X_\theta^{t,x}) + \varepsilon.$$

Etant donné $\alpha \in \mathcal{A}(t,x)$, on définit le processus :

$$\hat{\alpha}_s = \begin{cases} \alpha_s, & s \in [0,\theta] \\ \alpha_s^\varepsilon, & s \in [\theta,T]. \end{cases}$$

Le point délicat est de vérifier que $\hat{\alpha}$ est progressivement mesurable et alors bien dans $\mathcal{A}(t,x)$. Dans ce cas, on a :

$$\begin{aligned} v(t,x) \leq J(t,x,\hat{\alpha}) &= E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + J(\theta, X_\theta^{t,x},\alpha^\varepsilon)\right] \\ &\leq E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta, X_\theta^{t,x})\right] + \varepsilon. \end{aligned}$$

Ceci étant valable pour tout $\alpha \in \mathcal{A}(t,x)$, $\theta \in \mathcal{T}_{t,T}$ et $\varepsilon > 0$, on a l'inégalité :

$$v(t,x) \leq \inf_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (3.22)$$

En combinant les deux inégalités (3.21) et (3.22), on obtient le résultat voulu.

3.4 Equation d'Hamilton-Jacobi-Bellman

L'équation d'Hamilton-Jacobi-Bellman (HJB) est la version infinitésimale du principe de la programmation dynamique : elle décrit le comportement local de la fonction valeur $v(t,x)$ lorsqu'on fait tendre le temps d'arrêt θ dans (3.20) vers t.

Dans cette section, nous dérivons formellement l'équation d'HJB en supposant que la fonction valeur v est suffisamment régulière. L'obtention rigoureuse de cette équation sera développée au chapitre suivant grâce à la théorie des solutions de viscosité.

3.4.1 Dérivation formelle de HJB

Problème à horizon fini.
Considérons le temps $\theta = t+h$ et un contrôle constant $\alpha_s = a$, avec a arbitraire dans A, dans la relation (3.18) de la programmation dynamique :

$$v(t,x) \leq E\left[\int_t^{t+h} f(s, X_s^{t,x}, a)ds + v(t+h, X_{t+h}^{t,x})\right]. \quad (3.23)$$

En supposant que v est suffisamment régulière, on a par la formule d'Itô entre t et $t+h$:

$$\begin{aligned} v(t+h, X_{t+h}^{t,x}) &= v(t,x) \\ &\quad + \int_t^{t+h} \left(\frac{\partial v}{\partial t} + \mathcal{L}^a v\right)(s, X_s^{t,x})ds \quad + \quad \text{martingale (locale)} \end{aligned}$$

où $\mathcal{L}^a$ est l'opérateur associé à la diffusion (3.1) pour le contrôle constant a et défini par (voir paragraphe 1.3.1) :

$$\mathcal{L}^a v = b(x,a).D_x v + \frac{1}{2}\text{tr}\left(\sigma(x,a)\sigma'(x,a)D_x^2 v\right).$$

En substituant dans (3.23), on obtient alors :

$$0 \leq E\left[\int_t^{t+h} \left(\frac{\partial v}{\partial t} + \mathcal{L}^a v\right)(s, X_s^{t,x}) + f(s, X_s^{t,x}, a)ds\right].$$

En divisant par h et en faisant tendre h vers 0, on a :

$$0 \leq \frac{\partial v}{\partial t}(t,x) + \mathcal{L}^a v(t,x) + f(t,x,a).$$

Ceci étant valable pour tout $a \in A$, on a l'inégalité :

$$-\frac{\partial v}{\partial t}(t,x) + \sup_{a \in A}\left[-\mathcal{L}^a v(t,x) - f(t,x,a)\right] \leq 0. \tag{3.24}$$

D'autre part, supposons que α^* est un contrôle optimal. Alors dans (3.20), on a :

$$v(t,x) = E\left[\int_t^{t+h} f(s, X_s^*, \alpha_s^*)ds + v(t+h, X_{t+h}^*)\right],$$

où X^* est l'état du système solution de (3.1) partant de x en t avec le contrôle α^*. Par un argument similaire et avec des conditions de régularités sur v, on obtient :

$$-\frac{\partial v}{\partial t}(t,x) - \mathcal{L}^{\alpha_t^*} v(t,x) - f(t,x,\alpha_t^*) = 0, \tag{3.25}$$

ce qui combiné avec (3.24) suggère que v doit satisfaire :

$$-\frac{\partial v}{\partial t}(t,x) + \sup_{a \in A}\left[-\mathcal{L}^a v(t,x) - f(t,x,a)\right] = 0, \quad \forall (t,x) \in [0,T[\times \mathbb{R}^n, \tag{3.26}$$

si le supremum ci-dessus en a est fini. Nous verrons plus tard comment traiter le cas où ce supremum est infini, ce qui peut intervenir lorsque l'espace des contrôles A est non borné. On reécrit souvent cette équation aux dérivées partielles (EDP) sous la forme :

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) = 0, \quad \forall (t,x) \in [0,T[\times \mathbb{R}^n, \tag{3.27}$$

où pour $(t,x,p,M) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$ ($\mathcal{S}_n$ est l'ensemble des matrices $n \times n$ symétriques) :

$$H(t,x,p,M) = \sup_{a \in A}\left[-b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x,a)M\right) - f(t,x,a)\right].$$

Cette fonction H est appelée Hamiltonien du problème de contrôle considéré. Cette équation (3.27) est appelée équation de la programmation dynamique ou équation de Hamilton-Jacobi-Bellman (HJB). A cette équation aux dérivées partielles, il faut ajouter la condition terminale :

$$v(T,x) = g(x), \quad \forall x \in \mathbb{R}^n, \tag{3.28}$$

qui résulte immédiatement de la définition (3.7) de la fonction valeur v.

Remarque 3.4.4 1) Lorsque l'ensemble des contrôles est réduit à un singleton $\{a_0\}$, c'est à dire qu'il n'y a pas de contrôle sur l'état du système, l'EDP d'HJB se réduit au problème d'EDP linéaire de Cauchy :

$$-\frac{\partial v}{\partial t}(t,x) - \mathcal{L}^{a_0} v(t,x) = f(t,x,a_0), \quad \forall (t,x) \in [0,T[\times \mathbb{R}^n \tag{3.29}$$

$$v(T,x) = g(x), \quad \forall x \in \mathbb{R}^n. \tag{3.30}$$

2) L'argument d'optimalité de la programmation dynamique suggère que si l'on peut trouver un contrôle $\alpha^*(t,x)$ tel que :

$$\sup_{a\in A} \left[-\mathcal{L}^a v(t,x) - f(x,a)\right] = -\mathcal{L}^{\alpha^*(t,x)} v(t,x) - f(x,\alpha^*(t,x)),$$

c'est à dire que

$$\alpha^*(t,x) \in \arg\min_{a\in A} \left[\mathcal{L}^a v(t,x) + f(x,a)\right],$$

alors on aura :

$$-\frac{\partial v}{\partial t} - \mathcal{L}^{\alpha^*(t,x)} v(t,x) - f(x,\alpha^*(t,x)) = 0,$$

et donc

$$v(t,x) = E\left[\int_t^T f(X_s^*, \alpha^*(s,X_s^*))ds + g(X_T^*)\right],$$

où X^* est solution de l'équation différentielle stochastique :

$$\begin{aligned} dX_s^* &= b(X_s^*, \alpha^*(s,X_s^*)) + \sigma(X_s^*, \alpha^*(s,X_s^*))dW_s, \quad t \le s \le T \\ X_t^* &= x, \end{aligned}$$

et α^* est un contrôle optimal markovien.

Problème à horizon infini.
En reprenant les mêmes arguments que dans le cas d'un problème à horizon infini, on dérive l'équation d'Hamilton-Jacobi-Bellman pour la fonction valeur définie en (3.12) :

$$\beta v(x) + \sup_{a\in A} \left[-\mathcal{L}^a v(x) - f(x,a)\right] = 0, \quad \forall x \in \mathbb{R}^n,$$

qu'on reécrit aussi sous la forme :

$$\beta v + H(x, Dv(x), D^2 v(x)) = 0, \quad \forall x \in \mathbb{R}^n,$$

où pour $(x,p,M) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$:

$$H(x,p,M) = \sup_{a\in A} \left[-b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x,a)M\right) - f(x,a)\right].$$

3.4.2 Remarques-Extensions

1. Les résultats dans le cas horizon fini s'étendent aisément lorsque la fonction de coût J à minimiser a la forme plus générale suivante :

$$J(t,x,\alpha) = E\left[\int_t^T \Gamma(t,s) f(s, X_s^{t,x}, \alpha_s) ds + \Gamma(T) g(X_T^{t,x})\right],$$

où

$$\Gamma(t,s) = \exp\left(-\int_t^s \beta(u, X_u^{t,x}, \alpha_u) du\right), \quad t \leq s \leq T,$$

et $\beta(.,.,a)$ est une fonction continue positive sur $[0,T] \times \mathbb{R}^n$ pour tout $a \in A$. Dans ce cas l'Hamiltonien associé au problème de contrôle stochastique est :

$$\begin{aligned} & H(t,x,v,p,M) \\ & \quad = \sup_{a \in A}\left[\beta(t,x,a)v - b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma(x,a)\sigma'(x,a)M\right) - f(t,x,a)\right]. \end{aligned}$$

2. Lorsque l'espace des contrôles A est non borné, l'Hamiltonien

$$H(t,x,p,M) = \sup_{a \in A}\left[-b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x,a)M\right) - f(t,x,a)\right].$$

peut prendre la valeur $+\infty$ dans un certain domaine de (t,x,p,M). Plus précisément, supposons qu'il existe une fonction continue $G(t,x,p,M)$ sur $[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$ telle que :

$$H(t,x,p,M) \; < \; +\infty \Longleftrightarrow G(t,x,p,M) \; \leq \; 0.$$

Alors d'après le raisonnement conduisant à l'équation d'HJB (3.27), on doit avoir

$$G(t,x,D_x v(t,x), D_x^2 v(t,x)) \leq 0, \tag{3.31}$$

$$\text{et} \quad -\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) \leq 0 \tag{3.32}$$

De plus si l'inégalité (3.31) est stricte en un point $(t,x) \in [0,T[\times\mathbb{R}^n$, alors il existe un voisinage de $(t,x,D_x v(t,x), D_x^2 v(t,x))$ en lequel H est fini. Alors, formellement, le contrôle devrait être atteint au voisinage de (t,x) et par un raisonnement analogue à (3.27), on doit avoir l'égalité dans (3.32). On obtient ainsi une inéquation variationnelle d'HJB pour la programmation dynamique :

$$\begin{aligned} & \max\left[-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) \; , \; G(t,x,D_x v(t,x), D_x^2 v(t,x))\right] \\ & \quad = 0. \end{aligned}$$

On dit que le problème de contrôle est singulier. Un cas typique est lorsque le contrôle intervient de manière linéaire dans la dynamique du système et dans la fonction de coût. Par exemple, dans le case unidimensionnel $n = 1$, $A = \mathbb{R}_-$, et

$$\begin{aligned} b(x,a) &= a + \hat{b}(x), \ \ \sigma(x,a) = \hat{\sigma}(x), \\ f(t,x,a) &= a + \hat{f}(t,x), \end{aligned}$$

alors

$$H(t,x,p,M) = \begin{cases} -\hat{b}(x)p - \frac{1}{2}\hat{\sigma}(x)^2 M - \hat{f}(t,x) & \text{si } p+1 \leq 0 \\ +\infty & \text{si } p+1 > 0. \end{cases}$$

L'inéquation variationnelle d'HJB s'écrit ainsi :

$$\max\left[-\frac{\partial v}{\partial t}(t,x) - \hat{b}(x)\frac{\partial v}{\partial x}(t,x) - \frac{1}{2}\hat{\sigma}(x)^2\frac{\partial^2 v}{\partial x^2}(t,x)\ ,\ \frac{\partial v}{\partial x}(t,x) + 1\right] = 0.$$

On en verra un autre exemple dans le paragraphe 3.7. On étudiera aussi au chapitre suivant comment les équations (inéquations variationnelles) d'HJB s'obtiennent de façon rigoureuse, sans hypothèse de régularité sur la fonction valeur, grâce à la théorie des solutions de viscosité.

3. Lorsqu'on étudie un problème de maximisation

$$v(t,x) = \sup_{\alpha\in\mathcal{A}(t,x)} E\left[\int_t^T f(s, X_s^{t,x}, \alpha_s)ds + g(X_T^{t,x})\right],$$

on peut se ramener à un problème de minimisation en considérant la fonction valeur $-v$. Ceci revient alors à considérer l'Hamiltonien

$$H(t,x,p,M) = \inf_{a\in A}\left[-b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma(x,a)\sigma'(x,a)M\right) - f(t,x,a)\right],$$

avec une équation d'HJB :

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) = 0.$$

Lorsque H peut prendre la valeur $-\infty$ en supposant qu'il existe une fonction continue $G(t,x,p,M)$ sur $[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_n$ telle que :

$$H(t,x,p,M) \ > \ -\infty \Longleftrightarrow G(t,x,p,M) \ \geq \ 0,$$

l'inéquation variationnelle d'HJB s'écrit :

$$\min\left[-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x))\ ,\ G(t,x,D_x v(t,x), D_x^2 v(t,x))\right] = 0.$$

On a une remarque analogue dans le cas d'un problème de maximisation en horizon infini.

3.5 Théorème de vérification

L' étape cruciale dans l'approche classique de la programmation dynamique consiste à montrer, étant donnée une solution régulière à l'équation d'HJB, que ce candidat, sous des conditions suffisantes, coincide avec la fonction valeur. Ce résultat est appelé théorème de vérification et permet aussi d'obtenir un contrôle optimal. Il repose essentiellement sur la formule d'Itô. Les énoncés peuvent varier d'un problème à l'autre au niveau des conditions suffisantes requises. Celles-ci doivent être adaptées au cadre des hypothèses du problème particulier considéré. Nous proposons ici une version assez générale compte tenu du contexte défini au paragraphe 3.2.

Théorème 3.5.2 *(Horizon fini)*
Soit $w \in C^{1,2}([0,T[\times\mathbb{R}^n) \cap C^0([0,T]\times\mathbb{R}^n)$ à croissance quadratique, i.e. il existe une constante C telle que :

$$|w(t,x)| \leq C(1+|x|^2), \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n.$$

(i) Supposons que :

$$-\frac{\partial w}{\partial t}(t,x) + \sup_{a\in A}\left[-\mathcal{L}^a w(t,x) - f(t,x,a)\right] \leq 0, \quad (t,x) \in [0,T[\times\mathbb{R}^n, \tag{3.33}$$

$$w(T,x) \leq g(x), \quad x \in R^n. \tag{3.34}$$

Alors $w \leq v$ sur $[0,T]\times\mathbb{R}^n$.
(ii) De plus supposons que $w(T.) = g$, et pour tout $(t,x) \in [0,T[\times\mathbb{R}^n$, il existe $\hat{\alpha}(t,x)$ mesurable à valeurs dans A tel que :

$$\begin{aligned}-\frac{\partial w}{\partial t}(t,x) + \sup_{a\in A}\left[-\mathcal{L}^a w(t,x) - f(t,x,a)\right] = -\frac{\partial w}{\partial t}(t,x) - \mathcal{L}^{\hat{\alpha}(t,x)} w(t,x) \\ -f(t,x,\hat{\alpha}(t,x)) = 0,\end{aligned}$$

l'EDS :

$$dX_s = b(X_s, \hat{\alpha}(s,X_s))ds + \sigma(X_s,\hat{\alpha}(s,X_s))dW_s$$

admette une solution, notée $\hat{X}_s^{t,x}$, étant donnée une condition initiale $X_t = x$, et $\{\hat{\alpha}(s,\hat{X}_s^{t,x})\ t \leq s \leq T\} \in \mathcal{A}(t,x)$. Alors

$$w = v \quad \text{sur } [0,T]\times\mathbb{R}^n,$$

et $\hat{\alpha}$ est un contrôle optimal markovien.

Preuve. (i) Puisque $w \in C^{1,2}([0,T[\times\mathbb{R}^n)$, on a pour tout $(t,x) \in [0,T[\times\mathbb{R}^n$, $\alpha \in \mathcal{A}(t,x)$, $s \in [t,T[$, et pour tout temps d'arrêt τ à valeurs dans $[t,+\infty[$, par la formule d'Itô :

$$w(s\wedge\tau, X^{t,x}_{s\wedge\tau}) = w(t,x) + \int_t^{s\wedge\tau} |\frac{\partial w}{\partial t}(u, X^{t,x}_u) + \mathcal{L}^{\alpha_u} w(u, X^{t,x}_u) du + \int_t^{s\wedge\tau} D_x w(u, X^{t,x}_u)'\sigma(X^{t,x}_u, \alpha_u) dW_u.$$

On choisit $\tau = \tau_n = \inf\{s \geq t \ : \int_t^s |D_x w(u, X^{t,x}_u)'\sigma(X^{t,x}_u, \alpha_u)|^2 du \geq n\}$ en notant que $\tau_n \nearrow +\infty$ quand n tend vers l'infini. Le processus arrêté $\{\int_t^{s\wedge\tau_n} D_x w(u, X^{t,x}_u)'\ \sigma(X^{t,x}_u, \alpha_u) dW_u,\ t \leq s \leq T\}$ est donc une martingale et on a en prenant l'espérance :

$$E\left[w(s\wedge\tau_n, X^{t,x}_{s\wedge\tau_n})\right] = w(t,x) + E\left[\int_t^{s\wedge\tau_n} \frac{\partial w}{\partial t}(u, X^{t,x}_u) + \mathcal{L}^{\alpha_u} w(u, X^{t,x}_u) du\right].$$

Puisque w satisfait (3.33), on a :

$$\frac{\partial w}{\partial t}(u, X^{t,x}_u) + \mathcal{L}^{\alpha_u} w(u, X^{t,x}_u) + f(X^{t,x}_u, \alpha_u) \geq 0, \quad \forall \alpha \in \mathcal{A}(t,x),$$

d'où :

$$E\left[w(s\wedge\tau_n, X^{t,x}_{s\wedge\tau_n})\right] \geq w(t,x) - E\left[\int_t^{s\wedge\tau_n} f(X^{t,x}_u, \alpha_u) du\right],\ \forall \alpha \in \mathcal{A}(t,x). \tag{3.35}$$

On a

$$\left|\int_t^{s\wedge\tau_n} f(X^{t,x}_u, \alpha_u) du\right| \leq \int_t^T \left|f(X^{t,x}_u, \alpha_u)\right| du,$$

et le terme de droite est intégrable d'après la condition d'intégrablité sur $\mathcal{A}(t,x)$. Comme w est à croissance quadratique, on a :

$$\left|w(s\wedge\tau_n, X^{t,x}_{s\wedge\tau_n})\right| \leq C\left(1 + \sup_{s\in[t,T]} |X^{t,x}_s|^2\right),$$

et le terme de droite est intégrable d'après (3.4). On peut donc appliquer le théorème de convergence dominée et faire tendre n vers l'infini dans (3.35) :

$$E\left[w(s, X^{t,x}_s)\right] \geq w(t,x) - E\left[\int_t^s f(X^{t,x}_u, \alpha_u) du\right], \quad \forall \alpha \in \mathcal{A}(t,x).$$

Comme w est continue sur $[0,T] \times \mathbb{R}^n$, en faisant tendre s vers T, on a par le théorème de convergence dominée et en utilisant aussi (3.34) :

$$E\left[g(X^{t,x}_T)\right] \geq w(t,x) - E\left[\int_t^T f(X^{t,x}_u, \alpha_u) du\right], \quad \forall \alpha \in \mathcal{A}(t,x),$$

et donc $w(t,x) \leq v(t,x)$, $\forall (t,x) \in [0,T] \times \mathbb{R}^n$.

(ii) On applique la formule d'Itô à $w(u, \hat{X}_u^{t,x})$ entre $t \in [0,T[$ et $s \in [t,T[$ (après avoir localisé avec τ_n) :

$$E\left[w(s, \hat{X}_s^{t,x})\right] = w(t,x) + E\left[\int_t^s \frac{\partial w}{\partial t}(u, \hat{X}_u^{t,x}) + \mathcal{L}^{\hat{\alpha}(u,\hat{X}_u^{t,x})} w(u, \hat{X}_u^{t,x}) du\right].$$

Or par définition de $\hat{\alpha}(t,x)$, on a :

$$-\frac{\partial w}{\partial t} - \mathcal{L}^{\hat{\alpha}(t,x)} w(t,x) - f(t,x,\hat{\alpha}(t,x)) = 0,$$

d'où :

$$E\left[w(s, \hat{X}_s^{t,x})\right] = w(t,x) - E\left[\int_t^s f(\hat{X}_u^{t,x}, \hat{\alpha}(u, \hat{X}_u^{t,x})) du\right].$$

En faisant tendre s vers T, on obtient ainsi :

$$\begin{aligned} w(t,x) &= E\left[\int_t^T f(\hat{X}_u^{t,x}, \hat{\alpha}(u, \hat{X}_u^{t,x})) du + g(\hat{X}_T^{t,x})\right] \\ &= J(t,x,\hat{\alpha}). \end{aligned}$$

Donc $w(t,x) = J(t,x,\hat{\alpha}) \geq v(t,x)$ et finalement $w = v$ avec $\hat{\alpha}$ comme contrôle optimal markovien. □

Remarque 3.5.5 Dans le cas particulier où l'espace des contrôles A est réduit à un singleton $\{a_0\}$, ce théorème de vérification est une version du théorème de représentation de Feynman-Kac 1.3.16 : il stipule que si w est une fonction $C^{1,2}([0,T[\times\mathbb{R}^n) \cap C^0([0,T[\times\mathbb{R}^n)$ à croissance quadratique solution du problème de Cauchy (3.29)-(3.30), alors w admet la représentation

$$w(t,x) = E\left[\int_t^T f(X_s^{t,x}, a_0) ds + g(X_T^{t,x})\right].$$

Théorème 3.5.3 *(Horizon infini)*
Soit $w \in C^2(\mathbb{R}^n)$ à croissance quadratique.
(i) Supposons que :

$$\beta w(x) + \sup_{a \in A} \left[-\mathcal{L}^a w(x) - f(x,a)\right] \leq 0, \quad x \in \mathbb{R}^n, \tag{3.36}$$

$$\liminf_{T \to +\infty} e^{-\beta T} E[w(X_T^x)] \leq 0, \quad \forall x \in \mathbb{R}^n,\ \forall \alpha \in \mathcal{A}(x), \tag{3.37}$$

Alors $w \leq v$ sur $\mathbb{R}^n$.
(ii) Supposons de plus que pour tout $x \in \mathbb{R}^n$, il existe $\hat{\alpha}(x)$ mesurable à valeurs dans A tel que :

$$\beta w(x) + \sup_{a \in A} [-\mathcal{L}^a w(x) - f(x,a)] = \beta w(x) - \mathcal{L}^{\hat{\alpha}(x)} w(x) - f(x, \hat{\alpha}(x))$$
$$= 0,$$

l'EDS :

$$dX_s = b(X_s, \hat{\alpha}(X_s))ds + \sigma(X_s, \hat{\alpha}(X_s))dW_s$$

admette une solution, notée $\hat{X}_s^x$, étant donnée une condition initiale $X_0 = x$, avec

$$\limsup_{T \to +\infty} e^{-\beta T} E[w(\hat{X}_T^x] \geq 0, \tag{3.38}$$

et tel que $\{\hat{\alpha}(\hat{X}_s^x),\ s \geq 0\}, \in \mathcal{A}(x)$. Alors

$$w(x) = v(x), \quad \forall x \in \mathbb{R}^n$$

et $\hat{\alpha}$ est un contrôle optimal markovien.

Preuve. (i) Soit $w \in C^2(\mathbb{R}^n)$ et $\alpha \in \mathcal{A}(x)$. On a par la formule d'Itô à $e^{-\beta t} w(X_t^x)$ entre 0 et $T \wedge \tau_n$:

$$e^{-\beta(T \wedge \tau_n)} w(X_{T \wedge \tau_n}^x) = w(x) + \int_0^{T \wedge \tau_n} e^{-\beta u} \left[\mathcal{L}^{\alpha_u} w(X_u^x - \beta w(X_u^x)\right] du$$
$$+ \int_0^{T \wedge \tau_n} e^{-\beta u} Dw(X_u^x)' \sigma(X_u^x, \alpha_u) dW_u.$$

Ici, τ_n est le temps d'arrêt : $\tau_n = \inf\{t \geq 0\ :\ \int_0^t |D_x w(u, X_u^x)' \sigma(X_u^x, \alpha_u)|^2 du \geq n\}$. Comme le terme d'intégrale stochastique arrêtée est une martingale, on a en prenant l'espérance :

$$E\left[e^{-\beta(T \wedge \tau_n)} w(X_{T \wedge \tau_n}^x)\right] = w(x) + E\left[\int_0^{T \wedge \tau_n} e^{-\beta u} \left(-\beta w + \mathcal{L}^{\alpha_u} w\right)(X_u^x) du\right]$$
$$\geq w(x) - E\left[\int_0^{T \wedge \tau_n} e^{-\beta u} f(X_u^x, \alpha_u) du\right], \tag{3.39}$$

d'après (3.36). Par la condition de croissance quadratique de w et la condition d'intégrabilité (3.11), on peut appliquer le théorème de convergence dominée et faire tendre n vers l'infini :

$$E\left[e^{-\beta T} w(X_T^x)\right] \geq w(x) - E\left[\int_0^T e^{-\beta u} f(X_u^x, \alpha_u) du\right], \quad \forall \alpha \in \mathcal{A}(x).$$

En faisant tendre T vers l'infini, on a d'après (3.37) et le théorème de convergence dominée :

$$w(x) \leq E\left[\int_0^{+\infty} e^{-\beta t} f(X_t^x, \alpha_t) dt\right], \quad \forall \alpha \in \mathcal{A}(x),$$

et donc $w(x) \leq v(x)$, $\forall x \in \mathbb{R}^n$.

(ii) En appliquant la formule d'Itô à $e^{-\beta t} w(\hat{X}_s^x)$ (après avoir localisé avec τ_n) et en observant que le contrôle $\hat{\alpha}$ atteint l'égalité dans (3.39), on a :

$$E\left[e^{-\beta T} w(\hat{X}_T^x)\right] = w(x) - E\left[\int_0^T e^{-\beta u} f(\hat{X}_u^x, \hat{\alpha}(\hat{X}_u^x)) du\right].$$

En faisant tendre T vers l'infini et d'après (3.38), on obtient ainsi :

$$w(x) \geq J(x, \hat{\alpha}) = E\left[\int_0^{\infty} e^{-\beta u} f(\hat{X}_u^x, \hat{\alpha}(\hat{X}_u^x)) du\right]$$

et donc $w(x) = v(x) = J(x, \hat{\alpha})$. □

Remarque 3.5.6 Dans le cas particulier où l'espace des contrôles A est réduit à un singleton $\{a_0\}$, ce théorème de vérification est une version du théorème de représentation de Feynman-Kac 1.3.16 en horizon infini : il stipule que si w est une fonction $C^2(\mathbb{R}^n)$ à croissance quadratique solution du problème d'EDP elliptique

$$\begin{aligned} \beta w(x) - \mathcal{L}^{a_0} w(x) - f(x, a_0) &= 0, \quad x \in \mathbb{R}^n, \\ \lim_{T \to +\infty} e^{-\beta T} E[w(X_T^x)] &= 0, \quad x \in \mathbb{R}^n, \end{aligned}$$

alors w admet la représentation

$$w(t, x) = E\left[\int_0^{+\infty} e^{-\beta t} f(X_t^x, a_0) dt\right].$$

Le théorème précédent suggère la stratégie suivante pour résoudre le problème de contrôle stochastique. Dans le cas horizon fini, résoudre l'EDP non linéaire d'Hamilton-Jacobi-Bellman :

$$-\frac{\partial w}{\partial t} + \sup_{a \in A} \left[-\mathcal{L}^a w(t, x) - f(t, x, a)\right] = 0, \quad (t, x) \in [0, T[\times \mathbb{R}^n, \quad (3.40)$$

avec la condition terminale $w(T, x) = g(x)$. Fixer $(t, x) \in [0, T[\times \mathbb{R}^n$ et résoudre $\sup_{a \in A} \left[-\mathcal{L}^a w(t, x) - f(x, a)\right]$ comme un problème de maximum en $a \in A$. On note $a^*(t, x)$ la valeur de a qui réalise le maximum. Si cette EDP non linéaire avec condition terminale admet une solution régulière w, alors w est la fonction valeur du problème de contrôle stochastique et a^* est un contrôle optimal markovien. Cette méthode se justifie donc si l'EDP d'HJB (3.40) admet une solution $C^{1,2}$ satisfaisant les conditions d'application du théorème de vérification. Des résultats d'existence de solutions régulières à des équations de type HJB paraboliques (horizon fini) ou elliptiques (horizon infini) sont établis dans Fleming et Rishel [FR75], Gilbarg et Trudinger [GT85] ou encore Krylov [Kry87]. La condition principale imposée est une condition d'uniforme ellipticité :

$$\begin{aligned} &\text{Il existe une constante } c > 0 \text{ telle que} \\ &y'\sigma(x,a)\sigma'(x,a)y \;\geq\; c|y|^2, \qquad \forall x, y \in \mathbb{R}^n, \forall a \in A. \end{aligned}$$

Soulignons aussi dans la vérification des conditions (ii) des théorèmes 3.5.2 et 3.5.3 qu'il n'est pas toujours facile et parfois problématique d'obtenir l'existence d'une solution à l'EDS associé au candidat $\hat{\alpha}$ pour être le contrôle optimal.

3.6 Applications

3.6.1 Problème de choix de portefeuille de Merton en horizon fini

On reprend l'exemple décrit au paragraphe 2.2.1. Un agent investit à chaque date t une proportion α_t de sa richesse dans un actif risqué S et $1 - \alpha_t$ dans un actif sans risque S^0, avec la contrainte qu'à toute date, α_t doit être à valeurs dans A ensemble fermé convexe de $\mathbb{R}$. Son processus de richesse évolue selon l'EDS :

$$\begin{aligned} dX_t &= \frac{X_t\alpha_t}{S_t} dS_t + \frac{X_t(1-\alpha_t)}{S_t^0} dS_t^0 \\ &= X_t\left(\alpha_t\mu + (1-\alpha_t)r\right) dt + X_t\alpha_t\sigma dW_t. \end{aligned}$$

Partant d'une richesse initiale $X_t = x > 0$ au temps t, l'agent veut maximiser l'espérance de l'utilité de sa richesse terminale à un horizon $T > t$. Notons par $X^{t,x}$ le processus de richesse partant de x en t. Observons que les coefficients de X ne vérifient pas stricto-sensu la condition de Lipschitz uniforme en les contrôles a. En fait, on se ramène usuellement à ce cadre en considérant le logarithme de la richesse positive. La fonction valeur du problème de maximisation est donc :

$$v(t,x) = \sup_{\alpha \in \mathcal{A}} E\Big[U(X_T^{t,x})\Big], \quad (t,x) \in [0,T] \times \mathbb{R}_+,$$

où $\mathcal{A}$ est l'ensemble des processus α progressifs à valeurs dans A et tels que $E[\int_0^T |\alpha_s|^2 ds] < +\infty$.

La fonction d'utilité U est croissante et concave sur $\mathbb{R}_+$. Vérifions que $v(t,.)$ est croissante et concave en x. Soit $0 < x \leq y$ et α un processus de contrôle dans $\mathcal{A}$. Notons $Z_s = X_s^{t,x} - X_s^{t,y}$. Alors $dZ_s = Z_s\left[(\alpha_s\mu + (1-\alpha_s)r)\, ds + \alpha_s\sigma dW_s\right]$, $Z_t = y - x \geq 0$ et donc $Z_s \geq 0$ ou encore $X_s^{t,y} \geq X_s^{t,x}$ pour tout $s \geq t$. Puisque U est croissante, on a $U(X_T^{t,x}) \leq U(X_T^{t,y})$ d'où :

$$E\left[U(X_T^{t,x})\right] \leq E\left[U(X_T^{t,y})\right] \;\leq\; v(t,y), \quad \forall \alpha \;\in \mathcal{A}_t,$$

et donc $v(t,x) \leq v(t,y)$. Soit $0 < x_1, x_2$, α^1, α^2 deux processus de contrôle dans $\mathcal{A}$ et $\lambda \in [0,1]$. Posons $x_\lambda = \lambda x_1 + (1-\lambda)x_2$, X^{t,x_1} le processus de richesse partant de x_1 et controlé par α^1, X^{t,x_2} le processus de richesse partant de x_2 et controlé par α^2. Posons :

$$\alpha_s^\lambda = \frac{\lambda X_s^{t,x_1}\alpha_s^1 + (1-\lambda)X_s^{t,x_2}\alpha_s^2}{\lambda X_s^{t,x_1} + (1-\lambda)X_s^{t,x_2}}.$$

Notons que par convexité de A, le processus $\alpha^\lambda \in \mathcal{A}$. De plus, d'après la structure linéaire de l'évolution de l'équation de la richesse, le processus $X^\lambda := \lambda X^{t,x_1} + (1-\lambda)X^{t,x_2}$ est gouverné par :

$$\begin{aligned} dX_s^\lambda &= X_s^\lambda\left(\alpha_s^\lambda \mu + (1-\alpha_s^\lambda)r\right)ds + X_s^\lambda \alpha_s^\lambda \sigma dW_s, \quad s \geq t,\\ X_t^\lambda &= x_\lambda. \end{aligned}$$

Ceci prouve que $\lambda X^{t,x_1} + (1-\lambda)X^{t,x_2}$ est un processus de richesse partant de x_λ en t et contrôlé par α^λ. D'après la concavité de la fonction U, on a :

$$U\left(\lambda X_T^{t,x_1} + (1-\lambda)X_T^{t,x_2}\right) \geq \lambda U(X_T^{t,x_1}) + (1-\lambda)U(X_T^{t,x_2}),$$

d'où :

$$v(\lambda x_1 + (1-\lambda)x_2) \geq \lambda E\left[U(X_T^{t,x_1})\right] + (1-\lambda)E\left[U(X_T^{t,x_2})\right],$$

et ceci pour tout α^1, α^2 dans $\mathcal{A}$. On en déduit que :

$$v(\lambda x_1 + (1-\lambda)x_2) \geq \lambda v(x_1) + (1-\lambda)v(x_2).$$

En fait, on voit que si U est strictement concave et s'il existe un contrôle optimal, alors les arguments ci-dessus montrent que la fonction v est aussi strictement concave en x.

On va donc chercher à résoudre l'équation de Bellman :

$$-\frac{\partial w}{\partial t} + \inf_{a\in A}\left[-\mathcal{L}^a w(t,x)\right] = 0, \tag{3.41}$$

avec la condition terminale

$$w(T,x) = U(x), \quad x \in \mathbb{R}_+. \tag{3.42}$$

Ici $\mathcal{L}^a w(t,x) = x(a\mu + (1-a)r)\dfrac{\partial w}{\partial x} + \frac{1}{2}x^2a^2\sigma^2\dfrac{\partial^2 w}{\partial x^2}$. Le problème (3.41)–(3.42) n'a en général pas de solution explicite pour une fonction d'utilité U quelconque. Cependant, dans le cas particulier d'une fonction puissance :

$$U(x) = x^p, \quad x \geq 0,\ p < 1,$$

on peut résoudre explicitement ce problème. Cherchons une solution de la forme :

$$w(t, x) = \phi(t)x^p.$$

En substituant dans (3.41)-(3.42), on obtient que ϕ satisfait :

$$\begin{aligned} -\phi'(t) + \rho\phi(t) &= 0, \\ \phi(T) &= 1, \end{aligned}$$

où

$$\rho = \inf_{a \in A} \left[-ap(\mu - r) - pr + \frac{1}{2}a^2 p(1-p)\sigma^2 \right]. \tag{3.43}$$

On obtient alors $\phi(t) = \exp(-\rho(T-t))$ Ainsi, la fonction donnée par :

$$w(t, x) = \exp(-\rho(T-t))x^p, \quad (t, x) \in [0, T] \times \mathbb{R}_+, \tag{3.44}$$

est régulière, strictement croissante et strictement concave, et est solution de (3.41)-(3.42). De plus, la fonction $a \in A \mapsto -ap(\mu - r) - rp + \frac{1}{2}a^2p(1-p)\sigma^2$ est strictement convexe sur l'ensemble convexe A donc atteint son minimum en $\hat{a}$ constant. Par construction, $\hat{a}$ atteint l'infimum de $\inf_{a \in A}[-\mathcal{L}^a w(t, x)]$. De plus, l'équation de la richesse associée au contrôle constant $\hat{a}$:

$$dX_t = X_t \left(\hat{a}\mu + (1 - \hat{a})r \right) dt + X_t \hat{a} \sigma dW_t$$

admet bien une unique solution, étant donnée une condition initiale. Ceci prouve donc finalement d'après le théorème de vérification 3.5.2, que la fonction valeur du problème de maximisation est donnée par (3.44) et que la proportion optimale de richesse à investir dans l'actif risqué est constante et donnée par $\hat{a}$. Notons que lorsque $A = \mathbb{R}$, on a :

$$\hat{a} = \frac{\mu - r}{\sigma^2(1-p)}, \tag{3.45}$$

et

$$\rho = -\frac{(\mu - r)^2}{2\sigma^2} \frac{p}{1-p} - rp.$$

3.6.2 Un modèle de production et consommation en horizon infini

On reprend l'exemple du paragraphe 2.2.2. On considère le modèle suivant d'une unité de production. La valeur K_t de son capital à la date t varie selon le taux d'investissement I_t en capital et le prix S_t par unité de capital :

$$dK_t = K_t \frac{dS_t}{S_t} + I_t dt.$$

La dette L_t de cette unité de production évolue selon le taux d'intérêt r, la consommation C_t et le taux de productivité P_t du capital :

$$dL_t = rL_t dt - \frac{K_t}{S_t} dP_t + (I_t + C_t)dt.$$

On choisit un modèle d'évolution de $(Y_t = \ln S_t, P_t)$ selon :

$$\begin{aligned} dY_t &= \left(\mu - \frac{\sigma_1^2}{2}\right) dt + \sigma_1 dW_t^1 \\ dP_t &= bdt + \sigma_2 dW_t^2, \end{aligned}$$

où (W^1, W^2) est un mouvement brownien de dimension 2 sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, P)$ et μ, b, σ_1, σ_2 sont des constantes, $\sigma_1, \sigma_2 > 0$. La valeur nette de l'unité de production est

$$X_t = K_t - L_t.$$

On impose les contraintes

$$K_t \geq 0,\ C_t \geq 0,\ X_t > 0,\ \ t \geq 0.$$

On note par $k_t = K_t/X_t$ et $c_t = C_t/X_t$ les variables de contrôle d'investissement et de consommation. La dynamique du système contrôlé est donc gouvernée par :

$$\begin{aligned} dX_t &= X_t \left[k_t(\mu - r + be^{-Y_t}) + (r - c_t)\right] dt \\ &\quad + k_t X_t \sigma_1 dW_t^1 + k_t X_t e^{-Y_t} \sigma_2 dW_t^2 \qquad (3.46) \\ dY_t &= \left(\mu - \frac{\sigma_1^2}{2}\right) dt + \sigma_1 dW_t^1 \qquad (3.47) \end{aligned}$$

Etant donné un facteur d'actualisation $\beta > 0$ et une fonction d'utilité :

$$U(C) = \frac{C^\gamma}{\gamma}, \quad C \geq 0, \quad \text{où } 0 < \gamma < 1,$$

on note par $\mathcal{A}(x, y)$ l'ensemble des processus de contrôle (k, c) progressifs à valeurs dans $\mathbb{R}_+ \times \mathbb{R}_+$ tels que :

$$\begin{aligned} &\int_0^T k_t^2 dt + \int_0^T c_t^2 dt < +\infty, \quad p.s., \quad \forall T > 0, \\ E &\left[\int_0^{+\infty} e^{-\beta t} U(c_t X_t^{x,y}) dt\right] < \infty, \end{aligned}$$

où $(X^{x,y}, Y^y)$ est la solution de l'EDS (3.46)-(3.47) partant de (x, y) en $t = 0$. L'objectif est de déterminer l'investissement et la consommation optimale de l'unité de production. On cherche à résoudre le problème de contrôle stochastique en horizon infini :

$$v(x,y) = \sup_{(k,c)\in\mathcal{A}(x,y)} E\left[\int_0^{+\infty} e^{-\beta t} U(c_t X_t^{x,y}) dt\right]. \tag{3.48}$$

L'équation d'Hamilton-Jacobi-Bellman associée est :

$$\beta v - \left(\mu - \frac{\sigma_1^2}{2}\right)\frac{\partial v}{\partial y} - rx\frac{\partial v}{\partial x} - \frac{\sigma_1^2}{2}\frac{\partial^2 v}{\partial y^2} + \inf_{c\geq 0}\left[cx\frac{\partial v}{\partial x} - U(cx)\right] \tag{3.49}$$

$$+ \inf_{k\geq 0}\left[-k(\mu - r + be^{-y})x\frac{\partial v}{\partial x} - \frac{1}{2}k^2x^2(\sigma_1^2 + e^{-2y}\sigma_2^2)\frac{\partial^2 v}{\partial x^2} - kx\sigma_1^2\frac{\partial^2 v}{\partial x\partial y}\right] = 0.$$

En remarquant que $X^{x,y}$ s'exprime sous la forme $X^{x,y} = x\exp(Z(y))$ où $Z(y)$ s'écrit comme une exponentielle de processus en fonction de (k, c) et Y^y, on en déduit que :

$$v(x,y) = x^\gamma v(1,y).$$

On cherche donc une solution de HJB (3.49) sous la forme

$$v(x,y) = \frac{x^\gamma}{\gamma}\exp(\varphi(y)),$$

où $\varphi(y)$ est une fonction de $\mathbb{R}$ dand $\mathbb{R}$. En substituant cette forme dans (3.49), on obtient l'équation différentielle ordinaire (EDO) que doit satisfaire φ :

$$\beta - \gamma r - \frac{\sigma_1^2}{2}\left(\varphi_{yy} + \varphi_y^2\right) - \left(\mu - \frac{\sigma_1^2}{2}\right)\varphi_y$$
$$+ \inf_{c\geq 0}\left[c\gamma - c^\gamma e^{-\varphi}\right] + \gamma\inf_{k\geq 0} G(y, \varphi_y, k) = 0, \tag{3.50}$$

où

$$G(y,p,k) = \frac{k^2}{2}(1-\gamma)(\sigma_1^2 + \sigma_2^2 e^{-2y}) - k(\mu - r + be^{-y} + \sigma_1^2 p).$$

On peut montrer que pour β assez grand, il existe une unique solution régulière C^2 φ bornée à l'EDO (3.50). On consultera [FP04] pour les détails. Comme $G(y,p,0) = 0$, on a qu'en tout point y extrémum de φ_y, i.e. $\varphi_{yy}(y) = 0$:

$$0 \leq \beta - \gamma r - \left(\mu - \frac{\sigma_1^2}{2}\right)\varphi_y(y) - \frac{\sigma_1^2}{2}\varphi_y^2(y) - (1-\gamma)e^{\frac{\gamma\varphi(y)}{\gamma-1}}.$$

Comme φ est bornée, ceci prouve que φ_y est aussi bornée sur $\mathbb{R}$.

Par construction, la fonction positive $w(x,y) = (x^\gamma/\gamma)e^{\varphi(y)}$ est solution de HJB (3.49). D'après la première partie du théorème de vérification 3.5.3, on en déduit que $w \geq v$. D'autre part, considérons les fonctions :

$$\hat{k}(y) = \left(\frac{be^{-y} + \mu - r + \sigma_1^2\varphi_y}{(1-\gamma)(\sigma_1^2 + \sigma_2^2 e^{-2y})}\right)_+ \in \arg\min_{k\geq 0} G(y, \varphi_y, k)$$

$$\hat{c}(y) = \exp\left(\frac{\varphi(y)}{\gamma - 1}\right) \in \arg\min_{c\geq 0}\left[c\gamma - c^\gamma e^{-\varphi}\right].$$

Comme φ et φ_y sont bornées, ceci implique que les fonctions $\hat{c}$, $\hat{k}$, $e^{-y}\hat{k}$ sont aussi bornées en y. On en déduit qu'il existe une constante $M > 0$ telle que

$$E\left[\left|\hat{X}_t^{x,y}\right|^2\right] \leq x^2 \exp(Mt), \quad \forall t > 0,$$

où on a noté par $\hat{X}^{x,y}$ la solution de (3.46) contrôlée par $(\hat{k}(Y_t^y), \hat{c}(Y_t^y))_{t\geq 0}$. Ceci montre que le contrôle $(\hat{k}(Y_t^y), \hat{c}(Y_t^y))_{t\geq 0}$ est dans $\mathcal{A}(x,y)$:

$$E\left[\int_0^{+\infty} e^{-\beta t} U(\hat{c}(Y_t^y)\hat{X}_t^{x,y})dt\right] < +\infty, \tag{3.51}$$

De plus, comme φ est bornée, la fonction $\hat{c}(y)$ est bornée inférieurement par une constante strictement positive. On en déduit l'existence d'une constante $B > 0$ telle que :

$$0 \leq e^{-\beta T} w(\hat{X}_T^{x,y}, Y_T^y) \leq B e^{-\beta T} U(\hat{c}(Y_T^y)\hat{X}_T^{x,y}),$$

ce qui combiné avec (3.51), montre que :

$$\lim_{T\to+\infty} E\left[e^{-\beta T} w(\hat{X}_T^{x,y}, Y_T^y)\right] = 0.$$

On conclut avec la deuxième partie (ii) du théorème de vérification 3.5.3 que $w = v$ et que $(\hat{k}(Y_t^y), \hat{c}(Y_t^y))_{t\geq 0}$ est un contrôle optimal markovien.

3.7 Exemple de problème de contrôle stochastique singulier

On considère le processus contrôlé réel gouverné par :

$$dX_s = \alpha_s dW_s,$$

où α est un processus progressif à valeurs dans $A = \mathbb{R}$ et tel que $E[\int_0^T |\alpha_s|^2 ds]$ $< +\infty$. On note $\mathcal{A}$ l'ensemble de ces processus de contrôles. Soit g une fonction mesurable positive ou à croissance linéaire sur $\mathbb{R}$, et considérons le problème de contrôle stochastique :

$$v(t,x) = \inf_{\alpha\in\mathcal{A}} E[g(X_T^{t,x})], \quad (t,x) \in [0,T] \times \mathbb{R}. \tag{3.52}$$

Nous allons voir que pour un large choix de fonctions g, la fonction valeur v n'est pas régulière.

D'après le principe de la programmation dynamique, on a pour tout temps d'arrêt $\theta \in \mathcal{T}_{t,T}$, et tout contrôle constant $\alpha_t = a \in \mathbb{R}$:

$$v(t,x) \leq E\left[v(\theta, X_\theta^{t,x})\right]. \tag{3.53}$$

Supposons que v est régulière $C^{1,2}$ et appliquons la formule d'Itô à $v(s, X_s^{t,x})$ entre $s = t \in [0, T[$ et $s = \theta = (t+h) \wedge \tau$ où $\tau = \inf\{s \geq t \ : |X_s^{t,x} - x| \geq 1\}$. Alors l'intégrale stochastique apparaissant dans la formule d'Itô est une martingale arrêtée et en substituant dans (3.53), on obtient :

$$0 \leq E\left[\frac{1}{h}\int_t^{(t+h)\wedge\tau}\left(\frac{\partial v}{\partial t} + a^2\frac{\partial^2 v}{\partial x^2}\right)(s, X_s^{t,x})ds\right] \tag{3.54}$$

Notons que par continuité p.s. de la trajectoire de $X_s^{t,x}$, on a que pour $h \leq \bar{h}(\omega)$ suffisamment petit, $\theta = t + h$ p.s. On en déduit par le théorème de la moyenne que la variable aléatoire sous le signe espérance dans (3.54) converge p.s. vers $\left(\frac{\partial v}{\partial t} + a^2\frac{\partial^2 v}{\partial x^2}\right)(t, x)$ quand h tend vers zéro. De plus, cette variable aléatoire étant bornée par une constante indépendante de h, on peut appliquer le théorème de convergence dominée et obtenir quand h tend vers zéro :

$$0 \leq \frac{\partial v}{\partial t}(t, x) + a^2\frac{\partial^2 v}{\partial x^2}(t, x), \quad \forall (t, x) \in [0, T[\times\mathbb{R}. \tag{3.55}$$

Cette inégalité étant valable pour tout a dans $\mathbb{R}$, on doit avoir en particulier que $\frac{\partial^2 v}{\partial x^2} \geq 0$ sur $[0, T[\times\mathbb{R}$, autrement dit :

$$v(t, .) \ \text{ est convexe sur } \mathbb{R} \ \text{ pour tout } t \in [0, T[. \tag{3.56}$$

D'autre part, comme le contrôle constant nul est dans $\mathcal{A}$, il est immédiat que

$$v(t, x) \leq g(x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}.$$

En notant par g^{conv} l'enveloppe convexe de g, i.e. la plus grande fonction convexe minorant g, on en déduit d'après (3.56) :

$$v(t, x) \leq g^{conv}(x), \quad \forall (t, x) \in [0, T[\times\mathbb{R}. \tag{3.57}$$

En utilisant $g^{conv} \leq g$, l'inégalité de Jensen et la propriété de martingale de $X_s^{t,x}$ pour $\alpha \in \mathcal{A}$, on a :

$$\begin{aligned} v(t, x) &\geq \inf_{\alpha\in\mathcal{A}_t} E[g^{conv}(X_T^{t,x})] \\ &\geq \inf_{\alpha\in\mathcal{A}_t} g^{conv}\left(E[X_T^{t,x}]\right) \ = \ g^{conv}(x). \end{aligned}$$

En combinant avec (3.57), on en déduit que

$$v(t, x) = g^{conv}(x), \quad \forall (t, x) \in [0, T[\times\mathbb{R}. \tag{3.58}$$

On aboutit à une contradiction dès lors que la fonction g^{conv} n'est pas $C^2(\mathbb{R})$, par exemple si $g(x) = \max(x - \kappa, 0) = g^{conv}(x)$.

Remarque 3.7.7 On verra au chapitre suivant grâce à la théorie des solutions de viscosité que même si l'inégalité (3.55) ne peut pas être interprétée au sens classique, la relation (3.58) reste vraie. En particulier, on voit que la fonction valeur v est discontinue en T dès lors que $g \neq g^{conv}$. En fait, l'Hamiltonien de ce problème de contrôle singulier (3.52) est $H(M) = \sup_{a \in \mathbb{R}}[-\frac{1}{2}a^2 M]$. Ainsi, $H(M) < +\infty$ ssi $-M \leq 0$ et dans ce cas $H(M) = 0$. On montrera alors au chapitre suivant que v est solution de viscosité de l'inégalité variationnelle de HJB :

$$\max\left[-\frac{\partial v}{\partial t}, -\frac{\partial^2 v}{\partial x^2}\right] = 0.$$

.

3.8 Commentaires bibliographiques

Le principe d'optimalité de la programmation dynamique a été initié par Bellman dans les années 50 [Be57]. Simple et intuitif dans son énoncé, la preuve rigoureuse de ce principe est très technique et a été étudié par de nombreux auteurs selon diverses méthodes : Bensoussan et J.L. Lions [BL78], Krylov [Kry80], Nisio [Nis81], P.L. Lions [Lio83], Borkar [Bor89], Fleming et Soner [FSo93].

L'approche classique du contrôle optimal pour les processus de diffusion via un théorème de vérification sur les équations d'Hamilton-Jacobi-Bellman est traitée dans de nombreux livres : Fleming et Rishel [FR75], Krylov [Kry80], Fleming et Soner [FSo93], Yong et Zhou [YZ00]. Le cas de contrôle pour des processus de diffusion avec sauts est traité dans le récent livre d'Oksendal et Sulem [OS04]. D'autres aspects de la programmation dynamique dans un cadre plus général sont étudiés dans le cours de St-Flour par El Karoui [Elk81].

L'exemple du problème de contrôle stochastique singulier est inspiré du problème de la surréplication pour des modèles à volatilité stochastique qui sera développé au chapitre suivant.

Nous avons formulé ici les problèmes de contrôle stochastique sous une forme standard, où il s'agit d'optimiser une fonctionnelle s'exprimant sous forme d'espérance. Récemment, motivé par le problème de la surréplication en finance sous contraintes gamma, Soner et Touzi [ST00], [ST02] ont étudié d'autres formulations de problèmes de contrôle stochastiques et montré le principe de programmation dynamique associé.

L'approche de la programmation dynamique donne une caractérisation de la fonction valeur et du contrôle optimal lorsque cette fonction valeur est régulière. Le problème général d'existence d'un contrôle optimal n'est pas abordé ici. On se reportera aux travaux de Kushner [Ku75] et El Karoui, Huu Nguyen, Jeanblanc-Picqué [ElkNJ87] pour des résultats dans cette direction.

4

Approche des équations de Bellman par les solutions de viscosité

4.1 Introduction

Comme nous l'avons vu au chapitre précédent, la méthode de la programmation dynamique est un outil puissant pour étudier les problèmes de contrôle stochastique, via l'équation d'Hamilton-Jacobi-Bellman. Cependant, dans son approche classique, cette méthode suppose a priori que la fonction valeur soit régulière, ce qui n'est pas nécessairement le cas même pour des cas très simples.

Pour surmonter cette difficulté, Crandall et Lions ont introduit dans les années 80 la notion de solutions de viscosité pour les équations du premier ordre. Cette théorie a été ensuite généralisée aux équations du second ordre. Ce concept fournit un moyen très puissant pour étudier en toute généralité les problèmes de contrôle stochastique et permet de donner une formulation rigoureuse à l'équation d'HJB pour des fonctions supposées seulement localement bornées. En combinant avec des résultats de comparaison pour les solutions de viscosité, on obtient ainsi une caractérisation de la fonction valeur comme l'unique solution de viscosité de l'équation de la programmation dynamique associée.

Ce chapitre est une introduction basique à la notion de solutions de viscosité, et plus précisément aux outils nécessaires pour aborder les problémes de contrôle stochastique. Il existe par ailleurs une vaste littérature sur la théorie des solutions de viscosité et on se réferera par exemple à l'article de référence de Crandall, Ishii et Lions [CIL92] pour un état de l'art. Dans la section 4.2, on définit le concept de solutions de viscosité. Les résultats de comparaison sont discutés dans la section 4.3. Dans la section 4.4, on montre comment le principe de la programmation dynamique permet d'obtenir en toute généralité l'équation d'Hamilton-Jacobi-Bellman pour la fonction valeur au sens de la viscosité. Enfin, les sections 4.5 et 4.6 montrent comment utiliser la notion de solutions de viscosité pour résoudre deux problèmes de contrôle stochastique intervenant en finance.

4.2 Définition des solutions de viscosité

On considére des équations aux dérivées partielles non linéaires du second ordre :

$$F(x, w(x), Dw(x), D^2w(x)) = 0, \qquad x \in \mathcal{O}, \tag{4.1}$$

où $\mathcal{O}$ est un ouvert de $\mathbb{R}^N$ et F est une fonction continue de $\mathcal{O} \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}_N$ dans $\mathbb{R}$. La fonction F doit satisfaire la condition dite d'ellipticité : pour tout $x \in \mathcal{O}$, $r \in \mathbb{R}$, $p \in \mathbb{R}^N$, $M, \widehat{M} \in \mathcal{S}_N$,

$$M \leq \widehat{M} \Longrightarrow F(x, r, p, M) \;\geq\; F(x, r, p, \widehat{M}). \tag{4.2}$$

Dans le cas des problèmes d'évolution, un point de $\mathbb{R}^N$ doit être compris comme variable de temps t et d'espace x dans $\mathbb{R}^n$ $(N = n+1)$, $\mathcal{O}$ doit être vu comme un ouvert de la forme $[0, T[\times \mathcal{O}_n$ où $\mathcal{O}_n$ est un ouvert de $\mathbb{R}^n$ et $F(t, x, r, p_t, p, M)$ doit satisfaire de plus la condition de parabolicité : pour tout $t \in [0, T[$, $x \in \mathcal{O}_n$, $r \in \mathbb{R}$, $p_t, \hat{p}_t \in \mathbb{R}$, $p \in \mathbb{R}^n$, $M \in \mathcal{S}_n$,

$$p_t \leq \hat{p}_t \Longrightarrow F(t, x, r, p_t, p, M) \;\geq\; F(t, x, r, \hat{p}_t, p, M). \tag{4.3}$$

Le cas des équations d'Hamilton-Jacobi-Bellman étudié au chapitre précédent correspond dans le cas de problème à horizon infini à :

$$F(x, r, p, M) = \beta r + \sup_{a \in A} \left[-b(x, a).p - \frac{1}{2} \mathrm{tr}\,(\sigma\sigma'(x, a) M) - f(x, a) \right], \tag{4.4}$$

et dans le cas de problème à horizon fini à :

$$\begin{aligned} F(t, x, r, p_t, p, M) = -p_t + \sup_{a \in A} & \\ & \times \left[-b(x, a).p - \frac{1}{2} \mathrm{tr}\,(\sigma\sigma'(x, a) M) - f(t, x, a) \right]. \end{aligned} \tag{4.5}$$

La condition d'ellipticité est évidemment satisfaite car la matrice $\sigma\sigma'$ est définie nonnégative. La condition de parabolicité est aussi trivialement vérifiée pour le problème à horizon fini.

Etant donné une fonction w localement bornée de $\mathcal{O}$ dans $\mathbb{R}$ (i.e. pour tout x dans $\mathcal{O}$, il existe un voisinage compact V_x de x tel que v soit bornée sur V_x), on définit son envelope semicontinue supérieure $w^* : \bar{\mathcal{O}} \to \mathbb{R}$ par :

$$w^*(x) = \limsup_{x' \to x} w(x')$$

et son envelope semicontinue inférieure w_* par :

$$w_*(x) = \liminf_{x' \to x} w(x').$$

On rappelle que w^* (resp. w_*) est la plus petite (resp. grande) fonction semi-continue supérieurement (s.c.s) majorant (resp. semicontinue inférieurement (s.c.i) minorant) w sur $\mathcal{O}$. Notons qu'une fonction w localement bornée sur $\mathcal{O}$ est semicontinue inférieurement (resp. supérieurement) ssi $w = w_*$ (resp. w^*) sur $\mathcal{O}$, et qu'elle est continue sur $\mathcal{O}$ ssi $w = w_* = w^*$ sur $\mathcal{O}$.

Définition 4.2.1 *Soit $w : \mathcal{O} \to \mathbb{R}$ localement bornée.*
(i) On dit que w est une sous-solution (discontinue) de viscosité de (4.1) *si on a :*

$$F(\bar{x}, w^*(\bar{x}), D\varphi(\bar{x}), D^2\varphi(\bar{x})) \leq 0,$$

en tout point $\bar{x} \in \mathcal{O}$ et pour toute fonction $\varphi \in C^2(\mathcal{O})$ tel que $\bar{x}$ est un maximum local de $w^* - \varphi$.
(ii) On dit que w est une sursolution (discontinue) de viscosité de (4.1) *si on a :*

$$F(\bar{x}, w_*(\bar{x}), D\varphi(\bar{x}), D^2\varphi(\bar{x})) \geq 0,$$

en tout point $\bar{x} \in \mathcal{O}$ et pour toute fonction $\varphi \in C^2(\mathcal{O})$ tel que $\bar{x}$ est un minimum local de $w_* - \varphi$.
(iii) On dit que w est une solution (discontinue) de viscosité de (4.1) *si w est à la fois sous-solution et sursolution de viscosité de* (4.1).

Remarque 4.2.1 **1.** La définition ci-dessus reste inchangée si le maximum ou minimum $\bar{x}$ est local et/ou strict.
2. On peut aussi sans perte de généralité supposer dans la définition ci-dessus que $\varphi(\bar{x}) = w^*(\bar{x})$ (resp. $\varphi(\bar{x}) = w_*(\bar{x})$). En effet, il suffit sinon de considérer la fonction régulière $\psi(x) = \varphi(x) + w^*(\bar{x}) - \varphi(\bar{x})$ (resp. $\psi(x) = \varphi(x) + w_*(\bar{x}) - \varphi(\bar{x})$). Alors $\bar{x}$ est un maximum (resp. minimum) local de $w^* - \psi$ (resp. $w_* - \psi$) et $\psi(\bar{x}) = w^*(\bar{x})$ (resp. $w_*(\bar{x})$).

Remarque 4.2.2 Si v est une sous-solution (resp. sursolution) de viscosité de (4.1) alors v^* (resp. v_*) est une sous-solution s.c.s (resp. sursolution s.c.i) de viscosité de (4.1).

On vérifie d'abord que la notion de solution de viscosité étend bien la notion de solution classique.

Proposition 4.2.1 *Soit $w \in C^2(\mathcal{O})$. Alors w est sursolution (resp. sous-solution) de viscosité de* (4.1) *si et seulement si w est sursolution (resp. sous-solution) classique de* (4.1).

Preuve. On montre le résultat dans le cas d'EDP paraboliques. Supposons d'abord que w est sursolution de viscosité de (4.1). Puisque w est régulière, $\varphi \equiv w$ $(= w_*)$ peut être choisie comme fonction test. De plus, tout point $(t, x) \in [0, T[\times\mathcal{O}$ est un minimum global de $w_* - \varphi \equiv 0$. Donc on a par définition des sursolutions de viscosité :

$$F(t,x,\varphi(t,x),\frac{\partial\varphi}{\partial t}(t,x),D_x\varphi(t,x),D_x^2\varphi(t,x))\geq 0$$

pour tout $(t,x)\in[0,T[\times\mathcal{O}$ et donc $w=\varphi$ est sursolution classique de (4.1).

Supposons que w est sursolution classique de (4.1). Soit $\varphi\in C^2(\mathcal{O})$ et $(\bar{t},\bar{x})\in[0,T[\times\mathcal{O}$ un minimum global de $w_*-\varphi=w-\varphi$. Alors on a :

$$\frac{\partial(w-\varphi)}{\partial t}(\bar{t},\bar{x})\geq 0\quad(=0 \text{ si } \bar{t}>0)$$

$$D_xw(\bar{t},\bar{x})=D_x\varphi(\bar{t},\bar{x}) \text{ et } D_x^2w(\bar{t},\bar{x})\geq D_x^2\varphi(\bar{t},\bar{x}).$$

On en déduit par les conditions (4.2) et (4.3) :

$$F(\bar{t},\bar{x},\varphi(\bar{t},\bar{x}),\frac{\partial\varphi}{\partial t}(\bar{t},\bar{x}),D_x\varphi(\bar{t},\bar{x}),D_x^2\varphi(\bar{t},\bar{x}))$$

$$\geq F(\bar{t},\bar{x},w(\bar{t},\bar{x}),\frac{\partial w}{\partial t}(\bar{t},\bar{x}),D_xw(\bar{t},\bar{x}),D_x^2w(\bar{t},\bar{x}))\ \geq\ 0,$$

c'est à dire que w est sursolution de viscosité de (4.1).

La propriété de sous-solution est prouvée de manière similaire. □

4.3 Principe de comparaison

On dit que l'on a un principe de comparaison fort (pour les solutions discontinues) pour l'EDP (4.1) dans le cas d'un ouvert $\mathcal{O}$ borné si l'énoncé suivant est vrai :

Si v est une sous-solution de viscosité de (4.1) et w est une sursolution de viscosité de (4.1) tel que $v^*\leq w_*$ sur $\partial\mathcal{O}$ alors

$$v^*\leq w_*\quad\text{sur } \bar{\mathcal{O}}.$$

Remarque 4.3.3 Dans le cas d'une EDP elliptique (4.1) dans tout l'espace $\mathcal{O}=\mathbb{R}^N$, on a $\partial\mathcal{O}=\emptyset$ et il faut rajouter des conditions de croissance à l'infini sur v et w, comme par exemple une croissance quadratique. De même, dans le cas d'une EDP parabolique sur le domaine $\mathcal{O}=[0,T[\times\mathbb{R}^n$, on a $\partial\mathcal{O}=\{T\}\times\mathbb{R}^n$, et il faudra aussi rajouter des conditions de croissance à l'infini sur v et w.

Remarque 4.3.4 Le principe de comparaison se formule aussi de manière équivalente : Si v est une sous-solution s.c.s. de viscosité de (4.1) et w est une sursolution s.c.i. de viscosité de (4.1) tel que $v^*\leq w_*$ sur $\partial\mathcal{O}$ alors

$$v\leq w\quad\text{sur } \bar{\mathcal{O}}.$$

Remarque 4.3.5 Comme pour les principes de comparaison classiques (pour les solutions continues), le principe de comparaison fort permet de comparer une sous-solution et une sursolution sur tout le domaine à partir de leur comparaison sur le bord du domaine. En particulier, ceci prouve l'unicité d'une solution de viscosité à (4.1) étant donné une condition au bord : $v^* = v_* = g$ sur $\partial\mathcal{O}$. En effet, si v et w sont deux solutions de viscosité de (4.1) avec la même condition au bord, alors on aura $v^* \leq w_*$ et $w^* \leq v_*$ sur $\mathcal{O}$. Comme par construction, on a déjà $v_* \leq v^*$ et $w_* \leq w^*$, ceci implique les égalités suivantes :

$$v_* \ = \ v^* = w_* \ = \ w^*,$$

et donc $v = w$ sur $\mathcal{O}$. On a aussi montré en même temps que v et w sont continues sur $\mathcal{O}$. Ainsi, le résultat de comparaison fort est un outil très puissant qui permet de montrer en plus qu'une solution à priori discontinue est en fait continue sur $\mathcal{O}$.

On donne ci-dessous quelques exemples de fonctions F pour lesquels il y a un principe de comparaison fort. Des résultats généraux avec leurs preuves peuvent être trouvées dans Crandall, Ishii et P.L.Lions [CIL92] ou Barles [Ba95].

• On considère d'abord le cas où $\mathcal{O}$ est un ouvert borné de $\mathbb{R}^N$.

(1) $F(x, r, p, M) = \beta r + H(x, p) - \frac{1}{2}\mathrm{tr}\,(\sigma\sigma'(x)M)$ avec $\beta > 0$, $\sigma : \mathcal{O} \to \mathbb{R}^{N\times d}$ Lipschitzienne et $H : \mathcal{O} \times \mathbb{R}^N \to \mathbb{R}$ vérifiant l'hypothèse suivante :

(A1) $|H(x, p) - H(y, p)| \leq m(|x - y|(1 + |p|))$, où $m(z)$ tend vers zéro quand z tend vers zéro.

(2) $F(x, r, p, M) = H(x, p)$ avec $H : \mathcal{O} \times \mathbb{R}^N \to \mathbb{R}$ vérifiant **(A1)** et les hypothèses supplémentaires :

(A2) $H(x, p)$ est convexe en p, pour tout $x \in \mathcal{O}$

(A3) Il existe une fonction $\varphi \in C^1(\mathcal{O})$, continue sur $\bar{\mathcal{O}}$, et $\delta > 0$ telle que $H(x, D\varphi(x)) \leq -\delta$ sur $\mathcal{O}$.

• On considère maintenant le cas d'EDP (4.1) sur l'espace entier $\mathcal{O} = \mathbb{R}^N$ dans le cas elliptique, et $\mathcal{O} = [0, T[\times\mathbb{R}^n$, dans le cas parabolique, reliées à des problèmes de contrôle stochastique.

(3) $F(x, r, p, M)$ donné par (4.4) avec $\beta > 0$, A compact et b, σ, f vérifiant une condition de Lipschitz en x uniformément en a dans A.

(4) $F(t, x, r, p_t, p, M)$ donné par (4.5) avec A compact et b, σ, f vérifiant une condition de Lipschitz en x uniformément en t dans $[0, T]$ et a dans A.

Dans chacun des cas (3) ou (4), il y a un principe de comparaison pour les solutions de viscosité dans la classe des fonctions à croissance quadratique. En particulier, ceci prouve l'unicité d'une solution de viscosité continue à croissance quadratique de l'EDP (4.1) pour F défini en (4.4) et l'unicité d'une solution de viscosité continue à croissance quadratique de l'EDP (4.1) avec une condition terminale donnée $v_*(T,.) = v^*(T,.) = g$ (à croissance quadratique), pour F défini en FHJBfin.

4.4 De la programmation dynamique aux solutions de viscosité

Dans cette section, on revient au cadre des problèmes de contrôle stochastique pour les diffusions considérées à la section 3.2 du chapitre 3. On rappelle que l'Hamiltonien pour ces problèmes est donné par :

$$H(t,x,p,M) = \sup_{a\in A}\left[-b(x,a).p - \frac{1}{2}\mathrm{tr}\left(\sigma(x,a)\sigma'(x,a)M\right) - f(t,x,a)\right]. \quad (4.6)$$

Dans le cas à horizon infini, f donc H ne dépend pas de t.

La thérorie des solutions de viscosité permet de relacher la régularité a priori de la fonction valeur et de montrer que celle-ci satisfait l'équation d'HJB au sens de la viscosité. On montre d'abord la propriété de sous-solution.

Sous-solution

Proposition 4.4.2 *1) Horizon fini : Supposons que la fonction valeur v soit localement bornée sur $[0,T[\times\mathbb{R}^n$, que la fonction f soit à croissance quadratique* (3.8)*, et que $f(.,.,a)$ soit continue en (t,x) pour tout $a\in A$. Alors v est une sous-solution de viscosité de l'équation d'Hamilton-Jacobi-Bellman :*

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) = 0, \quad \forall (t,x)\in[0,T[\times\mathbb{R}^n. \quad (4.7)$$

2) Horizon infini : Supposons que la fonction valeur v soit localement bornée, que la fonction f soit à croissance quadratique (3.13)*, et que $f(.,a)$ soit continue en x pour tout $a\in A$. Alors pour $\beta>0$ assez grand, v est une sous-solution de viscosité de l'équation d'Hamilton-Jacobi-Bellman :*

$$\beta v(x) + H(x, Dv(x), D^2v(x)) = 0, \quad \forall x\in\mathbb{R}^n. \quad (4.8)$$

Preuve. On montre le résultat dans le cas d'un horizon fini. Soit $(\bar t,\bar x)\in[0,T[\times\mathbb{R}^n$ et $\varphi\in C^2([0,T[\times\mathbb{R}^n)$ une fonction test tel que :

$$0 = (v^*-\varphi)(\bar t,\bar x) = \max_{(t,x)\in[0,T[\times\mathbb{R}^n}(v^*-\varphi)(t,x). \quad (4.9)$$

Par définition de $v^*(\bar t,\bar x)$, il existe une suite (t_m,x_m) dans $[0,T[\times\mathbb{R}^n$ telle que

$$(t_m,x_m) \to (\bar t,\bar x) \text{ et } v(t_m,x_m) \to v^*(\bar t,\bar x),$$

quand m tend vers l'infini. Par continuité de φ et (4.9), on a aussi que

$$\gamma_m \ := \ v(t_m, x_m) - \varphi(t_m, x_m) \to 0,$$

quand m tend vers l'infini.

Soit $a \in A$, α le contrôle constant égal à a qui est bien dans $\mathcal{A}(t_m, x_m) = \mathcal{A}$ d'après la remarque 3.2.1. On note $X_s^{t_m,x_m}$ le processus d'état contrôlé associé. Soit τ_m le temps d'arrêt : $\tau_m = \inf\{s \geq t_m \ : |X_s^{t_m,x_m} - x_m| \geq \eta\}$ où $\eta > 0$ est une constante fixé. Soit (h_m) une suite strictement positive telle que :

$$h_m \ \to \ 0 \text{ et } \frac{\gamma_m}{h_m} \ \to \ 0,$$

quand m tend vers l'infini. On applique la première partie (3.18) du principe de la programmation dynamique pour $v(t_m, x_m)$ à $\theta_m := \tau_m \wedge (t_m + h_m)$:

$$v(t_m, x_m) \leq E\left[\int_{t_m}^{\theta_m} f\left(s, X_s^{t_m,x_m}, a\right) ds + v\left(\theta_m, X_{\theta_m}^{t_m,x_m}\right)\right].$$

D'après la relation (4.9) qui implique que $v \leq v^* \leq \varphi$, on en déduit :

$$\varphi(t_m, x_m) + \gamma_m \leq E\left[\int_{t_m}^{\theta_m} f\left(s, X_s^{t_m,x_m}, a\right) ds + \varphi\left(\theta_m, X_{\theta_m}^{t_m,x_m}\right)\right].$$

On applique alors la formule d'Itô à $\varphi(s, X_s^{t_m,x_m})$ entre t_m et θ_m, et on obtient après avoir noté que le terme d'intégrale stochastique s'annule en espérance pour cause d'intégrand borné :

$$\frac{\gamma_m}{h_m} \ + \ E\left[\frac{1}{h_m}\int_{t_m}^{\theta_m}\left(-\frac{\partial \varphi}{\partial t} - \mathcal{L}^a\varphi - f\right)\left(s, X_s^{t_m,x_m}, a\right) ds\right] \leq 0. \quad (4.10)$$

Par continuité p.s. de la trajectoire $X_s^{t_m,x_m}$, on a que pour m suffisamment grand ($m \geq N(\omega)$), $\theta_m(\omega) = t_m + h_m$, p.s. On en déduit par le théorème de la moyenne que la variable aléatoire sous le signe espérance de (4.10) converge p.s. vers $-\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{x}) - \mathcal{L}^a\varphi(\bar{t}, \bar{x}) - f(\bar{t}, \bar{x}, a)$ quand m tend vers l'infini. De plus, cette variable aléatoire est bornée par une constante indépendante de m. Par le théorème de convergence dominée, on obtient alors en envoyant m vers l'infini dans (4.10) :

$$-\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{x}) - \mathcal{L}^a\varphi(\bar{t}, \bar{x}) - f(\bar{t}, \bar{x}, a) \leq 0,$$

et on conclut puisque $a \in A$ est arbitraire.

Dans le cas à horizon infini, on suppose que $\beta > 0$ est assez grand pour que les contrôles constants soient dans $\mathcal{A}(x)$, pour tout x dans $\mathbb{R}^n$ (voir remarque 3.2.2). On fait ensuite le même raisonnement que pour le cas à horizon fini. □

Sursolution

L'inégalité inverse de sursolution de l'équation de HJB est obtenue sous une condition supplémentaire qui prend en compte le fait que l'Hamiltonien peut exploser lorsque typiquement l'espace des contrôles A est non borné. On note

$$\mathrm{dom}(H) = \{(t, x, p, M) \in [0, T[\times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n \ : H(t, x, p, M) < +\infty\},$$

et on fait l'hypothèse suivante :

$$\begin{aligned} H \text{ est continue sur } \mathrm{int}(\mathrm{dom}(H)) \\ \text{et il existe } G : [0, T[\times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n \text{ continue telle que } : \\ (t, x, p, M) \in \mathrm{dom}(H) \iff G(t, x, p, M) \leq 0 \end{aligned} \quad (4.11)$$

Bien évidemment, dans le cas de problème à horizon infini, la fonction H et donc G ne dépendent pas de t.

Proposition 4.4.3 *1) Horizon fini : Supposons que* (4.11) *soit vérifiée et que la fonction valeur v soit localement bornée sur $[0, T[\times \mathbb{R}^n$. Alors v est une sursolution de viscosité de l'inéquation variationnelle d'Hamilton-Jacobi-Bellman :*

$$\max\left\{-\frac{\partial v}{\partial t}(t, x) + H(t, x, D_x v(t, x), D_x^2 v(t, x)), G(t, x, D_x v(t, x), D_x^2 v(t, x))\right\} \\ = 0, \quad (t, x) \in [0, T[\times \mathbb{R}^n. \quad (4.12)$$

2) Horizon infini : Supposons que (4.11) *soit vérifiée et que la fonction valeur v soit localement bornée sur $\mathbb{R}^n$. Alors v est une sursolution de viscosité de :*

$$\max\left\{\beta v(x) + H(x, Dv(x), D^2 v(x)), G(x, Dv(x), D^2 v(x))\right\} = 0, \ x \in \mathbb{R}^n. \quad (4.13)$$

Preuve. On montre le résultat dans le cas d'un horizon infini. Soit $\bar{x} \in \mathbb{R}^n$ et $\varphi \in C^2(\mathbb{R}^n)$ une fonction test tel que :

$$0 = (v_* - \varphi)(\bar{x}) = \min_{x \in \mathbb{R}^n} (v_* - \varphi)(x). \quad (4.14)$$

On va prouver le résultat par l'absurde en supposant donc au contraire que

$$\begin{aligned} \beta\varphi(\bar{x}) + H(\bar{x}, D\varphi(\bar{x}), D^2\varphi(\bar{x})) &< 0, \\ \text{et } G(\bar{x}, D\varphi(\bar{x}), D^2\varphi(\bar{x})) &< 0. \end{aligned}$$

Alors par continuité de la fonction G, et celle de H sur l'intérieur de son domaine, il existe $\eta > 0$ et $\varepsilon > 0$ tels que :

$$\beta\varphi(y) + H(y, D\varphi(y), D^2\varphi(y)) \leq -\varepsilon,$$

pour tout $y \in B(\bar{x}, \eta) = \{y \in \mathbb{R}^n \ : |\bar{x} - y| < \eta\}$. Par définition de $v_*(\bar{x})$, il existe une suite x_m à valeurs dans $B(\bar{x}, \eta)$ telle que

$$x_m \;\to\; \bar{x} \text{ et } v(x_m) \;\to v_*(\bar{x}),$$

quand m tend vers l'infini. Par continuité de φ et (4.14), on a aussi que

$$\gamma_m \;:=\; v(x_m) - \varphi(x_m) \to 0,$$

quand m tend vers l'infini. Soit (h_m) une suite strictement positive telle que :

$$h_m \;\to\; 0 \text{ et } \frac{\gamma_m}{h_m} \;\to\; 0,$$

Alors, d'après la 2ème partie (3.19) de la programmation dynamique et en utilisant aussi (4.14), il existe $\hat{\alpha}^m \in \mathcal{A}(x_m)$, tel que

$$\varphi(x_m) + \gamma_m + \frac{\varepsilon h_m}{2} \geq E\left[\int_0^{\theta_m} e^{-\beta s} f(X_s^{x_m}, \hat{\alpha}_s^m)ds + e^{-\beta\theta_m}\varphi(X_{\theta_m}^{x_m})\right],$$

où l'on a choisi $\theta_m := \tau_m \wedge h_m$, $\tau_m = \inf\{s \geq 0 \,:\, |X_s^{x_m} - x_m| \geq \eta'\}$ et $0 < \eta' < \eta$. Puisque (x_m) tend vers $\bar{x}$, on peut toujours supposer que $B(x_m, \eta') \subset B(\bar{x}, \eta)$, de telle sorte que pour $0 \leq s < \theta_m$, $X_s^{x_m} \in B(\bar{x}, \eta)$. Ici $X_s^{x_m}$ correspond à la diffusion contrôlée par $\hat{\alpha}^m$. Par la formule d'Itô à $e^{-\beta s}\varphi(X_s^{x_m})$ entre $s = 0$ et $s = \theta_m$, on obtient alors :

$$\begin{aligned} 0 \leq \frac{\gamma_m}{h_m} + \frac{\varepsilon}{2} \;+\; & E\left[\frac{1}{h_m}\int_0^{\theta_m} L(X_s^{x_m}, \hat{\alpha}_s^m)ds\right] \\ & - E\left[\int_0^{\theta_m} D_x\varphi(X_s^{x_m})'\sigma(X_s^{x_m}, \hat{\alpha}_s^m)dW_s\right] \end{aligned} \tag{4.15}$$

avec

$$L(x,a) = \beta\varphi(x) - \mathcal{L}^a\varphi(x) - f(x,a).$$

On remarque que d'après la condition (3.2) sur σ, l'intégrand de l'intégrale stochastique en dW est borné sur $[0, \theta_m]$ par :

$$|D_x\varphi(X_s^{x_m})'\sigma(X_s^{x_m}, \hat{\alpha}_s^m)| \leq C_\eta(1 + |\sigma(0, \hat{\alpha}_s^m)|).$$

D'après la condition d'intégrabilité (3.3) sur $\hat{\alpha}^m \in \mathcal{A}(x_m)$, on en déduit que l'espérance de l'intégrale stochastique dans (4.15) est nulle.

De plus, en remarquant que pour $0 \leq s < \theta_m$:

$$\begin{aligned} L(X_s^{x_m}, \hat{\alpha}_s^m) &\leq \beta\varphi(X_s^{x_m}) + H(X_s^{x_m}, D\varphi(X_s^{x_m}), D^2\varphi(X_s^{x_m})) \\ &\leq -\varepsilon, \end{aligned}$$

on en déduit d'après (4.15) que

$$0 \leq \frac{\gamma_m}{h_m} + \varepsilon\left(\frac{1}{2} - \frac{1}{h_m}E[\theta_m]\right). \tag{4.16}$$

D'après l'inégalité de Chebicheff et l'estimation (3.5), on a :

$$\begin{aligned} P[\tau_m \leq h_m] &\leq P\left[\sup_{s\in[0,h_m]} |X_s^{x_m} - x_m| \geq \eta\right] \\ &\leq \frac{E\left|\sup_{s\in[0,h_m]} |X_s^{x_m} - x_m|^2\right|}{\eta^2} \\ &\to 0, \end{aligned}$$

quand h_m tend vers zéro, i.e. m tend vers l'infini. De plus, comme

$$P[\tau_m > h_m] \leq \frac{1}{h_m} E[\theta_m] \leq 1,$$

on en déduit que $\frac{1}{h_m} E[\theta_m]$ tend vers 1 quand h_m tend vers zéro. On obtient ainsi la contradiction voulue en faisant tendre m vers l'infini dans (4.16). □

Propriété de viscosité

En combinant les deux propositions précédentes, on obtient le résultat principal de cette section.

Théorème 4.4.1 *Sous les conditions des propositions 4.4.2 et 4.4.3, la fonction valeur v est solution de viscosité de l'inéquation variationnelle d'Hamilton-Jacobi-Bellman :*
1) Horizon fini :

$$\begin{aligned} &\max\left\{-\frac{\partial v}{\partial t}(t,x)+H(t,x,D_x v(t,x),D_x^2 v(t,x))\ ,\ G(t,x,D_x v(t,x),D_x^2 v(t,x))\right\} \\ &\quad = 0, \quad (t,x) \in [0,T[\times\mathbb{R}^n. \end{aligned} \tag{4.17}$$

2) Horizon infini :

$$\begin{aligned} \max\{\beta v(x) + H(x, Dv(x), D^2 v(x)), G(x, Dv(x), \\ D^2 v(x))\} = 0,\ x \in \mathbb{R}^n. \end{aligned} \tag{4.18}$$

Preuve. La propriété 4.4.2 de sous-solution de viscosité de v signifie, dans le cas à horizon fini, que :

$$-\frac{\partial \varphi}{\partial t}(\bar t, \bar x) + H(\bar t, \bar x, D_x\varphi(\bar t, \bar x), D_x^2\varphi(\bar t, \bar x)) \leq 0,$$

en tout $(\bar t, \bar x) \in [0,T[\times\mathbb{R}^n$ et $\varphi \in C^2([0,T[\times\mathbb{R}^n)$ tel que $(\bar t, \bar x)$ soit un maximum local de $v - \varphi$. Ceci implique d'après la condition (4.11) que

$$G(\bar t, \bar x, D_x\varphi(\bar t, \bar x), D_x^2\varphi(\bar t, \bar x)) \leq 0,$$

et donc que

$$\max\left\{-\frac{\partial \varphi}{\partial t}(\bar{t},\bar{x}) + H(\bar{t},\bar{x},D_x\varphi(\bar{t},\bar{x}),D_x^2\varphi(\bar{t},\bar{x})),\right.$$
$$\left. G(\bar{t},\bar{x},D_x\varphi(\bar{t},\bar{x}),D_x^2\varphi(\bar{t},\bar{x}))\right\} \leq 0.$$

Autrement dit, v est une sous-solution de viscosité de l'inéquation variationnelle (4.17). On conclut avec la propriété de sursolution de la proposition 4.4.3. □

Remarque 4.4.6 1. Lorsque A est compact, l'Hamiltonien H est fini sur tout le domaine $[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_n$ et le supremum est atteint. La condition (4.11) est satisfaite avec n'importe quel choix de fonction G continue strictement négative. L'inéquation variationnelle d'Hamiton-Jacobi-Bellman que doit satifaire (au sens de la viscosité) la fonction valeur devient dans ce cas simplement l'équation d'HJB :

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x),D_x^2 v(t,x)) = 0, \quad \forall (t,x)\in[0,T[\times\mathbb{R}^n.$$

2. Lorsque l'Hamiltonien n'est pas défini sur tout $[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_n$, on a un problème de contrôle stochastique singulier qui conduit à une inéquation variationnelle. On en verra deux exemples dans la section suivante.
3. Le théorème précédent combiné avec les résultats de comparaison forts montrent donc que la fonction valeur v d'un problème de contrôle stochastique est caractérisée comme l'unique solution de viscosité (avec des conditions de croissance à l'infini appropriées) de l'(in)équation (variationnelle) d'HJB associée avec une condition terminale $v_* = v^*$. La théorie des solutions de viscosité prouve toute sa puissance, ici en particulier dans le cadre des problèmes de contrôle stochastique. Elle peut être utilisée comme une méthode de vérification où il "suffira" de résoudre l'équation d'HJB et de déterminer la condition terminale pour calculer la fonction valeur. On donne une application dans la section suivante.

4.5 Un modèle d'investissement irréversible

4.5.1 Problème

On considère une firme produisant un bien (électicité, pétrole, ...). Elle peut augmenter sa capacité de production X_t en transférrant du capital à partir d'un autre secteur d'activité. L'évolution contrôlée de sa capacité de production évolue alors selon :

$$dX_t = X_t(-\delta dt + \sigma dW_t) + l_t dt.$$

$\delta \geq 0$ est le taux de dépréciation de la production, $\sigma > 0$ sa volatilité, $l_t dt$ est le nombre d'unités de capital acquis par la firme à un coût $\lambda l_t dt$, $\lambda > 0$ s'interprétant comme un facteur de conversion d'un secteur d'activité à l'autre. Le contrôle l_t est positif, non borné à valeurs dans $\mathbb{R}_+$, et il est commode de poser $L_t = \int_0^t l_t dt$: plus généralement, on considére comme contrôle $(L_t)_{t\geq 0}$ processus adapté, croissant, continue à droite, $L_{0^-} = 0$, et on note $L \in \mathcal{A}$: L_t représente le montant cumulé d'achat de capital jusquà la date t. Pour $x \geq 0$ et $L \in \mathcal{A}$, on note par X^x la solution de

$$dX_t = X_t(-\delta dt + \sigma dW_t) + dL_t, \quad X_{0^-} = x. \tag{4.19}$$

La fonction de profit de la firme est une fonction continue Π de $\mathbb{R}_+$ dans $\mathbb{R}$, concave, croissante et C^1 sur $]0, +\infty[$, avec $\Pi(0) = 0$ et satisfaisant les conditions usuelles d'Inada en 0 :

$$\Pi'(0^+) := \lim_{x\downarrow 0} \Pi'(x) \; = \; +\infty \;\text{ et }\; \Pi'(+\infty) := \lim_{x\to+\infty} \Pi'(x) \; = \; 0. \tag{4.20}$$

On considère la transformée de Fenchel-Legendre de Π qui est donc finie sur $\mathbb{R}_+$:

$$\tilde{\Pi}(y) := \sup_{x\geq 0} [\Pi(x) - xy] \; < \; +\infty, \quad \forall y > 0. \tag{4.21}$$

L'objectif de la firme est le suivant :

$$v(x) = \sup_{L \,\in\, \mathcal{A}} E\left[\int_0^{+\infty} e^{-\beta t}\left(\Pi(X_t^x)dt \; - \; \lambda dL_t\right)\right], \qquad x \geq 0. \tag{4.22}$$

L'Hamiltonien de ce problème de contrôle stochastique singulier est :

$$H(x,p,M) = \begin{cases} \delta x p - \frac{1}{2}\sigma^2 x^2 M - \Pi(x) & \text{si } \lambda - p \geq 0 \\ -\infty & \text{si } \lambda - p < 0. \end{cases}$$

L'inéquation variationnelle d'HJB s'écrit donc :

$$\min\left[\beta v - \mathcal{L}v - \Pi \;,\; \lambda - v'\right] = 0, \tag{4.23}$$

où

$$\mathcal{L}v = -\delta x v' + \frac{1}{2}\sigma^2 x^2 v''.$$

4.5.2 Régularité et construction de la fonction valeur

On commence par établir quelques propriétés utiles de la fonction valeur.

Lemme 4.5.2 *a) v est finie sur $\mathbb{R}_+$ et satisfait pour tout $\mu \in [0, \lambda]$:*

$$0 \leq \; v(x) \; \leq \frac{\tilde{\Pi}((\beta+\delta)\mu)}{\beta} + \mu x, \quad x \geq 0. \tag{4.24}$$

b) v est concave et continue sur $\mathbb{R}_+^*$.

Preuve. a) En considérant le contrôle nul $L = 0$, il est clair que $v \geq 0$. D'autre part, considérons pour $\mu \in [0, \lambda]$ la fonction positive

$$\varphi(x) = \mu x + \frac{\tilde{\Pi}((\beta + \delta)\mu)}{\beta}.$$

Alors $\varphi' \leq \lambda$ et on a pour tout $x \geq 0$,

$$\beta\varphi - \mathcal{L}\varphi - \Pi(x) = \tilde{\Pi}((\beta + \delta)\mu) + (\beta + \delta)\mu x - \Pi(x) \ \geq \ 0,$$

par définition de $\tilde{\Pi}$ en (4.21). Etant donné $L \in \mathcal{A}$, considérons le temps d'arrêt $\tau_n = \inf\{t \geq 0 \ : X_t^x \geq n\} \wedge n$, $n \in \mathbb{N}$, et appliquons la formule d'Itô à $e^{-\beta t}\varphi(X_t^x)$ entre 0 et τ_n. En prenant l'espérance et en notant que l'intégrand de l'intégrale stochastique est bornée sur $[0, \tau_n)$, on a

$$\begin{aligned} E\left[e^{-\beta\tau_n}\varphi(X_{\tau_n}^x)\right] &= \varphi(x) + E\left[\int_0^{\tau_n} e^{-\beta t}\left(-\beta\varphi + \mathcal{L}\varphi\right)(X_t^x)dt\right] \\ &+ E\left[\int_0^{\tau_n} e^{-\beta t}\varphi'(X_t^x)dL_t^c\right] + E\left[\sum_{0\leq t\leq\tau_n} e^{-\beta t}\left[\varphi(X_t^x) - \varphi(X_{t-}^x)\right]\right], \end{aligned} \tag{4.25}$$

où L^c est la partie continue de L. Puisque $\varphi' \leq \lambda$ et $X_t^x - X_{t-}^x = L_t - L_{t-}$, le théorème de la valeur moyenne implique que

$$\varphi(X_t^x) - \varphi(X_{t-}^x) \leq \lambda(L_t - L_{t-}).$$

En utilisant de nouveau l'inégalité $\varphi' \leq \lambda$ dans les intégrales en dL^c dans (4.25) et puisque $-r\varphi + \mathcal{L}\varphi \leq -\Pi$, on obtient :

$$\begin{aligned} E\left[e^{-\beta\tau_n}\varphi(X_{\tau_n}^x)\right] &\leq \varphi(x) \ - \ E\left[\int_0^{\tau_n} e^{-\beta t}\Pi(X_t^x)dt\right] \\ &\quad + E\left[\int_0^{\tau_n} e^{-\beta t}\lambda dL_t^c\right] \ + \ E\left[\sum_{0\leq t\leq\tau_n} e^{-\beta t}\lambda(L_t - L_{t-})\right] \\ &= \varphi(x) \ - \ E\left[\int_0^{\tau_n} e^{-\beta t}\Pi(X_t^x)dt\right] \ + \ E\left[\int_0^{\tau_n} e^{-\beta t}\lambda dL_t\right], \end{aligned}$$

et donc

$$E\left[\int_0^{\tau_n} e^{-\beta t}\left(\Pi(X_t^x)dt - \lambda dL_t\right)\right] \ + \ E\left[e^{-\beta\tau_n}\varphi(X_{\tau_n}^x)\right] \leq \varphi(x).$$

Puisque $\varphi \geq 0$, on a alors :

$$\varphi(x) \geq E\left[\int_0^{\tau_n} e^{-\beta t}\Pi(X_t^x)dt\right] \ - \ E\left[\int_0^{+\infty} e^{-\beta t}\lambda dL_t\right].$$

On peut appliquer le lemme de Fatou et obtenir en envoyant n vers l'infini :

$$E\left[\int_0^{+\infty} e^{-\beta t}\left(\Pi(X_t^x)dt - \lambda dL_t\right)\right] \leq \varphi(x),$$

et donc finalement $v(x) \leq \varphi(x)$ puisque L est quelconque.

b) La preuve de la concavité de v est standard : elle est établie (comme au paragraphe 3.6.1) en considérant des combinaisons convexes des états initiaux et des contrôles et en utilisant la linéarité de la dynamique de X et la concavité de Π. La continuité de v sur $\mathbb{R}_+$ est alors une conséquence générale de la propriété de continuité d'une fonction concave sur l'intérieur de son domaine. □

Puisque v est concave sur $\mathbb{R}_+^*$, elle admet une dérivée à droite $v'_+(x)$ et à gauche $v'_-(x)$ en tout $x > 0$, avec $v'_+(x) \leq v'_-(x)$.

Lemme 4.5.3

$$v'_-(x) \leq \lambda, \quad x > 0. \tag{4.26}$$

Preuve. Pour tous $x > 0$, $0 < l\ x$, $L \in \mathcal{A}$, considérons le contrôle $\tilde{L}$ défini par $\tilde{L}_{0^-} = 0$, et $\tilde{L}_t = L_t + l$, pour $t \geq 0$. Soit $\tilde{X}$ la diffusion contrôlée avec $\tilde{L}$ et condition initiale $\tilde{X}_{0^-} = x - l$. Alors, $\tilde{X}_t^{x-l} = X_t^x$ pour $t \geq 0$ et on a :

$$\begin{aligned} v(x-l) &\geq E\left[\int_0^{+\infty} e^{-\beta t}\left(\Pi(\tilde{X}_t^{x-l})dt - \lambda d\tilde{L}_t\right)\right] \\ &= E\left[\int_0^{+\infty} e^{-\beta t}\left(\Pi(X_t^x)dt - \lambda dL_t\right)\right] - \lambda l. \end{aligned}$$

L étant arbitraire, on obtient :

$$v(x-l) \geq -\lambda l + v(x), \quad x > 0.$$

On obtient le résultat voulu en divisant par l et en faisant tendre l vers zéro. □

Lemme 4.5.4 *Il existe* $x_b \in [0,\infty]$ *tel que :*

$$\mathcal{NT} := \left\{x > 0\ : v'_-(x) < \lambda\right\} =]x_b, +\infty[, \tag{4.27}$$

v *est différentiable sur* $]0, x_b[$ *et*

$$v'(x) = \lambda, \quad \textit{sur}\ \ \mathcal{B} =]0, x_b[. \tag{4.28}$$

Preuve. On pose $x_b = \inf\{x \geq 0\ : v'_+(x) < \lambda\}$. Alors, on a $\lambda \leq v'_+(x) \leq v'_-(x)$, pour tout $x < x_b$. En combinant avec (4.26), ceci prouve (4.28). Finalement la concavité de v montre (4.27). □

Par le principe de la programmation dynamique, on sait que la fonction valeur v est solution de viscosité de l'inéquation variationnelle d'HJB associée. En exploitant la concavité de v pour ce problème unidimensionnel, on montre que la fonction valeur est en fait une solution régulière C^2.

Théorème 4.5.2 *La fonction valeur v est une solution classique C^2 sur $\mathbb{R}_+^*$ de :*

$$\min\{\beta v - \mathcal{L}v - \Pi(x)\ ,\ -v'(x) + \lambda\} = 0, \qquad x > 0.$$

Preuve. *Etape 1.* On prouve d'abord que v est C^1 sur $\mathbb{R}_+^*$. Puisque v est concave, ses dérivées à gauche et à droite $v'_-(x)$ et $v'_+(x)$ existent pour tout $x > 0$ avec $v'_+(x) \leq v'_-(x)$. On raisonne par contradiction en supposant donc que $v'_+(x_0) < v'_-(x_0)$ pour un certain $x_0 > 0$. On fixe $\mu \in]v'_+(x_0), v'_-(x_0)[$ et on considère la fonction

$$\varphi_\varepsilon(x) = v(x_0) + \mu(x - x_0) - \frac{1}{2\varepsilon}(x - x_0)^2,$$

avec $\varepsilon > 0$. Alors x_0 est un maximum local de $(v - \varphi_\varepsilon)$ avec $\varphi_\varepsilon(x_0) = v(x_0)$. Puisque $\varphi'_\varepsilon(x_0) = \mu < \lambda$ d'après (4.26) et $\varphi''_\varepsilon(x_0) = 1/\varepsilon$, la propriété de sous-solution de viscosité implique :

$$\beta\varphi(x_0) + \delta x_0 \mu + \frac{1}{2\varepsilon}\sigma^2 x_0^2 - \Pi(x_0) \leq 0. \tag{4.29}$$

En prenant ε suffisamment petit, on a la contradiction voulue ce qui prouve que $v'_+(x_0) = v'_-(x_0)$.

Etape 2. D'après le lemme 4.5.4, v est C^2 sur $]0, x_b[$ et satisfait $v'(x) = \lambda$, $x \in]0, x_b[$. D'après l'étape 1, on a $\mathcal{NT} =]x_b, +\infty[= \{x > 0\ : v'(x) < \lambda\}$. Vérifions que v est une solution de viscosité de :

$$\beta v - \mathcal{L}v - \Pi = 0, \quad \text{sur }]x_b, +\infty[. \tag{4.30}$$

Soit $x_0 \in]x_b, +\infty[$ et φ une fonction C^2 sur $]x_b, +\infty[$ telle que x_0 soit un maximum local de $v - \varphi$, avec $(v - \varphi)(x_0) = 0$. Puique $\varphi'(x_0) = v'(x_0) < \lambda$, la propriété de sous-solution de viscosité implique :

$$\beta\varphi(x_0) - \mathcal{L}\varphi(x_0) - \Pi(x_0) \leq 0.$$

Ceci prouve que v est une sous-solution de viscosité de (4.30) sur $]x_b, +\infty[$. La preuve de la propriété de sursolution de viscosité pour (4.30) est similaire. Considérons pour $x_1 \leq x_2$ arbitraires dans $]x_b, +\infty[$, le problème de Dirichlet :

$$\beta V - \mathcal{L}V - \Pi(x) = 0, \quad \text{sur }]x_1, x_2[\tag{4.31}$$

$$V(x_1) = v(x_1), \quad V(x_2) = v(x_2). \tag{4.32}$$

Des résultats classiques donnent l'existence et l'unicité d'une function V régulière C^2 sur $]x_1, x_2[$ solution de (4.31)-(4.32). En particulier, cette solution régulière V est une solution de viscosité de (4.30) sur $]x_1, x_2[$. D'après le principe de comparaison sur les solutions de viscosité (ici pour des EDP linéaires sur un domaine borné), on en déduit que $v = V$ sur $]x_1, x_2[$. Comme x_1 et x_2 sont quelconques dans $]x_b, +\infty[$, ceci prouve que v est C^2 sur $]x_b, +\infty[$ et satisfait (4.30) au sens classique.

Etape 3. Il reste à prouver la condition C^2 en x_b dans le cas où $0 < x_b < +\infty$. Soit $x \in]0, x_b[$. Puisque v est C^2 sur $]0, x_b[$ avec $v'(x) = \lambda$, la propriété de sursolution de viscosité appliqué au point x et à la fonction test $\varphi = v$ implique que v satisfait au sens classique :

$$\beta v(x) - \mathcal{L}v(x) - \varPi(x) \geq 0, \quad 0 < x < x_b.$$

Comme la dérivée de v est constante égale à λ sur $]0, x_b[$, on a :

$$\beta v(x) + \delta x \lambda - \varPi(x) \geq 0, \quad 0 < x < x_b,$$

et donc :

$$\beta v(x_b) + \delta x_b \lambda - \varPi(x_b) \geq 0. \tag{4.33}$$

D'autre part, d'après la condition C^1 en x_b, on a en envoyant x vers x_b dans (4.30) :

$$\beta v(x_b) + \delta x_b \lambda - \varPi(x_b) = \frac{1}{2}\sigma^2 x_b^2 v''(x_b^+). \tag{4.34}$$

D'après la concavité de v, le terme de droite de (4.34) est négatif, ce qui combiné avec (4.33) prouve que $v''(x_b^+) = 0$. Ceci montre que v est C^2 en x_b avec

$$v''(x_b) = 0.$$

□

Nous pouvons à présent expliciter la forme de la fonction valeur. Considérons l'équation différentielle ordinaire (EDO) intervenant dans HJB :

$$\beta v - \mathcal{L}v - \varPi = 0. \tag{4.35}$$

On rappelle que la solution générale de (4.35) (avec $\varPi = 0$) est donnée par :

$$\hat{V}(x) = Ax^m + Bx^n,$$

où

$$\begin{aligned} m &= \frac{\delta}{\sigma^2} + \frac{1}{2} - \sqrt{\left(\frac{\delta}{\sigma^2} + \frac{1}{2}\right)^2 + \frac{2\beta}{\sigma^2}} < 0 \\ n &= \frac{\delta}{\sigma^2} + \frac{1}{2} + \sqrt{\left(\frac{\delta}{\sigma^2} + \frac{1}{2}\right)^2 + \frac{2\beta}{\sigma^2}} > 1. \end{aligned}$$

De plus, l'EDO (4.35) admet une solution particulière donnée par :

$$\begin{aligned}\hat{V}_0(x) &= E\left[\int_0^{+\infty} e^{-\beta t}\Pi(\hat{X}_t^x)\right] \\ &= \frac{2}{\sigma^2(n-m)}\left[x^n\int_x^{+\infty} s^{-n-1}\Pi(s)ds + x^m\int_0^x s^{-m-1}\Pi(s)ds\right], \quad x>0,\end{aligned}$$

où $\hat{X}^x$ est la solution de (4.19) avec $L = 0$. On vérifie (exercice laissé au lecteur) que sous la condition d'Inada (4.20) :

$$\hat{V}_0'(0^+) \;=\; +\infty \;\text{ et }\; \hat{V}_0'(+\infty) \;=\; 0\,.$$

Lemme 4.5.5 *La frontière x_b satisfait :*

$$0 < \; x_b \; < +\infty.$$

Preuve. On vérifie d'abord que $x_b > 0$. Sinon, la région $\mathcal{B}$ est vide et on aurait d'après le lemme 4.5.4 et le théorème 4.5.2

$$\beta v - \mathcal{L}v - \Pi = 0, \quad x > 0.$$

Alors, v devrait être de la forme :

$$v(x) = Ax^m + Bx^n + \hat{V}_0(x), \quad x > 0.$$

Comme $m < 0$ et $|v(0^+)| < \infty$, ceci implique $A = 0$. De plus, comme $n > 1$, on a $v'(0^+) = \hat{V}_0'(0^+) = +\infty$, ce qui est en contradiction avec le fait que $v'(x) \leq \lambda$ pour tout $x > 0$.

On a aussi $x_b < +\infty$. Sinon, v serait solution de $v' = \lambda$ sur $]0,+\infty[$ ce qui est contradiction avec la condition de croissance (4.24). □

Théorème 4.5.3 *La fonction valeur v a la forme explicite suivante :*

$$v(x) = \begin{cases} \lambda x + v(0^+) & x \leq x_b \\ Ax^m + \hat{V}_0(x) & x_b < x, \end{cases} \tag{4.36}$$

où les trois constantes $v(0^+)$, A and x_b sont déterminées par les conditions de continuité C^0, différentiabilité C^1 et C^2 de v en x_b :

$$Ax_b^m + \hat{V}_0(x_b) = \lambda x_b + v(0^+). \tag{4.37}$$

$$mAx_b^{m-1} + \hat{V}_0'(x_b) = \lambda, \tag{4.38}$$

$$m(m-1)Ax_b^{m-2} + \hat{V}_0''(x_b) = 0. \tag{4.39}$$

Preuve. On sait déjà d'après le lemme 4.5.4 que sur $]0, x_b[$, qui est non vide d'après le lemme 4.5.5, v a la structure décrite en (4.36). De plus, sur $]x_b, +\infty[$, on a $v' < \lambda$ d'après le lemme 4.5.4. Ainsi, avec le théorème 4.5.2, on en déduit que v satisfait $\beta v - \mathcal{L}v - \Pi = 0$, et donc est de la forme :

$$v(x) = Ax^m + Bx^n + \hat{V}_0(x), \quad x > x_b.$$

Puisque $m < 0$, $n > 1$, $\hat{V}_0'(x)$ tend vers 0 quand x tend vers l'infini et $0 \leq v'(x) \leq \lambda$, on doit avoir nécessairement $B = 0$. Ainsi v a la forme écrite en (4.36). Finalement, les trois conditions résultant de la continuité et différentiabilité C^1 et C^2 de v en x_b déterminent les constantes A, x_b et $v(0^+)$. □

4.5.3 Stratégie optimale

On rappelle le lemme de Skorohod suivant prouvé par exemple dans P.L. Lions et Snitzman [LS84].

Lemme 4.5.6 *Pour tout état initial $x \geq 0$ et toute frontière $x_b \geq 0$, il existe un unique processus adapté X^* et un processus croissant L^* continus à droite, satisfaisant le problème de Skorohod $\mathcal{S}(x, x_b)$:*

$$dX_t^* = X_t^* \left(-\delta dt + \sigma dW_t\right) + dL_t^*, \ t \geq 0, \qquad X_{0^-}^* = x, \tag{4.40}$$

$$X_t^* \in [x_b, +\infty[\quad p.s., \quad t \geq 0, \tag{4.41}$$

$$\int_0^{+\infty} 1_{X_u^* > x_b} dL_u^* = 0. \tag{4.42}$$

De plus, si $x \geq x_b$, alors L^* est continu. Lorsque $x < x_b$, $L_0^* = x_b - x$, et X_0^* $= x_b$.

La solution X^* des équations ci-dessus est une diffusion réfléchie sur la frontière x_b et le processus L^* est le temps local de X^* en x_b. La condition (4.42) signifie que L^* ne croit que lorsque X^* touche la frontière x_b. Le β-potentiel de L^* est fini, i.e. $E[\int_0^{+\infty} e^{-\beta t} dL_t^*]$, voir Chapitre X dans Revuz-Yor [ReY91], ce qui implique :

$$E\left[\int_0^{+\infty} e^{-\beta t} X_t^* dt\right] < +\infty. \tag{4.43}$$

Théorème 4.5.4 *Pour $x \geq 0$, soit (X^*, L^*) la solution du problème de Skorohod $\mathcal{S}(x, x_b)$. Alors on a :*

$$v(x) = E\left[\int_0^{+\infty} e^{-\beta t} \left(\Pi(X_t^*) dt - \lambda dL_t^*\right)\right].$$

Preuve. 1) On considère d'abord le cas où $x \geq x_b$. Alors les processus X^* et L^* sont continus (on a omis la dépendance en x dans X^*). D'après (4.41) et le théorème 4.5.2, on a

$$\beta v(X_t^*) - \mathcal{L} v(X_t^*) - \Pi(X_t^*) = 0, \ p.s. \quad t \geq 0.$$

En appliquant la formule d'Itô à $e^{-\beta t} v(X_t^*)$ entre 0 et T, on obtient :

$$E\left[e^{-\beta T} v(X_T^*)\right] = v(x) - \ E\left[\int_0^T e^{-\beta t} \Pi(X_t^*) dt\right] + \ E\left[\int_0^T e^{-\beta t} v'(X_t^*) dL_t^*\right]. \tag{4.44}$$

(L'intégrale stochastique apparaissant dans la formule d'Itô s'annule en espérance d'après (4.43)). Alors d'après (4.42), on a

$$E\left[\int_0^T e^{-\beta t} v'(X_t^*) dL_t^*\right] = E\left[\int_0^T e^{-\beta t} v'(X_t^*) 1_{X_t^* = x_b} dL_t^*\right]$$
$$= E\left[\int_0^T e^{-\beta t} \lambda dL_t^*\right],$$

car $v'(x_b) = \lambda$. En substituant dans (4.44), on obtient :

$$v(x) = E\left[e^{-\beta T} v(X_T^*)\right] + E\left[\int_0^T e^{-\beta t} \Pi(X_t^*) dt\right] - E\left[\int_0^T e^{-\beta t} \lambda dL_t^*\right]. \tag{4.45}$$

D'après (4.43), on a $\lim_{T\to+\infty} E[e^{-\beta T} X_T^*] = 0$ et donc par la croissance linéaire de v :

$$\lim_{T\to+\infty} E[e^{-\beta T} v(X_T^*)] = 0.$$

En envoyant alors T vers l'infini dans (4.45), on a l'égalité voulue.

2) Lorsque $x < x_b$ et puisque $L_0^* = x - x_b$, on a :

$$E\left[\int_0^{+\infty} e^{-\beta t} (\Pi(X_t^x) - \lambda dL_t^*)\right] = E\left[\int_0^{+\infty} e^{-\beta t} \left(\Pi(X_t^{x_b}) - \lambda dL_t^*\right)\right] - \lambda(x - x_b)$$
$$= v(x_b) - \lambda(x - x_b) = v(x),$$

en se rappelant que $v' = \lambda$ sur $]0, x_b[$. □

En conclusion, les théorèmes 4.5.3 et 4.5.4 donnent une solution explicite à ce modèle d'investissement irréversible. Ils valident l'intuition économique qu'une firme investit dans l'augmentation de capital de telle sorte de maintenir sa capacité de production au dessus d'un certain seuil x_b.

4.6 Calcul du coût de surréplication en volatilité incertaine

On considère la diffusion contrôlée :

$$dX_s = \alpha_s X_s dW_s, \quad t \le s \le T,$$

où le processus de contrôle α est progressif, à valeurs dans $A = [\underline{a}, \bar{a}]$, $0 \le \underline{a} \le \bar{a} \le +\infty$, et $E[\int_t^T |\alpha_s|^2 ds] < +\infty$. Pour éviter les cas triviaux dégénérés, on suppose $\bar{a} > 0$ et $\underline{a} \neq +\infty$. On note par $\mathcal{A}_t$ l'ensemble de tels processus de contrôles. Comme d'habitude, on se ramène au cas des hypothèses générales en faisant le changement de variable $x \in \mathbb{R}_+^* \mapsto \ln x \in \mathbb{R}$. Notons que lorsque $x = 0$, alors $X_s^{t,x} = 0$ pour tout $s \geq t$. En finance, α représente la volatilité incertaine du prix de l'action X. Etant donnée une fonction g s.c.i. à croissance linéaire, représentant le payoff d'une option, on cherche à calculer son coût de surréplication donné par :

$$v(t,x) = \sup_{\alpha \in \mathcal{A}_t} E\left[g(X_T^{t,x})\right], \quad (t,x) \in [0,T] \times \mathbb{R}_+^*. \tag{4.46}$$

Notons que le processus positif $\{X_s^{t,x}, t \leq s \leq T\}$ est une surmartingale pour tout $\alpha \in \mathcal{A}_t$. On en déduit qu'il existe C constante positive telle que :

$$|v(t,x)| \leq C(1+x), \quad \forall (t,x) \in [0,T] \times \mathbb{R}_+^*.$$

Donc v est aussi à croissance linéaire et est en particulier localement bornée sur $[0,T] \times \mathbb{R}_+^*$.

L'Hamiltonien du problème de maximisation (4.46) est donné par :

$$H(x,M) = \inf_{a \in [\underline{a},\bar{a}]} \left\{-\frac{1}{2}a^2x^2M\right\}, \quad (x,M) \in \mathbb{R}_+^* \times \mathbb{R}.$$

On va distinguer deux cas selon que la volatilité maximale est bornée ou non.

4.6.1 Volatilité bornée

Dans ce paragraphe, on suppose

$$\bar{a} < +\infty.$$

Dans ce cas, l'Hamiltonien est fini et continu sur tout $(x,M) \in \mathbb{R}_+^* \times \mathbb{R}$, et est donné par :

$$H(x,M) = -\frac{1}{2}\hat{a}^2(M)x^2M,$$

avec

$$\hat{a}(M) = \begin{cases} \bar{a} & \text{si } M \geq 0 \\ \underline{a} & \text{si } M < 0. \end{cases}$$

D'après le théorème 4.4.1 et en utilisant les résultats de comparaison, on obtient la caractérisation suivante :

Théorème 4.6.5 *Supposons $\bar{a} < +\infty$ et g continue. Alors v est l'unique solution de viscosité à croissance linéaire, continue sur $[0,T] \times \mathbb{R}_+^*$, de l'équation dite de Black-Scholes-Barenblatt :*

$$-\frac{\partial v}{\partial t} - \frac{1}{2}\hat{a}^2\left(\frac{\partial^2 v}{\partial x^2}\right)x^2\frac{\partial^2 v}{\partial x^2} = 0, \quad (t,x) \in [0,T[\times \mathbb{R}_+^*, \tag{4.47}$$

satisfaisant la condition terminale :

$$v(T,x) = g(x), \quad x \in \mathbb{R}_+^*. \tag{4.48}$$

Preuve. On sait d'après le théorème 4.4.1 que v est solution de viscosité de (4.47). Pour caractériser v, il faut déterminer $v_*(T,.)$ et $v^*(T,.)$. Soit $x \in \mathbb{R}_+^*$ quelconque. Soit (t_n, x_n) une suite dans $[0,T[\times\mathbb{R}_+^*$ telle que :

$$(t_n, x_n) \ \rightarrow \ (T, x) \text{ et } v(t_n, x_n) \ \rightarrow v_*(T, x),$$

quand n tend vers l'infini. Soit $a \in [\underline{a}, \bar{a}]$ et $X_s^{t_n,x_n}$ la diffusion contrôlée associée au contrôle constant a. On a donc $v(t_n, x_n) \geq E[g(X_T^{t_n,x_n})]$. Dans la suite C désigne une constante positive générique indépendante de n. Comme g est à croissance linéaire, on a :

$$\begin{aligned} E\left|g(X_T^{t_n,x_n})\right|^2 &\leq C(1 + E|X_T^{t_n,x_n}|^2) \\ &\leq C(1 + x^2 e^{a^2 T}). \end{aligned}$$

Ainsi la suite $\{g(X_T^{t_n,x_n}), n \geq 1\}$ est borné dans L^2, donc est uniformément intégrable. On en déduit par le théorème de convergence dominée, la continuité de g et p.s. de X_T en ses données initiales (t,x) que

$$v_*(T, x) \geq g(x). \tag{4.49}$$

D'autre part, considérons une suite (t_n, x_n) dans $[0,T[\times\mathbb{R}_+^*$ telle que :

$$(t_n, x_n) \ \rightarrow \ (T, x) \text{ et } v(t_n, x_n) \ \rightarrow v^*(T, x),$$

quand n tend vers l'infini. Pour tout $\varepsilon > 0$, il existe un contrôle $\hat{\alpha}^n \in \mathcal{A}_{t_n}$ tel que :

$$v(t_n, x_n) \leq E[g(X_T^{t_n,x_n})] + \varepsilon. \tag{4.50}$$

Ici $X_s^{t_n,x_n}$ est la diffusion contrôlée associée au contrôle $\hat{\alpha}^n$. Comme A est compact borné par $\bar{a}$, on a :

$$\begin{aligned} E\left|g(X_T^{t_n,x_n})\right|^2 &\leq C(1 + E|X_T^{t_n,x_n}|^2) \\ &\leq C(1 + x^2 e^{\bar{a}^2 T}). \end{aligned}$$

On obtient donc comme précédemment par le théorème de convergence dominée, en faisant tendre n vers l'infini, puis ε vers zéro dans (4.50), et avec (4.49) :

$$v_*(T, x) \ = \ v^*(T, x) = g(x).$$

On conclut avec le principe de comparaison fort, en notant aussi qu'on a la continuité jusqu'en $t = T$ puisque $v(T, x) = g(x)$. □

Remarque 4.6.7 Lorsque $\underline{a} > 0$, et g est assez régulière, il y a existence et unicité d'une solution régulière à l'EDP de Black-Scholes-Barenblatt (4.47) avec la condition de Cauchy (4.48). C'est essentiellement une conséquence de la condition d'ellipticité uniforme sur le terme de diffusion qui est obtenue après le changement de variable $x \rightarrow \ln x$.

Remarque 4.6.8 Si g est convexe, alors la fonction :

$$w(t,x) = E\left[g(\hat{X}_T^{t,x})\right], \quad (t,x) \in [0,T[\times\mathbb{R}_+^*,$$

où $\{\hat{X}_s^{t,x}, t \leq s \leq T\}$ est la solution de l'EDS :

$$d\hat{X}_s = \bar{a}\hat{X}_s dW_s, \;\; t \leq s \leq T, \quad \hat{X}_t \; = \; x,$$

est aussi convexe en x. De plus, w est solution régulière de l'équation de Black-Scholes :

$$-\frac{\partial w}{\partial t} - \frac{1}{2}\bar{a}^2 x^2 \frac{\partial^2 w}{\partial x^2} = 0, \quad (t,x) \; \in [0,T[\times\mathbb{R}_+^*,$$

et satisfait la condition terminale $w(T,x) = g(x)$. Comme $\hat{a}\left(\dfrac{\partial^2 w}{\partial x^2}\right) = \bar{a}$, ceci implique que w est en fait aussi solution de l'équation (4.47), et donc par unicité que $w = v$. On a une remarque analogue lorsque g est concave.

4.6.2 Volatilité non bornée

Dans ce paragraphe, on suppose

$$\bar{a} = +\infty.$$

Dans ce cas l'Hamiltonien est donné par :

$$H(x,M) = \begin{cases} -\frac{1}{2}\underline{a}^2 x^2 M & \text{si } \; -M \geq 0 \\ \quad -\infty & \text{si } \; -M < 0. \end{cases}$$

D'après le théorème 4.4.1, la fonction v est solution de viscosité de :

$$\min\left\{-\frac{\partial v}{\partial t} - \frac{1}{2}\underline{a}^2 x^2 \frac{\partial^2 v}{\partial x^2} \; , \; -\frac{\partial^2 v}{\partial x^2}\right\} = 0, \quad (t,x) \in [0,T[\times\mathbb{R}_+^*. \quad (4.51)$$

Ceci prouve en particulier que v est une sursolution de viscosité de $-\dfrac{\partial^2 v}{\partial x^2}$ ≥ 0 sur $[0,T[\times\mathbb{R}_+^*$. Le lemme général suivant montre alors que pour tout $t \in [0,T[$, $v(t,.)$ est sursolution de viscosité de $-\dfrac{\partial^2 v}{\partial x^2}(t,.) \geq 0$ sur $\mathbb{R}_+^*$. C'est une propriété triviale dans le cas de solutions régulières mais plus délicate à prouver avec la notion plus faible de solution de viscosité.

Lemme 4.6.7 *Soient $\mathcal{O}_1 \subset \mathbb{R}^n$ et $\mathcal{O}_2 \subset \mathbb{R}^d$ deux ouverts. Soit $v : \mathcal{O}_1 \times \mathcal{O}_2$ $\to \mathbb{R}$ une fonction localement bornée, sursolution s.c.i. de viscosité de :*

$$F(x,y,D_y v(x,y), D_y^2 v(x,y)) \geq 0, \quad \textit{sur } \mathcal{O}_1 \times \mathcal{O}_2,$$

où $F : \mathcal{O}_1 \times \mathcal{O}_2 \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R}$ est continue et vérifie la condition d'ellipticité (4.2)*. Alors pour tout $x_0 \in \mathcal{O}_1$, la fonction $v(x_0,.)$ est une sursolution s.c.i. de viscosité de :*

$$F(x_0,y,D_y v(x_0,y), D_y^2 v(x_0,y)) \geq 0, \quad \textit{sur } \mathcal{O}_2.$$

Preuve. Soit $x_0 \in \mathcal{O}_1$ fixé, $u(y) := v(x_0, y)$ et $\varphi \in C^2(\mathcal{O}_2)$, $y_0 \in \mathcal{O}_2$ tel que :

$$0 = (u - \varphi)(y_0) < (u - \varphi)(y), \quad \forall y \in \mathcal{O}_2 \setminus \{y_0\}, \tag{4.52}$$

i.e. y_0 est un minimum strict de $u - \varphi = u_* - \varphi$ (car v est s.c.i). On doit montrer que :

$$F(x_0, y_0, D\varphi(y_0), D^2\varphi(y_0)) \geq 0. \tag{4.53}$$

Pour cela, considérons la suite de fonctions dans $C^2(\mathcal{O}_1 \times \mathcal{O}_2)$:

$$\psi_n(x, y) = \varphi(y) - n|x - x_0|^2.$$

Soit $I_1 \times I_2$ un voisinage compact de (x_0, y_0) dans $\mathcal{O}_1 \times \mathcal{O}_2$. La fonction $v - \psi_n$ étant s.c.i, il existe $(x_n, y_n) \in I_1 \times I_2$ tel que :

$$(v - \psi_n)(x_n, y_n) = \min_{I_1 \times I_2} (v - \psi_n).$$

Par compacité de $I_1 \times I_2$, la suite (x_n, y_n) converge (à une sous-suite près) vers un élément $(\bar{x}, \bar{y}) \in I_1 \times I_2$. D'après la définition de ψ_n et de (x_n, y_n), on a :

$$\begin{aligned} v(x_n, y_n) - \varphi(y_n) &\leq (v - \psi_n)(x_n, y_n) \\ &\leq (v - \psi_n)(x_0, y_0) \;=\; u(y_0) - \varphi(y_0). \end{aligned}$$

En faisant tendre n vers l'infini et puisque v est s.c.i, on obtient :

$$\begin{aligned} v(\bar{x}, \bar{y}) - \varphi(\bar{y}) &\leq v(\bar{x}, \bar{y}) - \varphi(\bar{y}) + \liminf_{n \to +\infty} n|x_n - x_0|^2 \\ &\leq u(y_0) - \varphi(y_0). \end{aligned}$$

Ceci implique que $\bar{x} = x_0$ et ensuite que :

$$(u - \varphi)(\bar{y}) \;=\; v(x_0, \bar{y}) - \varphi(\bar{y}) \leq (u - \varphi)(y_0).$$

D'après (4.52), ceci prouve que $\bar{y} = y_0$ et donc finalement que

$$(x_n, y_n) \to (x_0, y_0),$$

quand n tend vers l'infini. De plus, par la propriété de sursolution de viscosité de v, on a :

$$\begin{aligned} 0 &\leq F(x_n, y_n, D_y\psi_n(x_n, y_n), D_y^2\psi_n(x_n, y_n)) \\ &= F(x_n, y_n, D\varphi(y_n), D^2\varphi(y_n)), \end{aligned}$$

et on obtient le résultat voulu (4.53) en faisant tendre n vers l'infini. □

Dans le cas où v est régulière, on aura aussi que $v(t, .)$ est concave sur $\mathbb{R}_+^*$ pour tout $t \in [0, T[$. Le lemme suivant montre que c'est encore vraie avec la notion de viscosité.

Lemme 4.6.8 *Soit $\mathcal{O}$ un intervalle ouvert de $\mathbb{R}$ et $w : \mathcal{O} \to \mathbb{R}$ une fonction localement bornée. Alors w est une sursolution s.c.i. de viscosité de $-D^2 w = 0$ sur $\mathcal{O}$ si et seulement si w est concave sur $\mathcal{O}$.*

Preuve.
1. Supposons d'abord que w est une sursolution s.c.i. de viscosité de $-D^2 w = 0$ sur $\mathcal{O}$. Soit $x_0 < x_1$ deux points quelconques de $\mathcal{O}$. Puisque w est localement bornée, w est bornée inférieurement sur $[x_0, x_1]$, et on peut donc supposer, quitte à ajouter une constante à w (ce qui ne changera pas la propriété de sursolution de w), que $w \geq 0$ sur $[x_0, x_1]$. On voit alors aisément que w est une sursolution de viscosité de $\varepsilon^2 w - D^2 w = 0$ sur $]x_0, x_1[$ pour tout $\varepsilon > 0$. Considérons l'équation :

$$\varepsilon u - D^2 u = 0 \quad \text{sur }]x_0, x_1[,$$

avec la condition de bord

$$\begin{aligned} u(x_0) &= w(x_0) (= w_*(x_0)) \\ u(x_1) &= w(x_1) (= w_*(x_1)). \end{aligned}$$

La solution de ce problème est régulière sur $[x_0, x_1]$ et est clairement donnée par :

$$u_\varepsilon(x) = \frac{w(x_0) \left[e^{\varepsilon(x_1 - x)} - 1 \right] + w(x_1) \left[e^{\varepsilon(x - x_0)} - 1 \right]}{e^{\varepsilon(x_1 - x_0)} - 1}.$$

Par le principe de comparaison, on a $w(x) \geq u_\varepsilon(x)$ pour tout $x \in [x_0, x_1]$. En faisant tendre ε vers zéro, on en déduit que :

$$\frac{w(x) - w(x_0)}{x - x_0} \geq \frac{w(x_1) - w(x_0)}{x_1 - x_0}, \quad \forall x \in]x_0, x_1[,$$

ce qui prouve la concavité de w.

2. Supposons réciproquement que w est une fonction concave sur $\mathcal{O}$. Notons en particulier que w est continue sur $\mathcal{O}$. Soit $\bar{x} \in \mathcal{O}$ et $\varphi \in C^2(\mathcal{O})$ tel que

$$0 \; = \; (w - \varphi)(\bar{x}) = \min_{\mathcal{O}} (w - \varphi).$$

Alors pour tout $h > 0$ assez petit tel que $\bar{x} - h$ et $\bar{x} + h \in \mathcal{O}$, on a :

$$\begin{aligned} -\frac{\varphi(\bar{x} + h) + \varphi(\bar{x} - h) - 2\varphi(\bar{x})}{h^2} &\geq -\frac{w(\bar{x} + h) + w(\bar{x} - h) - 2w(\bar{x})}{h^2} \\ &\geq 0. \end{aligned}$$

En faisant tendre h vers zéro, on obtient $-D^2 \varphi(\bar{x}) \geq 0$ qui est la propriété de sursolution requise. □

On note par g^{conc} l'envelope concave de g, i.e. la plus petite fonction concave majorant g. Notons que puisque g est à croissance linéaire, il en est de même de g^{conc}. On a alors la caractérisation explicite suivante de la fonction v.

Théorème 4.6.6 *Supposons $\bar{a} = +\infty$ et g borné inférieurement. Alors $v = w$ sur $[0, T[\times\mathbb{R}_+^*$ où*

$$w(t, x) = E\left[g^{conc}(\hat{X}_T^{t,x})\right], \quad \forall (t, x) \in [0, T] \times \mathbb{R}_+^*, \tag{4.54}$$

et $\{\hat{X}_s^{t,x}, t \leq s \leq T\}$ est la solution de l'EDS :

$$d\hat{X}_s = \underline{a}\hat{X}_s dW_s, \;\; t \leq s \leq T, \quad \hat{X}_t \; = \; x.$$

Preuve.
1. Montrons d'abord que v est s.c.i. sur $[0, T] \times \mathbb{R}_+^*$. Soit (t_n, x_n) une suite de $[0, T[\times\mathbb{R}_+^*$ convergeant vers $(t, x) \in [0, T] \times \mathbb{R}_+^*$. Comme g est s.c.i. et borné inférieurement, et par la continuité p.s. de $X_T^{t,x}$ en ses données initiales (t, x) pour tout contrôle $\alpha \in \mathcal{A}_t$, on a par le lemme de Fatou :

$$\begin{aligned}\liminf_{n\to+\infty} v(t_n, x_n) &\geq \liminf_{n\to+\infty} E[g(X_T^{t_n,x_n})] \\ &\geq E[g(X_T^{t,x})],\end{aligned}$$

et donc $\liminf_{n\to+\infty} v(t_n, x_n) \geq v(t, x)$.
2. D'après (4.51) et le lemme 4.6.7, pour tout $t \in [0, T[$, $v(t, .)$ est une sur-solution de viscosité de $-\dfrac{\partial^2 v}{\partial x^2}(t, .) \geq 0$ sur $\mathbb{R}_+^*$. D'après le lemme 4.6.8, ceci signifie que $v(t, .)$ est concave sur $[0, T[$ pour tout $t \in \mathbb{R}_+^*$. Autrement dit, pour tout $t \in [0, T[$, $x, y \in \mathbb{R}_+^*$,

$$v\left(t, \frac{x+y}{2}\right) \geq \frac{1}{2}v(t, x) + \frac{1}{2}v(t, y).$$

En passant à la liminf quand t tend vers T^-, on obtient :

$$v_*\left(T, \frac{x+y}{2}\right) \geq \frac{1}{2}v_*(T, x) + \frac{1}{2}v_*(T, y),$$

ce qui signifie que $v_*(T, .)$ est concave sur $\mathbb{R}_+^*$. De plus d'après la s.c.i. de v sur $[0, T] \times \mathbb{R}_+^*$, on a $v_*(T, x) \geq v(T, x) = g(x)$. On en déduit que :

$$v_*(T, x) \geq g^{conc}(x), \quad \forall x \in \mathbb{R}_+^*. \tag{4.55}$$

3. Soit $(t, x) \in [0, T] \times \mathbb{R}_+^*$. On note $\mathcal{A}_{t,b}$ le sous ensemble de $\mathcal{A}_t$ consitué des contrôles bornés p.s. sur $[t, T]$. On a donc $v(t, x) \geq \sup_{\alpha\in\mathcal{A}_{t,b}} E[g(X_T^{t,x})]$. Réciproquement, étant donné $\hat{\alpha} \in \mathcal{A}_t$ quelconque, on pose pour $n \in \mathbb{N}$, $\alpha^n := \hat{\alpha}1_{|\hat{\alpha}|\leq n} \in \mathcal{A}_{t,b}$. On note par $\hat{X}_s^{t,x}$ (resp. X_s^n), $t \leq s \leq T$, le processus de diffusion contrôlé par $\hat{\alpha}$ (resp. α^n). i.e. :

$$\begin{aligned}\hat{X}_s^{t,x} &= x \exp\left(\int_t^s \hat{\alpha}_u dW_u - \frac{1}{2}\int_t^s |\hat{\alpha}_u|^2 du\right), \\ X_s^n &= x \exp\left(\int_t^s \alpha_u^n dW_u - \frac{1}{2}\int_t^s |\alpha_u^n|^2 du\right).\end{aligned}$$

Alors X_T^n converge p.s. vers $\hat{X}_T^{t,x}$ quand n tend vers l'infini, et par le lemme de Fatou, on obtient :

$$\begin{aligned}\sup_{\alpha\in\mathcal{A}_{t,b}} E[g(X_T^{t,x})] &\geq \liminf_{n\to+\infty} E[g(X_T^n)] \\ &\geq E[g(\hat{X}_T^{t,x})].\end{aligned}$$

Puisque $\hat{\alpha}\in\mathcal{A}_t$ est arbitraire, on a l'inégalité réciproque voulue d'où :

$$v(t,x) = \sup_{\alpha\in\mathcal{A}_{t,b}} E[g(X_T^{t,x})].$$

En remarquant que pour tout $\alpha\in\mathcal{A}_{t,b}$, le processus $X_s^{t,x}$, $t\leq s\leq T$, est une martingale, on a par l'inégalité de Jensen :

$$\begin{aligned}v(t,x) &\leq \sup_{\alpha\in\mathcal{A}_{t,b}} E[g^{conc}(X_T^{t,x})] \\ &\leq \sup_{\alpha\in\mathcal{A}_{t,b}} g^{conc}\left(E[X_T^{t,x}]\right) \;=\; g^{conc}(x).\end{aligned} \tag{4.56}$$

Ceci implique en particulier que $v^*(T,x)\leq g^{conc}(x)$, et montre avec la relation (4.55) que la condition terminale pour le problème de contrôle stochastique (4.46) est :

$$v^*(T,x) \;=\; v_*(T,x) = g^{conc}(x), \quad x\in\mathbb{R}_+^*. \tag{4.57}$$

4. La fonction g^{conc} étant continue et à croissance linéaire, on montre par les mêmes arguments que dans la preuve du théorème 4.6.5, que w donné par (4.54) est continue en T avec $w^*(T,x) = w_*(T,x) = g^{conc}(x)$. Comme cas particulier du théorème 4.4.1 et par un résultat d'unicité fort, on en déduit que w est l'unique solution de viscosité à croissance linéaire, continue sur $[0,T]\times\mathbb{R}_+^*$ de l'équation de Black-Scholes :

$$-\frac{\partial w}{\partial t} - \frac{1}{2}\underline{a}^2 x^2\frac{\partial^2 w}{\partial x^2} = 0, \quad (t,x)\;\in[0,T[\times\mathbb{R}_+^*, \tag{4.58}$$

avec la condition terminale :

$$w^*(T,x) \;=\; w_*(T,x) = g^{conc}(x), \quad x\in\mathbb{R}_+^*. \tag{4.59}$$

D'après (4.51), v est sursolution s.c.i. de viscosité de (4.58). Comme v et w vérifient la même condition terminale, on en déduit par le principe de comparaison fort que :

$$v\geq w \quad \text{sur } [0,T[\times\mathbb{R}_+^*. \tag{4.60}$$

5. La fonction g^{conc} étant concave, il en est de même pour la fonction $w(t,.)$. D'après le lemme 4.6.8, ceci implique que $w(t,.)$ est une sursolution de viscosité de $-\dfrac{\partial^2 w}{\partial x^2}(t,.)\geq 0$ sur $\mathbb{R}_+^*$ pour tout $t\in[0,T[$. A fortiori, on en déduit que w est sursolution de viscosité de

$$-\frac{\partial^2 w}{\partial x^2}(t,x) \geq 0, \quad \text{sur } [0,T[\times\mathbb{R}_+^*.$$

En combinant avec (4.58), ceci prouve que w est une solution de viscosité de la même équation que v, i.e. :

$$\min\left\{-\frac{\partial w}{\partial t} - \frac{1}{2}\underline{a}^2 x^2 \frac{\partial^2 w}{\partial x^2}, -\frac{\partial^2 w}{\partial x^2}\right\} = 0, \quad (t,x) \in [0,T[\times\mathbb{R}_+^*. \quad (4.61)$$

De plus comme v et w satisfont la même condition terminale, on peut conclure que $v = w$ sur $[0,T[\times\mathbb{R}_+^*$ par un résultat d'unicité fort sur l'inéquation variationnelle (4.61). On donne ici un argument qui n'utilise pas directement un tel résultat d'unicité fort. Lorsque $\underline{a} = 0$, on a $w = g^{conc} \geq v$ d'après (4.56) et donc l'égalité voulue avec (4.60). Lorsque $\underline{a} > 0$, la fonction w est en fait régulière $C^{1,2}([0,T[\times\mathbb{R}_+^*) \cap C^0([0,T]\times\mathbb{R}_+^*)$. En appliquant alors la formule d'Itô à $w(s, X_s^{t,x})$ pour tout $\alpha \in \mathcal{A}_t$ (après avoir éventuellement localisé pour annuler en espérance le terme d'intégrale stochastique), on obtient :

$$\begin{aligned}
E\left[g^{conc}(X_T^{t,x})\right] &= w(t,x) + E \\
&\times \left[\int_t^T \frac{\partial w}{\partial t}(s, X_s^{t,x}) + (\alpha_s)^2 (X_s^{t,x})^2 \frac{\partial^2 w}{\partial x^2}(s,, X_s^{t,x})\, ds\right].
\end{aligned}$$

Comme w est concave et $\alpha_s \geq \underline{a}$, on en déduit que :

$$\begin{aligned}
E\left[g(X_T^{t,x})\right] &\leq E\left[g^{conc}(X_T^{t,x})\right] \\
&\leq w(t,x) + E\left[\int_t^T \frac{\partial w}{\partial t}(s, X_s^{t,x})\right. \\
&\qquad \left. + \underline{a}^2 (X_s^{t,x})^2 \frac{\partial^2 w}{\partial x^2}(s,, X_s^{t,x})\, ds\right] = w(t,x),
\end{aligned}$$

puisque w est solution de (4.58). Ceci étant valide pour tout $\alpha \in \mathcal{A}_t$, on conclut que $v \leq w$ et finalement :

$$v = w \quad \text{sur } [0,T[\times\mathbb{R}_+^*.$$

□

Remarque 4.6.9 Le théorème précédent montre donc que le coût de surréplication dans un modèle à volatilité non bornée est égal au prix de Black-Scholes d'un payoff "concavifié" et pour la volatilité inférieure $\underline{a}$. Il n'y a pas de contrôle optimal lorsque $\bar{a} = +\infty$ et l'effet de la volatilité maximale non bornée est de concavifier la fonction valeur de ce problème de contrôle stochastique singulier.

Remarque 4.6.10 Le théorème précédent montre que lorsque $\underline{a} = 0$ alors $v(t,.) = g^{conc}$ pour tout $t \in [0,T[$. Ceci est aussi vraie pour $\underline{a} \geq 0$ lorsqu'on suppose de plus que g est convexe. En effet, on sait déjà d'après (4.56) que $v \leq g^{conc}$. De plus, par l'inégalité de Jensen, si g est convexe, on a d'après (4.46) que $v(t,x) \geq g(x)$ pour tout $(t,x) \in [0,T] \times \mathbb{R}_+^*$. Par concavité de v, on en déduit que $v(t,x) \geq g^{conc}(x)$ pour tout $(t,x) \in [0,T[\times\mathbb{R}_+^*$.

4.7 Commentaires bibliographiques

Les solutions de viscosité pour les équations du second ordre de type HJB liés à des problèmes de contrôle stochastique ont été introduites par P.L. Lions [Lio83]. Les résultats de comparaison et d'unicité pour les solutions de viscosité ont fait l'objet d'importants travaux, parmi lesquels Jensen [Je88], Ishii [Ish89], voir aussi l'article de synthèse de Crandall, Ishii et P.L. Lions [CIL92].

L'application au modèle d'investissement irréversible est un cas limite d'un modèle d'investissement réversible étudié par Guo et Pham [GP05]. D'autres exemples d'utilisation des solutions de viscosité pour des modèles de contrôle singulier en finance sont traités dans Shreve et Soner [ShSo94] pour des modèles avec coûts de transaction proportionnels ou encore dans Choulli, Taksar et Zhou [CTZ03] pour des modèles de versement de dividende d'une firme. L'application au calcul du coût de surréplication en volatilité incertaine est inspirée de Cvitanic, Pham et Touzi [CPT99]. Une extension au cadre multidimensionnel pour les actifs risqués est étudiée dans Gozzi et Vargiolu [GV02].

5

Méthodes d'équations différentielles stochastiques rétrogrades

5.1 Introduction

Depuis l'article de référence de Pardoux et Peng [PaPe90], la théorie des équations différentielles stochastiques rétrogrades (EDSR) a connu un grand développement ces dernières années grâce notamment à ses diverses applications en mathématiques. Formellement, les EDSR sont des équations différentielles stochastiques où l'on se donne la condition terminale. Dans le cas déterministe, la donnée d'une condition terminale ou d'une condition terminale pour une équation différentielle (ordinaire donc) est équivalente par inversion du temps. Dans le cas stochastique, les choses sont fondamentalement différentes lorsqu'on cherche des solutions qui restent adaptées par rapport à une filtration donnée. En effet, en inversant simplement le temps, on perd la propriété de non anticipation de la solution. Le premier point est donc de formuler correctement la notion de solution adaptée à une EDSR.

Ce chapitre est une introduction basique à cette théorie des EDSR en mettant l'accent sur ses applications au contrôle stochastique et mathématiques financières. On présentera dans la section suivante 2 quelques résultats généraux sur les EDSR : existence et unicité de solution et principe de comparaison. La section 3 montrera que certaines EDP non linéaires peuvent se représenter via les solutions d'EDSR, généralisant ainsi les formules de Feynman-Kac. Dans la section 4, on décrit comment les EDSR apparaissent comme un outil puissant pour étudier des problèmes de contrôle stochastique. Enfin, la section 5 illustrera ces résultats sur deux exemples d'applications issus de la finance.

5.2 Propriétés générales

5.2.1 Résultats d'existence et d'unicité

Soit $W = (W_t)_{0\leq t\leq T}$ un mouvement Brownien standard d-dimensionnel défini sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F}, P)$ où $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ est la filtration naturelle de W. On note par $\mathbb{H}^2(0,T)$ l'ensemble des processus progressifs X tel que $E[\int_0^T |X_t|^2 dt] < \infty$. Ici T est un horizon fini fixé.

On se donne un couple (ξ, f) appelé condition terminale et générateur vérifiant :

- (A) $\xi \in L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^m)$
- (B) $f : \Omega \times [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ tel que :

  - $f(., t, y, z)$ noté pour simplifier $f(t, y, z)$ est progressif pour tous y, z

  - $f(t, 0, 0) \in \mathbb{H}^2(0,T)^m$, i.e. $E\left[\int_0^T |f(t,0,0)|^2 dt\right] < +\infty$.

  - f satisfait une condition de Lipschitz uniforme : il existe une constante positive K_f telle que

$$
\begin{aligned}
&|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq K_f\left(|y_1 - y_2| + |z_1 - z_2|\right),\\
&\qquad \forall y_1, y_2,\ \forall z_1, z_2, \quad dt \times dP \ p.p.
\end{aligned}
$$

On considère l'équation différentielle stochastique rétrograde (EDSR) :

$$
-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \quad Y_T = \xi. \tag{5.1}
$$

Définition 5.2.1 *Une solution de l'EDSR* (5.1) *est un couple* (Y, Z) *de processus progressifs à valeurs dans* $\mathbb{R}^m \times \mathbb{R}^{m\times d}$ *tel que :* $Z \in \mathbb{H}^2(0,T)^{m\times d}$*, i.e.* $E\left[\int_0^T |Z_t|^2 dt\right] < +\infty$ *et*

$$
Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T.
$$

Théorème 5.2.1 *Etant donné un couple* (ξ, f) *vérifiant (A) et (B), il existe une unique solution* (Y, Z) *à l'EDSR* (5.1).

Preuve. Nous donnons ici une preuve de ce théorème basée sur une méthode de point fixe. On considère la fonction Φ sur $\mathbb{H}^2(0,T)^m \times \mathbb{H}^2(0,T)^{m\times d}$, qui à tout $(U, V) \in \mathbb{H}^2(0,T)^m \times \mathbb{H}^2(0,T)^{m\times d}$ associe $(Y, Z) = \Phi(U, V)$ défini par :

$$
Y_t = \xi + \int_t^T f(s, U_s, V_s)ds - \int_t^T Z_s dW_s. \tag{5.2}
$$

En fait, le couple (Y, Z) est construit ainsi : on considère la martingale $M_t = E[\xi + \int_0^T f(s, U_s, V_s)ds|\mathcal{F}_t]$ qui est de carré intégrable sous les hypothèses sur (ξ, f). On applique alors le théorème de représentation d'Itô qui fournit l'existence et l'unicité de $Z \in \mathbb{H}^2(0,T)^{m\times d}$ tel que :

$$
M_t = M_0 + \int_0^t Z_s dW_s. \tag{5.3}
$$

On définit alors Y par :

$$Y_t = E\left[\xi + \int_t^T f(s, U_s, V_s)ds \middle| \mathcal{F}_t\right] \;=\; M_t - \int_0^t f(s, U_s, V_s)ds, \quad 0 \leq t \leq T.$$

En utilisant la représentation (5.3) de M dans l'égalité précédente et en notant que $Y_T = \xi$, on voit que Y satisfait (5.2). Notons par l'inégalité de Doob que :

$$E\left[\sup_{0\leq t\leq T}\left|\int_t^T Z_s dW_s\right|^2\right] \leq 4E\left[\int_0^T |Z_s|^2 ds\right] \;<\; +\infty.$$

On en déduit sous les hypothèses sur (ξ, f) que

$$E\left[\sup_{0\leq t\leq T} |Y_t|^2\right] < +\infty. \tag{5.4}$$

En particulier, $Y \in \mathbb{H}^2(0,T)^m$ et Φ est donc bien une fonction de $\mathbb{H}^2(0,T)^m \times \mathbb{H}^2(0,T)^{m\times d}$ dans lui-même.

Soit (U, V), $(U', V') \in \mathbb{H}^2(0,T)^m \times \mathbb{H}^2(0,T)^{m\times d}$ et $(Y, Z) = \Phi(U, V)$, $(Y', Z') = \Phi(U', V')$. On pose $(\bar{U}, \bar{V}) = (U - U', V - V')$, $(\bar{Y}, \bar{Z}) = (Y - Y', Z - Z')$ et $\bar{f}_t = f(t, U_t, V_t) - f(t, U'_t, V'_t)$. Soit $\beta > 0$ à choisir plus tard. On applique la formule d'Itô à $e^{\beta s}|\bar{Y}_s|^2$ entre $s = 0$ et $s = T$:

$$\begin{aligned}|\bar{Y}_0|^2 = &- \int_0^T e^{\beta s}\left(\beta|\bar{Y}_s|^2 - 2\bar{Y}_s.\bar{f}_s\right) ds \\ &- \int_0^T e^{\beta s}|\bar{Z}_s|^2 ds - 2\int_0^T e^{\beta s}\bar{Y}_s'\bar{Z}_s dW_s >\end{aligned} \tag{5.5}$$

Notons que

$$E\left[\left(\int_0^T e^{2\beta t}|Y_t|^2|Z_t|^2 dt\right)^{\frac{1}{2}}\right] \leq \frac{e^{\beta T}}{2} E\left[\sup_{0\leq t\leq T}|Y_t|^2 + \int_0^T |Z_t|^2 dt\right] \;<\; +\infty,$$

ce qui montre que la martingale locale $\int_0^t e^{\beta s}\bar{Y}_s'\bar{Z}_s dW_s >$ est en fait une martingale uniformément intégrable d'après l'inégalité de Burkholder-Davis-Gundy. En prenant l'espérance dans (5.5), on a donc :

$$\begin{aligned}&E|\bar{Y}_0|^2 + E\left[\int_0^T e^{\beta s}\left(\beta|\bar{Y}_s|^2 + |\bar{Z}_s|^2\right) ds\right] \;=\; 2E\left[\int_0^T e^{\beta s}\bar{Y}_s.\bar{f}_s ds\right] \\ &\leq 2K_f E\left[\int_0^T e^{\beta s}|\bar{Y}_s|(|\bar{U}_s| + |\bar{V}_s|) ds\right] \\ &\leq 4K_f^2 E\left[\int_0^T e^{\beta s}|\bar{Y}_s|^2 ds\right] \;+\; \frac{1}{2}E\left[\int_0^T e^{\beta s}(|\bar{U}_s|^2 + |\bar{V}_s|^2) ds\right]\end{aligned}$$

En choisissant $\beta = 1 + 4K_f^2$, on obtient alors :

$$E\left[\int_0^T e^{\beta s}\left(|\bar{Y}_s|^2 + |\bar{Z}_s|^2\right) ds\right] \le \frac{1}{2} E\left[\int_0^T e^{\beta s}(|\bar{U}_s|^2 + |\bar{V}_s|^2) ds\right].$$

Ceci prouve que Φ est une contraction stricte sur l'espace de Banach $\mathbb{H}^2(0,T)^m \times \mathbb{H}^2(0,T)^{m\times d}$ muni de la norme :

$$\|(Y,Z)\|_\beta = \left(E\left[\int_0^T e^{\beta s}\left(|Y_s|^2 + |Z_s|^2\right) ds\right]\right)^{\frac{1}{2}}.$$

On conclut que Φ admet un unique point fixe qui est la solution de l'EDSR (5.1). □

5.2.2 EDSR linéaires

On s'intéresse au cas particulier où le générateur f est linéaire en y et z. Pour simplifier les notations, on considère le cas unidimensionnel $m = 1$ et on a donc une EDSR de la forme :

$$-dY_t = (A_t Y_t + Z_t B_t + C_t)\, dt - Z_t dW_t, \quad Y_T = \xi, \tag{5.6}$$

où A, B sont des processus bornés progressifs à valeurs dans $\mathbb{R}$ et $\mathbb{R}^d$ et C est un processus dans $\mathbb{H}^2(0,T)$. Dans ce cas, la solution de cette EDSR peut être explicitée.

Proposition 5.2.1 *La solution unique (Y,Z) de l'EDSR (5.6) est donnée par :*

$$\Gamma_t Y_t = E\left[\Gamma_T \xi + \int_t^T \Gamma_s C_s ds \middle| \mathcal{F}_t\right], \tag{5.7}$$

où Γ est le processus adjoint (ou dual) donné par l'EDS :

$$d\Gamma_t = \Gamma_t\left(A_t dt + B_t' dW_t\right), \quad \Gamma_0 = 1.$$

Preuve. Par la formule d'Itô à $\Gamma_t Y_t$:

$$d(\Gamma_t Y_t) = -\Gamma_t C_t dt + \Gamma_t (Z_t + Y_t B_t') dW_t,$$

soit

$$\Gamma_t Y_t + \int_0^t \Gamma_s C_s ds = Y_0 + \int_0^t \Gamma_s (Z_s + Y_s B_s') dW_t. \tag{5.8}$$

Comme A et B sont bornés, on a que $E[\sup_t |\Gamma_t|^2] < +\infty$ et en notant b_∞ la borne supérieure de B, on a :

$$
\begin{aligned}
E & \left[\left(\int_0^T \Gamma_s^2 (Z_s + Y_s B_s)^2 ds \right)^{\frac{1}{2}} \right] \\
& \leq \frac{1}{2} E \left[\sup_t |\Gamma_t|^2 + 2 \int_0^T |Z_t|^2 dt + 2 b_\infty^2 \int_0^T |Y_t|^2 dt \right] \\
& < +\infty.
\end{aligned}
$$

Par l'inégalité de Burkholder-Davis-Gundy, ceci prouve que la martingale locale dans (5.8) est une martingale uniformément intégrable. On en déduit que

$$
\begin{aligned}
\Gamma_t Y_t + \int_0^t \Gamma_s C_s ds &= E \left[\Gamma_T Y_T + \int_0^T \Gamma_s C_s ds \middle| \mathcal{F}_t \right] \\
&= E \left[\Gamma_T \xi + \int_0^T \Gamma_s C_s ds \middle| \mathcal{F}_t \right], \qquad (5.9)
\end{aligned}
$$

ce qui donne l'expression (5.7) de Y. Notons que Z est donné par la représentation d'Itô (5.8) de la martingale (5.9). □

5.2.3 Principe de comparaison

On établit dans l'énoncé suivant un principe de comparaison fort utile pour les EDSR unidimensionnels.

Théorème 5.2.2 *Soit (ξ^1, f^1) et (ξ^2, f^2) deux couples de données condition terminale-générateur satisfaisant les conditions (A) et (B), et (Y^1, Z^1), (Y^2, Z^2) les solutions de leurs EDSR associées. On suppose que :*

- $\xi^1 \leq \xi^2$ *p.s.*
- $f^1(t, Y_t^1, Z_t^1) \leq f^2(t, Y_t^1, Z_t^1)$ $dt \times dP$ *p.p.*
- $f^2(t, Y_t^1, Z_t^1) \in \mathbb{H}^2(0, T)$

Alors $Y_t^1 \leq Y_t^2$ p.s., pour tout $0 \leq t \leq T$.

De plus, si $Y_0^2 \leq Y_0^1$, alors $Y_t^1 = Y_t^2$, $0 \leq t \leq T$. En particulier, si $P(\xi^1 < \xi^2) > 0$ ou $f^1(t, ., .) < f^2(t, ., .)$ sur un ensemble de mesure $dt \times dP$ strictement positif, alors $Y_0^1 < Y_0^2$.

Preuve. Pour simplifier les notations, on suppose $d = 1$. On note $\bar{Y} = Y^2 - Y^1$, $\bar{Z} = Z^2 - Z^1$. Alors $(\bar{Y}, \bar{Z})$ satisfait l'EDSR linéaire

$$
-d\bar{Y}_t = \left(\Delta_t^y \, \bar{Y}_t + \Delta_t^z \, \bar{Z}_t + \bar{f}_t \right) dt - \bar{Z}_t dW_t, \quad \bar{Y}_T = \xi^2 - \xi^1. \qquad (5.10)
$$

où

$$\begin{aligned}
\Delta_t^y &= \frac{f^2(t, Y_t^2, Z_t^2) - f^2(t, Y_t^1, Z_t^2)}{Y_t^2 - Y_t^1} \, 1_{Y_t^2 - Y_t^1 \neq 0} \\
\Delta_t^z &= \frac{f^2(t, Y_t^1, Z_t^2) - f^2(t, Y_t^1, Z_t^1)}{Z_t^2 - Z_t^1} \, 1_{Z_t^2 - Z_t^1 \neq 0} \\
\bar{f}_t &= f^2(t, Y_t^1, Z_t^1) - f^1(t, Y_t^1, Z_t^1).
\end{aligned}$$

Comme le générateur f^2 est uniformément Lipschitz en y et z, les processus Δ^y et Δ^z sont bornés. De plus, $\bar{f}_t$ est un processus positif dans $\mathbb{H}^2(0, T)$. D'après la proposition 5.2.1, $\bar{Y}$ est donné par :

$$\Gamma_t \bar{Y}_t = E\left[\Gamma_T(\xi^2 - \xi^1) + \int_t^T \Gamma_s \bar{f}_s ds \middle| \mathcal{F}_t \right],$$

où Γ est le processus adjoint strictement positif. Ceci conclut la preuve avec cette formule sous forme d'espérance et la positivité de $\xi^2 - \xi^1$ et $\bar{f}$. □

Remarque 5.2.1 Notons que dans la démonstration du théorème 5.2.2, il n'est pas nécessaire de supposer de la régularité sur le générateur f_1, et en particulier la condition de Lipschitz uniforme sur f_1. Cette condition est uniquement requise sur f_2.

Corollaire 5.2.1 *Si le couple* (ξ, f) *vérifie* $\xi \geq 0$ *a.s. et* $f(t, 0, 0) \geq 0$ $dt \times dP$ *a.e., alors* $Y_t \geq 0$, $0 \leq t \leq T$ *a.s. De plus si* $P[\xi > 0] > 0$ *ou* $f(t, 0, 0) > 0$ $dt \times dP$ *p.p. alors* $Y_0 > 0$.

Preuve. C'est une conséquence immédiate du théorème de comparaison 5.2.2 avec $(\xi^1, f^1) = (0, 0)$ dont la solution est $(Y^1, Z^1) = (0, 0)$. □

5.3 EDSR, EDP et formules de type Feynman-Kac

On rappelle le résultat bien connu (voir paragraphe 1.3.3) que la solution de l'EDP parabolique linéaire

$$-\frac{\partial v}{\partial t} - \mathcal{L}v - \beta(t, x)v - f(t, x) = 0 \tag{5.11}$$

$$v(T, x) = g(x) \tag{5.12}$$

peut se représenter de façon probabiliste par

$$\begin{aligned}
v(t, x) = E\Bigg[&\int_t^T e^{-\int_t^s \beta(u, X_u^{t,x}) du} f(s, X_s^{t,x}) ds \\
&+ e^{-\int_t^T \beta(u, X_u^{t,x}) du} g(X_T^{t,x}) \Bigg],
\end{aligned} \tag{5.13}$$

où $\{X_s^{t,x}, t \leq s \leq T\}$ est la solution de l'EDS

$$dX_s = b(X_s)ds + \sigma(X_s)dW_s, \quad t \leq s \leq T, \; X_t = x,$$

et $\mathcal{L}$ est l'opérateur du second ordre

$$\mathcal{L}v = b(x).D_x v + \frac{1}{2}\mathrm{tr}(\sigma(x)\sigma'(x)D^2_{xx}v).$$

On a vu au chapitre précédent une généralisation de la formule de Feynman-Kac (5.13) pour des EDP non linéaires de la forme :

$$-\frac{\partial v}{\partial t} + \sup_{a\in A}\left[-\mathcal{L}^a v - \beta(t,x,a)v - f(t,x,a)\right] = 0 \tag{5.14}$$

$$v(T,x) = g(x) \tag{5.15}$$

où pour tout $a \in A$, sous espace de $\mathbb{R}^m$,

$$\mathcal{L}^a v = b(x,a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma(x,a)\sigma'(x,a)D^2_{xx}v).$$

La solution (de viscosité) de (5.14)-(5.15) peut se représenter à l'aide d'un problème de contrôle stochastique :

$$\begin{aligned} v(t,x) = \inf_{\alpha\in\mathcal{A}} E\Bigg[\int_t^T & e^{-\int_t^s \beta(u,X_u^{t,x},\alpha_u)du} f(s,X_s^{t,x},\alpha_s)ds \\ & + \; e^{-\int_t^T \beta(u,X_u^{t,x},\alpha_u)du} g(X_T^{t,x})\Bigg], \end{aligned} \tag{5.16}$$

où $\mathcal{A}$ est l'ensemble des processus α progressifs à valeurs dans A et pour $\alpha \in \mathcal{A}$, $\{X_s^{t,x}, t \leq s \leq T\}$ est solution de la diffusion contrôlée :

$$dX_s = b(X_s,\alpha_s)ds + \sigma(X_s,\alpha_s)dW_s, \quad t \leq s \leq T, \; X_t = x.$$

Dans ce chapitre, on étudie une autre extension de la formule de Feynman-Kac pour des EDP non linéaires de la forme :

$$-\frac{\partial v}{\partial t} - \mathcal{L}v - f(t,x,v,(D_x v)'\sigma(x)) = 0 \tag{5.17}$$

$$v(T,x) = g(x). \tag{5.18}$$

Nous allons représenter la solution de cette EDP à l'aide de l'EDSR unidimensionnelle :

$$-dY_s = f(s,X_s,Y_s,Z_s)ds - Z_s dW_s, \quad t \leq s \leq T, \;\; Y_T = g(X_T), \tag{5.19}$$

et de l'EDS à valeurs dans $\mathbb{R}^n$:

$$dX_s = b(X_s)ds + \sigma(X_s)dW_s, \quad t \leq s \leq T, \;\; X_t = x. \tag{5.20}$$

Les fonction b et σ satisfont une condition de Lipschitz et $f(t,x,y,z)$ est une fonction continue vérifiant une condition de Lipschitz en (y,z) uniformément en (t,x). La fonction continue g satisfait une condition de croissance linéaire. Ainsi le générateur f vérifie la condition (B) énoncée au paragraphe 5.2. On note alors par $\{X_s^{t,x}, t \le s \le T\}$ la solution de l'EDS (5.20) et $\{(Y_s^{t,x}, Z_s^{t,x}), t \le s \le T\}$ la solution de l'EDSR (5.19) avec $X_s = X_s^{t,x}$. Nous avons alors la proposition de vérification suivante pour l'EDSR (5.19), qui est un résultat analogue au théorème de vérification pour les équations d'HJB par la programmation dynamique.

Proposition 5.3.2 *Soit* $v \in C^{1,2}([0,T[\times\mathbb{R}^n) \cap C^0([0,T]\times\mathbb{R}^n)$ *une solution classique de* (5.17)-(5.18) *telle que :*

$$|D_x v(t,x)| \le C(1+|x|), \quad (t,x) \in [0,T]\times\mathbb{R}^n, \tag{5.21}$$

où C est une constante positive. Alors $\{(v(s,X_s^{t,x}), D_x v(s,X_s^{t,x})'\sigma(s,X_s^{t,x})), t \le s \le T\} = \{(Y_s^{t,x}, Z_s^{t,x}), t \le s \le T\}$ *est la solution de l'EDSR* (5.19). *En particulier,* $v(t,x) = Y_t^{t,x}$.

Preuve. C'est une conséquence directe de la formule d'Itô appliquée à $v(s,X_s^{t,x})$ et en notant que d'après la condition (5.21), $Z_s^{t,x} = D_x v(s,X_s^{t,x})'$ $\sigma(s,X_s^{t,x})$ appartient bien à $\mathbb{H}^2(t,T)^{1\times d}$ i.e. :

$$E\left[\int_t^T \left|D_x v(s,X_s^{t,x})'\sigma(s,X_s^{t,x})\right|^2 ds\right] < +\infty.$$

□

On s'intéresse maintenant à la réciproque de ce résultat : on montre que la solution de l'EDSR (5.19) fournit une solution à l'EDP (5.17)-(5.18).

Théorème 5.3.3 *La fonction* $v(t,x) = Y_t^{t,x}$ *est une fonction continue de* $(t,x) \in [0,T]\times\mathbb{R}^n$, *et est solution de viscosité de* (5.17) *avec la condition terminale* $v(T,x) = g(x)$.

Preuve. 1) Pour (t_1,x_1), $(t_2,x_2) \in [0,T]\times\mathbb{R}^n$, avec sans perte de généralité $t_1 \le t_2$, on note pour $i = 1,2$, $X_s^i = X_s^{t_i,x_i}$, avec la convention $X_s^2 = x_2$ pour $t_1 \le s \le t_2$. et $(Y_s^i, Z_s^i) = (Y_s^{t_i,x_i}, Z_s^{t_i,x_i})$, qui est donc bien défini pour $t_1 \le s \le T$. En appliquant la formule d'Itô à $|Y_s^1 - Y_s^2|^2$ entre $s = t \in [t_1,T]$ et $s = T$, on a :

$$\begin{aligned}|Y_t^1 - Y_t^2|^2 = |g(X_T^1) - g(X_T^2)|^2 &- \int_t^T |Z_s^1 - Z_s^2|^2 ds \\ &+ 2\int_t^T (Y_s^1 - Y_s^2).(f(s,X_s^1,Y_s^1,Z_s^1) - f(s,X_s^2,Y_s^2,Z_s^2))ds \\ &- 2\int_t^T (Y_s^1 - Y_s^2)'(Z_s^1 - Z_s^2)dW_s.\end{aligned}$$

Comme dans la preuve du théorème 5.2.1, la martingale locale $\int_t^s (Y_u^1 - Y_u^2)'(Z_u^1 - Z_u^2)dW_u$, $t \leq s \leq T$, est en fait uniformément intégrable et donc en prenant l'espérance dans la relation ci-dessus, on obtient :

$$
\begin{aligned}
& E\Big[|Y_t^1 - Y_t^2|^2\Big] + E\left[\int_t^T |Z_s^1 - Z_s^2|^2 ds\right] \\
&= E\Big[|g(X_T^1) - g(X_T^2)|^2\Big] \\
&\quad + 2\,E\left[\int_t^T (Y_s^1 - Y_s^2).(f(s, X_s^1, Y_s^1, Z_s^1) - f(s, X_s^2, Y_s^2, Z_s^2))ds\right] \\
&\leq E\Big[|g(X_T^1) - g(X_T^2)|^2\Big] \\
&\quad + 2\,E\left[\int_t^T |Y_s^1 - Y_s^2|\,|f(s, X_s^1, Y_s^1, Z_s^1) - f(s, X_s^2, Y_s^1, Z_s^1)|ds\right] \\
&\quad + 2K_f E\left[\int_t^T |Y_s^1 - Y_s^2|\left(|Y_s^1 - Y_s^2| + |Z_s^1 - Z_s^2|\right)ds\right] \\
&\leq E\Big[|g(X_T^1) - g(X_T^2)|^2\Big] \\
&\quad + E\left[\int_t^T |f(s, X_s^1, Y_s^1, Z_s^1) - f(s, X_s^2, Y_s^1, Z_s^1)|^2 ds\right] \\
&\quad + (1 + 4K_f^2)E\left[\int_t^T |Y_s^1 - Y_s^2|^2 ds + \frac{1}{2}E\int_t^T |Z_s^1 - Z_s^2|^2 ds\right],
\end{aligned}
$$

où K_f est la constante de Lipschitz uniforme de f par rapport à y et z. On a alors,

$$
\begin{aligned}
E\Big[|Y_t^1 - Y_t^2|^2\Big] \leq\; & E\Big[|g(X_T^1) - g(X_T^2)|^2\Big] \\
& + E\left[\int_t^T |f(s, X_s^1, Y_s^1, Z_s^1) - f(s, X_s^2, Y_s^1, Z_s^1)|^2 ds\right] \\
& + (1 + 4K_f^2)E\Big[\int_t^T |Y_s^1 - Y_s^2|^2 ds\Big]
\end{aligned}
$$

et par le lemme de Gronwall,

$$
\begin{aligned}
E\Big[|Y_t^1 - Y_t^2|^2\Big] \leq C\Bigg\{ & E\Big[|g(X_T^1) - g(X_T^2)|^2\Big] \\
& + E\left[\int_t^T |f(s, X_s^1, Y_s^1, Z_s^1) - f(s, X_s^2, Y_s^1, Z_s^1)|^2 ds\right]\Bigg\}.
\end{aligned}
$$

Cette dernière inégalité combinée avec la continuité de f et g en x et $X^{t,x}$ par rapport à (t, x) montre la continuité de $(t, x) \in [0, T] \times \mathbb{R}^n \to EY_s^{t,x}$ pour tout $t \leq s \leq T$. Puisqu'on a aussi la continuité de $s \in [t,T] \to EY_s^{t,x}$ pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$, ceci prouve la continuité de $Y_t^{t,x} = EY_t^{t,x}$ en (t,x).

2) Nous montrons maintenant que $u(t,x) = Y_t^{t,x}$ est une solution de viscosité de (5.17). On montre la propriété de sous-solution, celle de sursolution étant prouvée de manière similaire. Soit donc φ une fonction test régulière et $(t,x) \in [0,T[\times\mathbb{R}^n$ tel que (t,x) soit un maximum local de $v - \varphi$ avec $u(t,x) = \varphi(t,x)$. On raisonne par contradiction en supposant que

$$-\frac{\partial\varphi}{\partial t}(t,x) - \mathcal{L}\varphi(t,x) - f(t,x,v(t,x),(D_x\varphi)'(t,x)\sigma(x)) > 0.$$

Par continuité de f et de φ et ses dérivées, il existe h, $\varepsilon > 0$ tel que pour tous $t \leq s \leq t+h$, $|x-y| \leq \varepsilon$,

$$v(s,y) \leq \varphi(s,y) \tag{5.22}$$

$$-\frac{\partial\varphi}{\partial t}(s,y) - \mathcal{L}\varphi(s,y) - f(s,y,v(s,y),(D_x\varphi)'(s,y)\sigma(y)) > 0. \tag{5.23}$$

Soit $\tau = \inf\{s \geq t \ : |X_s^{t,x} - x| \geq \varepsilon\} \wedge (t+h)$. Considérons le couple

$$(Y_s^1, Z_s^1) = (Y_{s\wedge\tau}^{t,x}, 1_{[0,\tau]}(s)Z_s^{t,x}), \quad t \leq s \leq t+h.$$

Alors par construction, (Y_s^1, Z_s^1) est solution de l'EDSR :

$$\begin{aligned}-dY_s^1 &= 1_{[0,\tau]}(s)f(s,X_s^{t,x},u(s,X_s^{t,x}),Z_s^1)ds - Z_s^1 dW_s, \quad t \leq s \leq t+h,\\ Y_{t+h}^1 &= u(\tau, X_\tau^{t,x}).\end{aligned}$$

D'autre part, le couple

$$(Y_s^2, Z_s^2) = (\varphi(s, X_{s\wedge\tau}^{t,x}), 1_{[0,\tau]}(s)D_x\varphi(s,X_s^{t,x})'\sigma(X_s^{t,x})), \quad t \leq s \leq t+h.$$

satisfait, d'après la formule d'Itô, l'EDSR

$$\begin{aligned}-dY_s^2 &= 1_{[0,\tau]}(s)(\frac{\partial\varphi}{\partial t} + \mathcal{L}\varphi)(s,X_s^{t,x}) - Z_s^2 dW_s, \quad t \leq s \leq t+h,\\ Y_{t+h}^1 &= \varphi(\tau, X_\tau^{t,x}).\end{aligned}$$

D'après les inégalités (5.22)-(5.23), et le principe de comparaison strict dans le théorème 5.2.2, on a $Y_0^1 < Y_0^2$, i.e. $u(t,x) < \varphi(t,x)$ qui est la contradiction requise. □

5.4 Contrôle et EDSR

Dans cette section, nous indiquons comment les EDSR peuvent être utilisées pour étudier des problèmes de contrôle stochastique.

5.4.1 Optimisation d'une famille d'EDSR

Théorème 5.4.4 *Soient* (ξ, f) *et* (ξ^α, f^α), $\alpha \in \mathcal{A}$ *sous-ensemble de processus progressifs, une famille de couple condition terminale-générateur, et* (Y, Z), (Y^α, Z^α) *les solutions de leurs EDSR (unidimensionnelles) associées. On suppose qu'il existe* $\hat{\alpha} \in \mathcal{A}$ *tel que*

$$
\begin{aligned}
f(t, Y_t, Z_t) = \operatorname*{ess\,inf}_{\alpha} f^\alpha(t, Y_t, Z_t) \ &= \ f^{\hat{\alpha}}(t, Y_t, Z_t), \quad dt \times dP \ \ p.p. \\
\xi = \operatorname*{ess\,inf}_{\alpha} \xi^\alpha \ &= \ \xi^{\hat{\alpha}}.
\end{aligned}
$$

Alors,

$$
Y_t = \operatorname*{ess\,inf}_{\alpha} Y_t^\alpha \ = \ Y_t^{\hat{\alpha}}, \quad 0 \leq t \leq T, \ p.s.
$$

Preuve. D'après le théorème de comparaison 5.2.2, puisque $\xi \leq \xi^\alpha$ et $f(t, Y_t, Z_t) \leq f^\alpha(t, Y_t, Z_t)$, on a $Y_t \leq Y_t^\alpha$ pour tout α et donc

$$
Y_t \leq \operatorname*{ess\,inf}_{\alpha} Y_t^\alpha.
$$

D'autre part, s'il existe $\hat{\alpha}$ tel que $\xi = \xi^{\hat{\alpha}}$ et $f(t, Y_t, Z_t) = f^{\hat{\alpha}}(t, Y_t, Z_t)$ alors (Y, Z) et $(Y^{\hat{\alpha}}, Z^{\hat{\alpha}})$ sont tous les deux solutions de la même EDSR avec condition $(\xi^{\hat{\alpha}}, f^{\hat{\alpha}})$: par unicité, ils sont égaux et donc

$$
\operatorname*{ess\,inf}_{\alpha} Y_t^\alpha \leq \ Y_t^{\hat{\alpha}} \ = \ Y_t \ \leq \operatorname*{ess\,inf}_{\alpha} Y_t^\alpha,
$$

ce qui conclut la preuve. □

A l'aide du résultat précédent, on montre comment la solution d'une EDSR avec un générateur concave peut se représenter comme la fonction valeur d'un problème de contrôle.

Soit $f(t, y, z)$ un générateur d'EDSR, concave en (y, z) et (Y, Z) la solution de l'EDSR associée à (ξ, f). On considère la transformée polaire de f :

$$
F(t, b, c) = \sup_{(y,z) \in \mathbb{R} \times \mathbb{R}^{1 \times d}} \left[f(t, y, z) - yb - zc \right], \quad (b, c) \in \mathbb{R} \times \mathbb{R}^d. \tag{5.24}
$$

Comme f est concave, on a la relation de dualité :

$$
f(t, y, z) = \inf_{(b,c) \in \mathbb{R} \times \mathbb{R}^d} \left[F(t, b, c) + yb + zc \right], \quad (y, z) \in \mathbb{R} \times \mathbb{R}^{1 \times d}. \tag{5.25}
$$

On note par $\mathcal{A}$ l'ensemble des processus progressifs (β, γ) bornés, à valeurs dans $\mathbb{R} \times \mathbb{R}^d$ tels que

$$
E \left[\int_0^T |F(t, \beta_t, \gamma_t)|^2 dt \right] < +\infty.
$$

La condition de bornitude sur $\mathcal{A}$ signifie que pour tout $(\beta,\gamma) \in \mathcal{A}$, il existe une constante (dépendante de (β,γ)) telle que $|\beta_t| + |\gamma_t| \leq C$, $dt \times dP$ p.p. On considére la famille de générateurs linéaires :

$$f^{\beta,\gamma}(t,y,z) = F(t,\beta_t,\gamma_t) + y\beta_t + z\gamma_t, \quad (\beta,\gamma) \in \mathcal{A}.$$

Etant donné $(\beta,\gamma) \in \mathcal{A}$, on note par $(Y^{\beta,\gamma}, Z^{\beta,\gamma})$ la solution de l'EDSR linéaire associée à $(\xi, f^{\beta,\gamma})$.

Théorème 5.4.5 *Y s'écrit comme la fonction valeur du problème de contrôle*

$$Y_t = \operatorname{ess}\inf_{\beta,\gamma\in\mathcal{A}} Y_t^{\beta,\gamma}, \quad 0 \leq t \leq T, \ p.s. \tag{5.26}$$

$$Y_t^{\beta,\gamma} = E^{Q^\gamma}\left[\int_t^T e^{\int_t^s \beta_u du} F(s,\beta_s,\gamma_s)ds + e^{\int_t^T \beta_u du}\xi \,\middle|\, \mathcal{F}_t\right],$$

où Q^γ est la probabilité de processus de densité martingale :

$$dL_t = L_t\gamma_t' dW_t, \quad L_0 \ = \ 1.$$

Preuve. 1) Notons que d'après la relation (5.25), on a $f(t,Y_t,Z_t) \leq f^{\beta,\gamma}$ (t,Y_t,Z_t) pour tout $(\beta,\gamma) \in \mathcal{A}$. De plus, comme f est concave à croissance linéaire, pour chaque (t,ω,y,z), l'infimum dans la relation (5.25) est atteint en $(\hat{b}(t,y,z), \hat{c}(t,y,z))$ appartenant au sous-différentiel de $-f$ et donc borné par la constante de Lipschitz de f. Par un argument de sélection mesurable (voir e.g. Appendice au chapitre III dans Dellacherie et Meyer [DM75]), comme Y,Z sont progressifs, on peut trouverun couple de processus $(\hat{\beta},\hat{\gamma})$ progressifs et bornés tel que

$$f(t,Y_t,Z_t) = f^{\hat{\beta},\hat{\gamma}}(t,Y_t,Z_t) \ = \ F(t,\hat{\beta}_t,\hat{\gamma}_t) + Y_t\hat{\beta}_t + Z_t\hat{\gamma}_t, \quad 0 \leq t \leq T, \ p.s.$$

On en déduit la relation (5.26) d'après le théorème 5.4.4.

2) D'autre part, d'après la proposition 5.2.1, la solution $Y^{\beta,\gamma}$ de l'EDSR linéaire associée à $(\xi, f^{\beta,\gamma})$ s'exprime comme :

$$\Gamma_t Y_t^{\beta,\gamma} = E\left[\int_t^T \Gamma_s F(s,\beta_s,\gamma_s)ds + \Gamma_T\xi \,\middle|\, \mathcal{F}_t\right],$$

où Γ est le processus adjoint (ou dual) donné par l'EDS :

$$d\Gamma_t = \Gamma_t\left(\beta_t dt + \gamma_t' dW_t\right), \quad \Gamma_0 = 1.$$

On conclut en notant que $\Gamma_t = e^{\int_0^t \beta_u du} L_t$ et en utilisant la formule de Bayes.
□

5.4.2 Principe du maximum stochastique

Dans le chapitre précédent, nous avons vu comment résoudre un problème de contrôle stochastique par la méthode de la programmation dynamique de Bellman. Nous présentons dans ce paragraphe une approche alternative appelée principe du maximum de Pontryagin et basée sur des conditions d'optimalité du contrôle.

On se place dans le cadre d'un problème de contrôle stochastique à horizon fini défini au chapitre 3 : soit la diffusion contrôlée dans $\mathbb{R}^n$

$$dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s, \tag{5.27}$$

où W est un mouvement Brownien standard d-dimensionnel, $\alpha \in \mathcal{A}$ est le processus de contrôle progressif à valeurs dans A. La fonctionnelle de coût à minimiser est :

$$J(\alpha) = E\left[\int_0^T f(t, X_t, \alpha_t)dt + g(X_T)\right].$$

où $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ est continue en (t,x) pour tout a dans A, $g : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe C^1, et f, g sont à croissance quadratique en x.

On définit l'Hamiltonien généralisé $\mathcal{H} : [0,T] \times \mathbb{R}^n \times A \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}$:

$$\mathcal{H}(t,x,a,y,z) = b(x,a).y + \mathrm{tr}(\sigma'(x,a)z) + f(t,x,a), \tag{5.28}$$

et on suppose que $\mathcal{H}$ est dérivable en x de dérivée notée $D_x\mathcal{H}$. On considère alors pour tout $\alpha \in \mathcal{A}$, l'EDSR, appelée aussi équation adjointe

$$-dY_t = D_x\mathcal{H}(t, X_t, \alpha_t, Y_t, Z_t)dt - Z_t dW_t, \quad Y_T \;=\; D_x g(X_T). \tag{5.29}$$

Théorème 5.4.6 *Soit $\hat{\alpha} \in \mathcal{A}$ et $\hat{X}$ la diffusion contrôlée associée. Supposons qu'il existe une solution $(\hat{Y}, \hat{Z})$ à l'EDSR correspondante* (5.29) *telle que :*

$$\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) = \min_{a\in A} \mathcal{H}(t, \hat{X}_t, a, \hat{Y}_t, \hat{Z}_t), \quad 0 \le t \le T, \;\; p.s. \tag{5.30}$$

et

$$(x,a) \to \mathcal{H}(t, x, a, \hat{Y}_t, \hat{Z}_t) \quad \textit{est une fonction convexe}\;, \tag{5.31}$$

pour tout $t \in [0,T]$. Alors $\hat{\alpha}$ est un contrôle optimal, i.e.

$$J(\hat{\alpha}) = \inf_{\alpha\in\mathcal{A}} J(\alpha).$$

Preuve. Pour tout $\alpha \in \mathcal{A}$, on a

$$J(\hat{\alpha}) - J(\alpha) = E\left[\int_0^T f(t, \hat{X}_t, \hat{\alpha}_t) - f(t, X_t, \alpha_t)dt + g(\hat{X}_T) - g(X_T)\right]. \tag{5.32}$$

D'après la convexité de g et le produit d'Itô, on a

$$
\begin{aligned}
E\Big[g(\hat{X}_T) - g(X_T)\Big] \;\le\; E\Big[(\hat{X}_T - X_T).D_x g(\hat{X}_T)\Big] \;=\; E\Big[(\hat{X}_T - X_T).\hat{Y}_T\Big] \\
= E\Bigg\{\int_0^T (\hat{X}_t - X_t).d\hat{Y}_t + \int_0^T \hat{Y}_t.(d\hat{X}_t - dX_t) \\
+ \int_0^T \mathrm{tr}\Big[(\sigma(\hat{X}_t,\hat{\alpha}_t) - \sigma(X_t,\alpha_t))'\hat{Z}_t\Big]\,dt\Bigg\} \\
= E\Bigg\{\int_0^T (\hat{X}_t - X_t).(-D_x\mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,\hat{Y}_t,\hat{Z}_t))dt \\
+ \int_0^T \hat{Y}_t.(b(\hat{X}_t,\hat{\alpha}_t) - b(X_t,\alpha_t))dt \\
+ \int_0^T \mathrm{tr}\Big[(\sigma(\hat{X}_t,\hat{\alpha}_t) - \sigma(X_t,\alpha_t))'\hat{Z}_t\Big]\,dt\Bigg\}.
\end{aligned}
\tag{5.33}
$$

D'autre part, d'après la définition de $\mathcal{H}$, on a :

$$
\begin{aligned}
E\left[\int_0^T f(t,\hat{X}_t,\hat{\alpha}_t) - f(t,X_t,\alpha_t)dt\right] = E\Bigg\{\int_0^T \mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,\hat{Y}_t,\hat{Z}_t) \\
-\mathcal{H}(t,X_t,\alpha_t,\hat{Y}_t,\hat{Z}_t)dt - \int_0^T (b(\hat{X}_t,\hat{\alpha}_t) - b(X_t,\alpha_t)).\hat{Y}_t \\
- \int_0^T \mathrm{tr}\Big[(\sigma(\hat{X}_t,\hat{\alpha}_t) - \sigma(X_t,\alpha_t))'\hat{Z}_t\Big]\,dt\Bigg\}
\end{aligned}
\tag{5.34}
$$

En ajoutant (5.33) et (5.34) dans (5.32), on obtient :

$$
\begin{aligned}
J(\hat{\alpha}) - J(\alpha) \le E\Bigg[\int_0^T \mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,\hat{Y}_t,\hat{Z}_t) - \mathcal{H}(t,X_t,\alpha_t,\hat{Y}_t,\hat{Z}_t)dt \\
- \int_0^T (\hat{X}_t - X_t).D_x\mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,\hat{Y}_t,\hat{Z}_t)dt\Bigg].
\end{aligned}
$$

Sous les conditions (5.30) et (5.31), le terme entre crochet dans la relation ci-dessus est négatif ce qui conclut la preuve. □

On termine ce paragraphe en donnant le lien entre le principe du maximum et la programmation dynamique. On définit la fonction valeur du problème de contrôle stochastique considéré ci-dessus :

$$
v(t,x) = \inf_{\alpha\in\mathcal{A}} E\left[\int_t^T f(s,X_s^{t,x},\alpha_s)ds + g(X_T^{t,x})\right], \tag{5.35}
$$

où $\{X_s^{t,x}, t \leq s \leq T\}$ est la solution de (5.27) partant de x en t. On rappelle que l'EDP d'Hamilton-Jacobi-Bellman s'écrit :

$$-\frac{\partial v}{\partial t} + \sup_{a \in A} \left[-\mathcal{G}(t,x,a,D_x v, D_x^2 v) \right] = 0, \tag{5.36}$$

où pour $(t,x,a,p,M) \in [0,T] \times \mathbb{R}^n \times A \times \mathbb{R}^n \times \mathcal{S}_n$,

$$\mathcal{G}(t,x,a,p,M) = b(x,a).p + \frac{1}{2}\mathrm{tr}(\sigma\sigma'(x,a)M) + f(t,x,a). \tag{5.37}$$

Théorème 5.4.7 *Supposons que $v \in C^{1,3}([0,T[\times\mathbb{R}^n) \cap C^0([0,T] \times \mathbb{R}^n)$ et qu'il existe un contrôle optimal $\hat{\alpha} \in \mathcal{A}$ à* (5.35) *de diffusion contrôlée associée $\hat{X}$. Alors*

$$\mathcal{G}(t,\hat{X}_t,\hat{\alpha}_t,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)) = \min_{a \in A} \mathcal{G}(t,\hat{X}_t,a,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)), \tag{5.38}$$

et le couple

$$(\hat{Y}_t,\hat{Z}_t) = (D_x v(t,\hat{X}_t)\ ,\ D_x^2 v(t,\hat{X}_t)\ \sigma(\hat{X}_t,\hat{\alpha}_t)), \tag{5.39}$$

est solution de l'EDSR adjointe (5.29)*.*

Preuve. Puisque $\hat{\alpha}$ est un contrôle optimal, on a :

$$\begin{aligned} v(t,\hat{X}_t) &= E\left[\int_t^T f(s,\hat{X}_s,\hat{\alpha}_s)ds + g(\hat{X}_T) \middle| \mathcal{F}_t \right] \\ &= -\int_0^t f(s,\hat{X}_s,\hat{\alpha}_s)ds + M_t, \quad 0 \leq t \leq T,\ p.s. \end{aligned} \tag{5.40}$$

où M est la martingale $M_t = E\left[\int_0^T f(s,\hat{X}_s,\hat{\alpha}_s)ds + g(\hat{X}_T) \middle| \mathcal{F}_t \right]$. En appliquant la formule d'Itô à $v(t,\hat{X}_t)$ et en identifiant les termes en dt dans la relation (5.40), on obtient :

$$-\frac{\partial v}{\partial t}(t,\hat{X}_t)\ -\ \mathcal{G}(t,\hat{X}_t,\hat{\alpha}_t,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)) = 0. \tag{5.41}$$

Comme v est régulière, v satisfait l'EDP d'HJB (5.36), ce qui implique (5.38).

D'après (5.36) et (5.41), on a :

$$\begin{aligned} 0 &= \frac{\partial v}{\partial t}(t,\hat{X}_t)\ +\ \mathcal{G}(t,\hat{X}_t,\hat{\alpha}_t,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)) \\ &\leq \frac{\partial v}{\partial t}(t,x)\ +\ \mathcal{G}(t,x,\hat{\alpha}_t,D_x v(t,x),D_x^2 v(t,x)), \quad \forall x \in \mathbb{R}^n, \end{aligned}$$

ce qui implique puisque v est $C^{1,3}$:

$$\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial t}(t,x) \;+\; \mathcal{G}(t,x,\hat{\alpha}_t,D_x v(t,x),D_x^2 v(t,x))\right)\Big|_{x=\hat{X}_t} = 0.$$

En se rappelant l'expression (5.37) de $\mathcal{G}$ et celle (5.28) de $\mathcal{H}$, l'égalité précédente s'écrit :

$$\frac{\partial^2 v}{\partial t \partial x}(t,\hat{X}_t) + D_x^2 v(t,\hat{X}_t) b(\hat{X}_t,\hat{\alpha}_t) + \frac{1}{2}\mathrm{tr}(\sigma\sigma'(\hat{X}_t,\hat{\alpha}_t) D_x^3 v(t,\hat{X}_t)) \\ + D_x\mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)\;\sigma(\hat{X}_t,\hat{\alpha}_t)) = 0. \tag{5.42}$$

En appliquant alors la formule d'Itô à $D_x v(t,\hat{X}_t)$ et en utilisant (5.42), on a :

$$\begin{aligned} -dD_x v(t,\hat{X}_t) &= -\left[\frac{\partial^2 v}{\partial t \partial x}(t,\hat{X}_t) + D_x^2 v(t,\hat{X}_t) b(\hat{X}_t,\hat{\alpha}_t)\right. \\ &\qquad \left. + \frac{1}{2}\mathrm{tr}(\sigma\sigma'(\hat{X}_t,\hat{\alpha}_t) D_x^3 v(t,\hat{X}_t))\right] dt \\ &\qquad - D_x^2 v(t,\hat{X}_t)\;\sigma(\hat{X}_t,\hat{\alpha}_t) dW_t \\ &= D_x\mathcal{H}(t,\hat{X}_t,\hat{\alpha}_t,D_x v(t,\hat{X}_t),D_x^2 v(t,\hat{X}_t)\;\sigma(\hat{X}_t,\hat{\alpha}_t))\,dt \\ &\qquad - \; D_x^2 v(t,\hat{X}_t)\;\sigma(\hat{X}_t,\hat{\alpha}_t) dW_t. \end{aligned}$$

Comme de plus $v(T,.) = g(.)$, on a

$$D_x v(T,\hat{X}_T) = D_x g(\hat{X}_T).$$

Ceci prouve le résultat (5.39). □

5.5 Applications

5.5.1 Maximisation d'utilité exponentielle avec actif contingent

On considère un marché financier avec un actif sans risque de prix $S^0 = 1$ et un actif risqué de processus de prix :

$$dS_t = S_t(b_t dt + \sigma_t dW_t),$$

où W est un mouvement Brownien réel standard sur $(\Omega,\mathcal{F},\mathbb{F} = (\mathcal{F}_t)_t,P)$ avec $\mathbb{F}$ filtration naturelle de W, b et σ sont deux processus progressifs bornés, $\sigma_t \geq \varepsilon$, pour tout t, p.s. où $\varepsilon > 0$. Un agent partant d'un capital x, investit un montant α_t à toute date t dans l'actif risqué. Son processus de richesse, contrôlée par α , est donc donné par :

$$X_t^{x,\alpha} = x + \int_0^t \alpha_u \frac{dS_u}{S_u} \;=\; x + \int_0^t \alpha_u(b_u du + \sigma_u dW_u), \quad 0 \leq t \leq T. \tag{5.43}$$

On note $\mathcal{A}$ l' ensemble des processus progressifs α à valeurs dans $\mathbb{R}$, tel que $\int_0^T |\alpha_t|^2 dt < +\infty$ p.s. et $X^{x,\alpha}$ est borné inférieurement. En échange du capital reçu x initialement, l'agent doit verser à l'horizon T un actif contingent représenté par une variable aléatoire ξ $\mathcal{F}_T$-mesurable et supposée bornée. Etant donnée son aversion pour le risque caractérisée par une fonction d'utilité exponentielle

$$U(x) = -\exp(-\eta x), \quad x \in \mathbb{R}, \ \ \eta > 0, \tag{5.44}$$

l'objectif de l'agent est de résoudre le problème de maximisation :

$$v(x) = \sup_{\alpha \in \mathcal{A}} E[U(X_T^{x,\alpha} - \xi)]. \tag{5.45}$$

L'approche adoptée pour déterminer la fonction valeur v et le contrôle optimal $\hat{\alpha}$ est générale et basée sur le principe suivant. On construit une famille de processus $(J_t^\alpha)_{0 \leq t \leq T}$, $\alpha \in \mathcal{A}$, satisfaisant les propriétés :

(**i**) $J_T^\alpha = U(X_T^{x,\alpha} - \xi)$ pour tout $\alpha \in \mathcal{A}$

(**ii**) J_0^α est une constante indépendante de $\alpha \in \mathcal{A}$

(**iii**) J^α est une surmartingale pour tout $\alpha \in \mathcal{A}$ et il existe $\hat{\alpha} \in \mathcal{A}$ tel que $J^{\hat{\alpha}}$ soit une martingale.

En effet, dans ce cas, on aura pour un tel $\hat{\alpha}$ et pour tout $\alpha \in \mathcal{A}$,

$$E[U(X_T^{x,\alpha} - \xi)] = E[J_T^\alpha] \leq J_0^\alpha = J_0^{\hat{\alpha}} = E[J_T^{\hat{\alpha}}] = E[U(X_T^{x,\hat{\alpha}} - \xi)] = v(x),$$

ce qui prouvera que $\hat{\alpha}$ est un contrôle optimal.

Pour construite une telle famille (J_t^α), on la cherche de la forme :

$$J_t^\alpha = U(X_t^{x,\alpha} - Y_t), \quad 0 \leq t \leq T, \ \alpha \in \mathcal{A}, \tag{5.46}$$

avec (Y, Z) solution de l'EDSR

$$Y_t = \xi + \int_t^T f(s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \tag{5.47}$$

où f est un générateur à déterminer. Les conditions (**i**) et (**ii**) sont clairement satisfaites et on a alors la fonction valeur

$$v(x) = J_0^\alpha = U(x - Y_0).$$

Pour satisfaire la condition (**iii**), on va exploiter la structure particulière exponentielle de la fonction d'utilité U. En effet, en substituant (5.43), (5.47) dans (5.46) avec U donnée par (5.44), on obtient :

$$J_t^\alpha = M_t^\alpha \, C_t^\alpha,$$

où M^α est la martingale (locale) donnée par :

$$\begin{aligned} M_t^\alpha &= \exp(-\eta(x - Y_0)) \\ &\quad \times \exp\left(- \int_0^t \eta(\alpha_u \sigma_u - Z_u) dW_u - \frac{1}{2} \int_0^t |\eta(\alpha_u \sigma_u - Z_u)|^2 du \right), \end{aligned}$$

et

$$C_t^\alpha = - \exp\Big(\int_0^t \rho(u, \alpha_u, Z_u) du \Big),$$

avec

$$\rho(t, a, z) = \eta \left(\frac{\eta}{2} |a\sigma_t - z|^2 - a b_t - f(t, z) \right).$$

Pour l'obtention de la condition (**iii**), on va donc chercher un générateur f tel que le processus (C_t^α) soit décroissant pour tout $\alpha \in \mathcal{A}$ et constant pour un $\hat{\alpha} \in \mathcal{A}$. Autrement dit, cela revient à déterminer f tel que :

$$\rho(t, \alpha_t, Z_t) \geq 0, \quad 0 \leq t \leq T, \;\; \forall \alpha \in \mathcal{A} \tag{5.48}$$

et

$$\rho(t, \hat{\alpha}_t, Z_t) = 0, \quad 0 \leq t \leq T. \tag{5.49}$$

En réecrivant ρ sous la forme,

$$\frac{1}{\eta} \rho(t, a, z) = \frac{\eta}{2} \left| a\sigma_t - z - \frac{1}{\eta} \frac{b_t}{\sigma_t} \right|^2 - z \frac{b_t}{\sigma_t} - \frac{1}{2\eta} \left| \frac{b_t}{\sigma_t} \right|^2 - f(t, z),$$

on voit clairement que les conditions (5.48) et (5.49) seront vérifiées avec

$$f(t, z) = -z \frac{b_t}{\sigma_t} - \frac{1}{2\eta} \left| \frac{b_t}{\sigma_t} \right|^2, \tag{5.50}$$

et

$$\hat{\alpha}_t = \frac{1}{\sigma_t} \left(Z_t + \frac{1}{\eta} \frac{b_t}{\sigma_t} \right), \quad 0 \leq t \leq T. \tag{5.51}$$

Théorème 5.5.8 *La fonction valeur du problème* (5.45) *est donnée par*

$$v(x) = U(x - Y_0) \; = \; - \exp(-\eta(x - Y_0)),$$

où (Y, Z) *est l'unique solution de l'EDSR*

$$-dY_t = f(t, Z_t) dt - Z_t dW_t, \;\; Y_T \; = \; \xi, \tag{5.52}$$

avec un générateur f *donné par* (5.50). *Le contrôle optimal* $\hat{\alpha}$ *est donné par* (5.51).

Preuve. Au vu des arguments établis ci-dessus, il reste à vérifier rigoureusement la condition (**iii**) sur J^α. Notons que puisque b/σ est bornée, le générateur $f(t,z)$ satisfait une condition de Lipschitz, et donc à fortiori de croissance quadratique en z, uniformément en (t,ω). De plus, comme ξ est supposée bornée, on admettra alors (voir Kobylanski [Ko00]) que la solution de l'EDSR (5.52) est telle que Y est aussi bornée.

Pour tout $\alpha \in \mathcal{A}$, le processus M^α est une martingale locale et il existe donc une suite de temps d'arrêts (τ_n), $\tau_n \to +\infty$ p.s., tel que $(M^\alpha_{t\wedge\tau_n})$ soit une martingale (positive). De plus avec le choix de f en (5.50), le processus C^α est décroissant et donc $(J^\alpha_{t\wedge\tau_n}) = (M^\alpha_{t\wedge\tau_n} C^\alpha_{t\wedge\tau_n})$ est une surmartingale. Comme $X^{x,\alpha}$ est borné inférieurement et Y est borné, le processus J^α, donné par (5.46), est borné inférieurement. Par le lemme de Fatou, on en déduit que J^α est une surmartingale.

Finalement, pour le choix de $\hat{\alpha}$ donné en (5.51), on a :

$$J^{\hat{\alpha}}_t = M^{\hat{\alpha}}_t \;=\; \exp(-\eta(x - Y_0))\; \exp\left(- \int_0^t \frac{b_u}{\sigma_u} dW_u - \frac{1}{2}\int_0^t \left|\frac{b_u}{\sigma_u}\right|^2 du \right).$$

Comme b/σ est borné, on conclut que $J^{\hat{\alpha}}$ est une martingale. □

Remarque 5.5.2 Le modèle financier présenté dans cet exemple est un modèle de marché complet : tout actif contingent ξ, $\mathcal{F}_T$-mesurable et borné, est parfaitement couvert par la richesse d'un portefeuille autofinançant. Autrement dit, il existe $\pi \in \mathcal{A}$ tel que $\xi = X_T^{x_\xi,\pi}$ où x_ξ est le prix d'arbitrage de ξ donné par $x_\xi = E^Q[\xi]$ et Q est l'unique probabilité équivalent à P rendant martingale (locale) le prix S, et appelé aussi probabilité risque-neutre. Le problème (5.45) peut donc se formuler comme :

$$v(x) = \sup_{\alpha\in\mathcal{A}} E[U(X_T^{x - x_\xi, \alpha - \pi})].$$

On est donc ramené à un problème de maximisation d'utilité exponentielle sans actif contingent. Ainsi, la stratégie optimale (5.51) du problème original se décompose en la somme $\alpha_t = \pi_t + \alpha^0_t$ de la stratégie de couverture $\pi_t = Z_t/\sigma_t$ de l'actif contingent ξ et la stratégie optimale $\alpha^0_t = \frac{1}{\eta} b_t/\sigma_t^2$ du problème de maximisation d'utilité exponentielle sans actif contingent.

Dans un contexte plus général de marché incomplet, i.e. où l'actif contingent ξ n'est pas parfaitement couvert, la même démarche **(i)**, **(ii)**, **(iii)**, s'applique mais conduit à un générateur f plus complexe faisant intervenir un terme quadratique en z, voir El Karoui et Rouge [ElkR00].

5.5.2 Critère moyenne-variance d'allocation de portefeuille

On considère un modèle financier de Black Scholes. Il y a un actif sans risque de processus de prix

$$dS_t^0 = rS_t^0 dt$$

et un actif risqué de processus de prix

$$dS_t = S_t(bdt + \sigma dW_t),$$

avec des constantes $b > r$ et $\sigma > 0$. Un agent investit un montant α_t dans l'actif risqué et son processus de richesse évolue alors selon :

$$\begin{aligned} dX_t &= \alpha_t \frac{dS_t}{S_t} + (X_t - \alpha_t)\frac{dS_t^0}{S_t^0} \\ &= [rX_t + \alpha_t(b-r)]\, dt + \sigma\alpha_t dW_t, \quad X_0 = x. \end{aligned} \tag{5.53}$$

On note par $\mathcal{A}$ l'ensemble des processus de contrôle α progressifs à valeurs dans $\mathbb{R}$, tel que $E[\int_0^T |\alpha_t|^2 dt] < +\infty$.

Le critère moyenne-variance d'allocation de portefeuille consiste à minimiser la variance du portefeuille sous contrainte que l'espérance soit égale à une constante donnée :

$$V(m) = \inf_{\alpha\in\mathcal{A}} \left\{ \mathrm{Var}(X_T) \ : \ E(X_T) = m \right\}, \quad m \in \mathbb{R}. \tag{5.54}$$

Nous verrons dans la Proposition 5.5.3, par la méthode du Lagrangien, que cela revient à résoudre le problème de contrôle auxiliaire dual

$$\tilde{V}(\lambda) = \inf_{\alpha\in\mathcal{A}} E[X_T - \lambda]^2, \quad \lambda \in \mathbb{R}. \tag{5.55}$$

Nous allons résoudre ce problème (5.55) par le principe du maximum stochastique décrit au paragraphe 5.4.2. Dans ce cas, l'Hamiltonien (5.28) a la forme :

$$\mathcal{H}(x,a,y,z) = [rx + a(b-r)]\, y + \sigma a z.$$

L'EDSR adjointe (5.29) s'écrit pour $\alpha \in \mathcal{A}$:

$$-dY_t = rY_t dt - Z_t dW_t, \quad Y_T \ = \ 2(X_T - \lambda). \tag{5.56}$$

Soit $\hat{\alpha} \in \mathcal{A}$ un candidat pour le contrôle optimal et $\hat{X}$, $(\hat{Y}, \hat{Z})$ les processus associés. Alors

$$\mathcal{H}(x,a,\hat{Y}_t,\hat{Z}_t) = rx\hat{Y}_t + a\left[(b-r)\hat{Y}_t + \sigma\hat{Z}_t\right].$$

Cette expression étant linéaire en a, on voit donc que les conditions (5.30) et (5.31) seront satisfaites si

$$(b-r)\hat{Y}_t + \sigma\hat{Z}_t = 0, \quad 0 \le t \le T, \ p.s. \tag{5.57}$$

On cherche $(\hat{Y}, \hat{Z})$ solution de (5.56) de la forme

$$\hat{Y}_t = \varphi(t)\hat{X}_t + \psi(t), \tag{5.58}$$

où φ et ψ sont deux fonctions déterministes C^1. En substituant dans (5.56) et en utilisant l'expression (5.53), on voit que φ, ψ et $\hat{\alpha}$ doivent satisfaire :

$$\varphi'(t)\hat{X}_t + \varphi(t)(r\hat{X}_t + \hat{\alpha}_t(b-r)) + \psi'(t) = -r(\varphi(t)\hat{X}_t + \psi(t)), \tag{5.59}$$
$$\varphi(t)\sigma\hat{\alpha}_t = \hat{Z}_t \tag{5.60}$$

et les conditions terminales

$$\varphi(T) \;=\; 2, \;\; \psi(T) \;=\; -2\lambda. \tag{5.61}$$

En utilisant les relations (5.57), (5.58) et (5.60), on obtient l'expression de $\hat{\alpha}$:

$$\hat{\alpha}_t = \frac{(r-b)\hat{Y}_t}{\sigma^2\varphi(t)} \;=\; \frac{(r-b)(\varphi(t)\hat{X}_t + \psi(t))}{\sigma^2\varphi(t)}. \tag{5.62}$$

D'autre part, d'après (5.59), on a :

$$\hat{\alpha}_t = \frac{(\varphi'(t) + 2r\varphi(t))\hat{X}_t + \psi'(t) + r\psi(t)}{(r-b)\varphi(t)}. \tag{5.63}$$

En comparant avec (5.62), on obtient les équations différentielles satisfaites par φ et ψ :

$$\varphi'(t) + \left(2r - \frac{(b-r)^2}{\sigma^2}\right)\varphi(t) = 0, \quad \varphi(T) \;=\; 2 \tag{5.64}$$
$$\psi'(t) + \left(r - \frac{(b-r)^2}{\sigma^2}\right)\psi(t) = 0, \quad \varphi(T) \;=\; -2\lambda, \tag{5.65}$$

dont les solutions explicites sont (seul $\psi = \psi_\lambda$ dépend de λ)

$$\varphi(t) = 2\exp\left[\left(2r - \frac{(b-r)^2}{\sigma^2}\right)(T-t)\right] \tag{5.66}$$
$$\psi_\lambda(t) = \lambda\psi_1(t) \;=\; -2\lambda\exp\left[\left(r - \frac{(b-r)^2}{\sigma^2}\right)(T-t)\right] \tag{5.67}$$

Avec ce choix de φ, ψ_λ, les processus $(\hat{Y}, \hat{Z})$ résolvent l'EDSR adjointe (5.56) et les conditions du principe du maximum stochastique (Théorème 5.4.6) sont vérifiées : le contrôle optimal est donné par (5.62) ou encore sous forme Markovienne par :

$$\hat{\alpha}_\lambda(t,x) = \frac{(r-b)(\varphi(t)x + \psi_\lambda(t))}{\sigma^2\varphi(t)}. \tag{5.68}$$

Pour calculer la fonction valeur $\tilde{V}(\lambda)$, on procède comme suit. Pour tout $\alpha \in \mathcal{A}$, on applique la formule d'Itô à $\frac{1}{2}\varphi(t)X_t^2 + \psi_\lambda(t)X_t$ entre 0 et T, en utilisant la dynamique (5.53) de X et les EDO (5.64)-(5.65) satisfaites par φ et ψ_λ. On obtient alors en prenant l'espérance :

$$\begin{aligned} E\left[X_T-\lambda\right]^2 &= \frac{1}{2}\varphi(0)x^2+\psi_\lambda(0)x+\lambda^2 \\ &\quad + E\left[\int_0^T \frac{\varphi(t)\sigma^2}{2}\left(\alpha_t-\frac{(r-b)(\varphi(t)X_t+\psi_\lambda(t))}{\sigma^2\varphi(t)}\right)^2 dt\right] \\ &\quad - \frac{1}{2}\int_0^T\left(\frac{b-r}{\sigma}\right)^2\frac{\psi_\lambda(t)^2}{\varphi(t)}dt. \end{aligned}$$

Ceci montre de nouveau que le contrôle optimal est donné par (5.62) et que la fonction valeur est :

$$\tilde{V}(\lambda)=\frac{1}{2}\varphi(0)x^2+\psi_\lambda(0)x+\lambda^2-\frac{1}{2}\int_0^T\left(\frac{b-r}{\sigma}\right)^2\frac{\psi_\lambda(t)^2}{\varphi(t)}dt,$$

soit avec les expressions explicites (5.66)-(5.67) de φ et ψ_λ :

$$\tilde{V}(\lambda)=e^{-\frac{(b-r)^2}{\sigma^2}T}(\lambda-e^{rT}x)^2,\quad \lambda\in\mathbb{R}. \tag{5.69}$$

Nous montrons maintenant comment les deux problèmes (5.54) et (5.55) sont liés.

Proposition 5.5.3 *On a les relations de conjugaison :*

$$\tilde{V}(\lambda)=\inf_{m\in\mathbb{R}}\left[V(m)+(m-\lambda)^2\right],\quad \lambda\in\mathbb{R}, \tag{5.70}$$

$$V(m)=\sup_{\lambda\in\mathbb{R}}\left[\tilde{V}(\lambda)-(m-\lambda)^2\right],\quad m\in\mathbb{R}. \tag{5.71}$$

Pour tout m dans $\mathbb{R}$, le contrôle optimal de $V(m)$ est égal à $\hat{\alpha}_{\lambda_m}$ donné par (5.68) *où λ_m atteint l'argument maximum dans* (5.71)*, soit :*

$$\lambda_m=\frac{m-\exp\left[\left(r-\frac{(b-r)^2}{\sigma^2}\right)T\right]x}{1-\exp\left[-\frac{(b-r)^2}{\sigma^2}T\right]}. \tag{5.72}$$

Preuve. Notons d'abord que pour tout $\alpha\in\mathcal{A}$, $\lambda\in\mathbb{R}$, on a :

$$E[X_T-\lambda]^2=\mathrm{Var}(X_T)+(E(X_T)-\lambda)^2. \tag{5.73}$$

Pour tout $m\in\mathbb{R}$, pour tout $\varepsilon>0$, on peut trouver $\alpha^\varepsilon\in\mathcal{A}$ de diffusion associée X^ε, tel que $E(X_T^\varepsilon)=m$ et $\mathrm{Var}(X_T^\varepsilon)\leq V(m)+\varepsilon$. On en déduit avec (5.73) que

$$E[X_T^\varepsilon-\lambda]^2\leq V(m)+(m-\lambda)^2+\varepsilon,$$

et donc

$$\tilde{V}(\lambda) \leq V(m) + (m-\lambda)^2, \quad \forall\, m,\ \lambda \in \mathbb{R}. \tag{5.74}$$

D'autre part, pour $\lambda \in \mathbb{R}$, soit $\hat{\alpha}_\lambda \in \mathcal{A}$ de diffusion associée $\hat{X}^\lambda$, un contrôle optimal de $\tilde{V}(\lambda)$. Posons $m_\lambda = E(\hat{X}_T^\lambda)$. Alors d'après (5.73), on a :

$$\begin{aligned}\tilde{V}(\lambda) &= \mathrm{Var}(\hat{X}_T^\lambda) + (m_\lambda - \lambda)^2 \\ &\geq V(m_\lambda) + (m_\lambda - \lambda)^2.\end{aligned}$$

Cette dernière inégalité combinée avec (5.74) prouve (5.70) :

$$\begin{aligned}\tilde{V}(\lambda) &= \inf_{m \in \mathbb{R}} \left[V(m) + (m-\lambda)^2\right] \\ &= V(m_\lambda) + (m_\lambda - \lambda)^2,\end{aligned}$$

et aussi que $\hat{\alpha}_\lambda$ est la solution de $V(m_\lambda)$.

On vérifie facilement que la fonction V est convexe en m. En reécrivant la relation (5.70) sous la forme $(\lambda^2 - \tilde{V}(\lambda))/2 = \sup_m[m\lambda - (V(m)+m^2)/2]$, on obtient que la fonction $\lambda \to (\lambda^2 - \tilde{V}(\lambda))/2$ est la transformée polaire (ou de Fenchel-Legendre) de la fonction convexe $m \to (V(m)+m^2)/2$. On a alors la relation de conjugaison $(V(m)+m^2)/2 = \sup_\lambda[m\lambda - (\lambda^2 - \tilde{V}(\lambda))/2]$, ce qui donne (5.71).

Finalement, pour tout $m \in \mathbb{R}$, soit $\lambda_m \in \mathbb{R}$ l'argument maximum de $V(m)$ dans (5.71) qui est explicitement donné par (5.72) d'après l'expression (5.69) de $\tilde{V}$. Alors m est un argument minimum de $\tilde{V}(\lambda_m)$ dans (5.70). Comme la fonction $m \to V(m) + (m-\lambda)^2$ est strictement convexe, cet argument minimum est unique et donc $m = m_{\lambda_m} = E(\hat{X}_T^{\lambda_m})$. On a donc

$$\begin{aligned}V(m) &= \tilde{V}(\lambda_m) + (m-\lambda_m)^2 \\ &= E\left[\hat{X}_T^{\lambda_m} - \lambda_m\right]^2 + \left[E(\hat{X}_T^{\lambda_m}) - \lambda_m\right]^2 \ = \ \mathrm{Var}(\hat{X}_T^{\lambda_m}),\end{aligned}$$

ce qui prouve que $\hat{\alpha}_{\lambda_m}$ est solution de $V(m)$. □

Remarque 5.5.3 Il y a une interprétation financière claire de la stratégie de portefeuille optimale (5.68) du problème (5.55). En effet, notons qu'elle s'écrit aussi comme

$$\hat{\alpha}_t^{(\lambda)} \ := \ \hat{\alpha}_\lambda(t, X_t) = -\frac{b-r}{\sigma^2}(X_t - R_\lambda(t)), \quad 0 \leq t \leq T,$$

où le processus (ici déterministe) $R_\lambda(t) = -\psi_\lambda(t)/\varphi(t)$ est explicitement déterminé par :

$$dR_\lambda(t) = rR_\lambda(t)dt, \qquad R_\lambda(T) = \lambda.$$

R_λ est donc le processus de richesse du portefeuille d'investissement nul dans l'actif risqué, et répliquant parfaitement l'actif contingent constant λ. D'autre part, considérons le problème d'un investisseur de richesse autofinançante $\bar{X}_t$ et cherchant à minimiser $E[(\bar{X}_T)^2]$ dans ce modèle de marché complet. Sa stratégie optimale de portefeuille est le portefeuille de Merton pour une fonction d'utilité $U(x) = -x^2$, et est donnée par

$$\bar{\alpha}_t = -\frac{b-r}{\sigma^2}\bar{X}_t, \quad 0 \leq t \leq T. \tag{5.75}$$

La stratégie optimale du problème (5.55) est donc la stratégie selon (5.75) de portefeuille de richesse $X_t - R_\lambda(t)$, et aurait pu ainsi être obtenue plus directement avec cette remarque. Nous avons seulement voulu illustré sur cet exemple simple de marché complet, comment appliquer l'approche par principe du maximum stochastique. En fait, cette approche s'utilise avec succès pour traiter le cas plus complexe de coefficients aléatoires sur les prix et de marchés incomplets, et conduit à des équations différentielles stochastiques rétrogrades pour $\varphi(t)$ et $\psi_\lambda(t)$, voir e.g. Kohlmann et Zhou [KZ00].

5.6 Commentaires bibliographiques

Les EDSR ont été introduites dans le cas linéaire par Bismut [Bis76] comme l'équation adjointe associée à la version stochastique du principe du maximum de Pontryagine en théorie du contrôle. Le cas général non linéaire d'existence et d'unicité de solution d'EDSR a été résolu dans l'article de référence de Pardoux et Peng [PaPe90]. Il y a eu ensuite de nombreuses extensions portant sur les hypothèses sur le générateur. Citons notamment l'article de Kobylanski [Ko00] qui montre l'existence de solution bornée pour un générateur à croissance quadratique en z. Ce résultat est fort utile dans de nombreuses applications en finance. Pour d'autres extensions, on se réferrera au livre édité par El Karoui et Mazliak [ElkM97] ou à celui de Ma et Yong [MY00]. Le lien entre les EDSR et les EDP non linéaires et leur représentation par des formules de type Feynman-Kac est étudié plus en détail dans l'article de Pardoux [Pa98]. Les applications des EDSR au contrôle et aux mathématiques financières ont été étudiées par El Karoui, Peng et Quenez [ElkPQ97]. La présentation du paragraphe 5.4.1 est très largement due à cet article. D'autres applications des EDSR pour le contrôle sont étudiées dans Hamadène et Lepeltier [HL95]. Le théorème de vérification suffisante du principe du maximum et la relation avec la programmation dynamique énoncés au paragraphe 5.4.2 sont traités dans le livre de Yong et Zhou [YZ00]. L'utilisation des EDSR pour la résolution du problème de maximisation d'utilité exponentielle avec actif contingent a été étudiée par El Karoui et Rouge [ElkR00], voir aussi les articles de Sekine [Se02] et Hu, Imkeller, Müller [HIM04] pour des fonctions d'utilité puissance. Les applications des EDSR pour les problèmes de couverture moyenne-variance et plus généralement pour des problèmes avec état linéaire et coût quadratique ont été initiées par Bismut [Bis78] et étendues dans les articles de Kohlmann et Zhou [KZ00], Zhou et Li [ZL00] ou Kohlmann et Tang [KT02].

6

Méthodes martingales de dualité convexe

6.1 Introduction

Dans les méthodes d'optimisation par la programmation dynamique ou par les équations différentielles stochastiques rétrogrades vues aux chapitres précédents, l'optimisation portait essentiellement sur le processus de contrôle α agissant sur la dynamique du processus d'état X. L'idée générale et formelle des méthodes martingales de dualité est de ramener de façon équivalente l'optimisation sur la variable d'état contrôlée grâce à une représentation linéaire sous forme d'espérance pondérée par une variable dite duale. Illustrons cette idée sur un exemple. Considérons une variable d'état X, contrôlée par un processus α progressif à valeurs réelles et vérifiant $\int_0^T |\alpha_t|^2 dt < +\infty$:

$$dX_t = \alpha_t(dt + dW_t), \quad 0 \leq t \leq T,$$

où W est un mouvement Brownien standard sur $(\Omega, \mathcal{F}, \mathbb{F}, P)$. On suppose que $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ est la filtration naturelle de W. Pour $x \in \mathbb{R}_+$ et α contrôle, on note X^x la solution de l'EDS ci-dessus partant de x en $t = 0$ et $\mathcal{A}(x)$ l'ensemble des processus de contrôle α tel que $X_t^x \geq 0$, $0 \leq t \leq T$. Etant donné une fonction de gain croissante et concave g de $\mathbb{R}_+$ dans $\mathbb{R}$, on veut résoudre

$$v(x) = \sup_{\alpha \in \mathcal{A}(x)} E[g(X_T^x)], \quad x \geq 0. \tag{6.1}$$

Introduisons la probabilité $Q \sim P$ qui fait du processus $B_t = W_t + t$ un mouvement Brownien, par le théorème de Girsanov. Alors d'après le théorème de représentation d'Itô sous Q, pour toute variable aléatoire positive X_T, $\mathcal{F}_T$-mesurable, i.e. $X_T \in L_+^0(\Omega, \mathcal{F}_T, P)$ satisfaisant la contrainte $E^Q[X_T] \leq x$, il existe $\alpha \in \mathcal{A}(x)$ tel que :

$$X_T = E^Q[X_T] + \int_0^T \alpha_t dB_t \ \leq \ X_T^x \ = \ x + \int_0^T \alpha_t dB_t.$$

Réciproquement pour tout $\alpha \in \mathcal{A}(x)$, le processus $X^x = x + \int \alpha dB$ est une Q-martingale locale positive donc une Q surmartingale et on a $E^Q[X_T^x] \leq x$. On en déduit que le problème d'optimisation (6.1) se reformule de façon équivalente comme :

$$v(x) = \sup_{X_T \in L^0_+(\Omega, \mathcal{F}_T, P)} E[g(X_T)] \quad \text{sous la contrainte } E\left[\frac{dQ}{dP} X_T\right] \leq x. \tag{6.2}$$

Ainsi, on est ramené à un problème d'optimisation concave dans $L^0_+(\Omega, \mathcal{F}_T, P)$ sous contrainte linéaire représentée par la variable duale dQ/dP. On peut alors appliquer les méthodes d'analyse et d'optimisation convexe pour résoudre (6.2).

L'outil clé dans l'approche de résolution duale du problème d'optimisation ci-dessus est le célèbre théorème de représentation d'Itô qui est aussi l'argument central dans la réplication d'actifs contingents en marché complet. L'extension de cette approche à des problèmes d'optimisation plus généraux, est basée sur un puissant théorème d'analyse stochastique, appelé théorème de décomposition optionnelle des surmartinagles. Ce récent théorème a été motivé initialement par le problème de la surréplication en marché incomplet et a été établi initialement dans le cadre de processus d'Itô par El Karoui et Quenez [ElkQ95]. Il a été ensuite généralisé dans un cadre très général de processus semimartingale. Nous décrivons ce théorème de décomposition optionnelle dans le paragraphe suivant. Losque le problème initial de contrôle est transformé en un problème dit primal d'optimisation convexe sous contraintes linéaires, on peut essayer de le résoudre par des méthodes d'analyse convexe. Cela conduit à la formulation et résolution d'un problème dual issu du Lagrangien sur le problème primal contraint. Nous détaillons cette méthode de résolution duale pour le problème de maximisation d'utilité de la richesse terminale. Il est à noter que l'approche duale permet d'obtenir des résultats d'existence et de caractérisation dans un cadre général de processus de prix semimartingale, alors que l'approche par programmation dynamique et équation d'Hamilton-Jacobi-Bellman nécessite de se placer dans un cadre Markovien.

Dans la dernière partie de ce chapitre, nous étudions le problème de couverture moyenne-variance d'actif contingents. Il se formule comme un problème de projection dans L^2 d'une variable aléatoire de carré intégrable sur un espace d'intégrales stochastiques. Nous verrons comment résoudre ce problème d'optimisation quadratique en combinant un théorème de projection de Kunita-Watanabe, une approche par dualité et une méthode de changement de numéraire.

6.2 Représentation duale du problème de surréplication

6.2.1 Formulation du problème de surréplication

Soit S une semimartingale continue à valeurs dans $\mathbb{R}^n$ sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ satisfaisant les conditions habituelles. Pour simplifier, on suppose $\mathcal{F} = \mathcal{F}_T$ et $\mathcal{F}_0$ trivial, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$. On considère ici un horizon fini T. S représente le processus de prix actualisé de n actifs risqués. On note $L(S)$ l'ensemble des processus progressifs, intégrable par rapport à S. Un élément $\alpha \in L(S)$ représente une stratégie de contrôle de portefeuille d'un investisseur : α_t est le nombre d'unités investi dans l'actif risqué à la date t. Ainsi, partant d'un capital initial $x \in \mathbb{R}$, le processus de richesse de l'investisseur qui utilise le contrôle α est :

$$x + \int_0^t \alpha_s dS_s, \quad 0 \leq t \leq T.$$

Un processus de portefeuille $\alpha \in L(S)$ est dit admissible si $\int \alpha dS$ est borné inférieurement et on notera $\mathcal{A}(S)$ l'ensemble de tels processus. Cela signifie économiquement que l'investisseur n'est pas autorisé à avoir un découvert infini. Cette condition d'admissibilité empêche en fait les stratégies d'arbitrage doubling (voir Harrison-Pliska [HP81]) : en effet, on pourrait construire une suite de stratégies de portefeuille $(\alpha^n)_{n \geq 1} \in L(S)$ tel que $\int_0^T \alpha_t^n dS_t \to +\infty$ p.s., ce qui représente un moyen de gagner autant d'argent que l'on veut en T à partir d'un capital nul !

On se donne un actif contingent de maturité T, représenté par une variable aléatoire $\mathcal{F}_T$-mesurable $X_T \geq 0$. Le problème de la surréplication de X_T consiste à déterminer le capital initial minimum qui permet de surcouvrir sans risque à la maturité T de l'actif contingent, son flux X_T par une stratégie de portefeuille admissible. Mathématiquement, le problème se formule ainsi :

$$v_0 = \inf \left\{ x \in \mathbb{R} \; : \exists \alpha \in \mathcal{A}(S), \; x + \int_0^T \alpha_t dS_t \; \geq \; X_T \; p.s. \right\}. \qquad (6.3)$$

v_0 est appelé coût de surréplication de X_T.

On note $L_+^0(\Omega, \mathcal{F}_T, P)$ l'ensemble des variables aléatoires $\mathcal{F}_T$-mesurables et positives p.s. Pour tout $x \in \mathbb{R}$, on considèrera aussi l'ensemble

$$\mathcal{C}(x) = \left\{ X_T \in L_+^0(\Omega, \mathcal{F}_T, P) \; : \exists \alpha \in \mathcal{A}(S), x + \int_0^T \alpha_t dS_t \geq X_T \, p.s. \right\}. \qquad (6.4)$$

$\mathcal{C}(x)$ représente l'ensemble des actifs contingents qui peuvent être surcouverts sans risque à partir d'un capital initial x et d'une stratégie de portefeuille admissible.

L'objet de cette section est de donner une représentation et caractérisation de v_0 et $\mathcal{C}(x)$ en termes d'un certain ensemble dual de probabilités.

6.2.2 Probabilités martingales et arbitrage

On définit

$$\mathcal{M}_e(S) = \{Q \sim P \ \text{ sur } (\Omega, \mathcal{F}_T) \ : \ S \ \text{ est une } Q - \text{martingale locale}\}.$$

$\mathcal{M}_e(S)$ est appelé ensemble des probabilités martingales ou risque-neutre.

Dans toute la suite du chapitre, on fera l'hypothèse cruciale :

$$\mathcal{M}_e(S) \neq \emptyset. \tag{6.5}$$

Cette hypothèse sera présente dans tous les énoncés de propositions ou théorèmes et ne sera pas rappelée. Elle est équivalente à une condition de non-arbitrage et ce résultat, appelé premier théorème fondamental de la finance, est développé en profondeur dans l'article de Delbaen et Schachermayer [DS94]. Signalons ici simplement que pour tout $Q \in \mathcal{M}_e(S)$ et $\alpha \in \mathcal{A}(S)$, l'intégrale stochastique bornée inférieurement $\int \alpha dS$ est une Q-martingale locale et donc aussi par le lemme de Fatou, une Q-surmartingale. On a ainsi $E^Q[\int_0^T \alpha_t dS_t] \leq 0$. L'hypothèse (6.5) implique donc :

$$\nexists \ \alpha \in \mathcal{A}(S), \quad \int_0^T \alpha_t dS_t \geq 0, \ p.s. \quad \text{et} \quad P\left[\int_0^T \alpha_t dS_t > 0\right] > 0.$$

Autrement dit, on ne peut pas trouver de stratégie de portefeuille admissible qui permette, partant d'un capital initial nul, d'atteindre p.s. une richesse terminale positive, et avec une probabilité non nulle d'être strictement positive. C'est la condition économique de non arbitrage.

6.2.3 Le théorème de décomposition optionnelle et la représentation duale du coût de surréplication

Le problème de la surréplication a en fait motivé un très joli résultat d'analyse stochastique que nous énonçons en toute généralité dans le cas de processus S semimartingale continue. Nous donnerons une preuve de ce théorème dit de décomposition optionnelle dans un cadre particulier au paragraphe suivant.

Théorème 6.2.1 *Soit X un processus positif qui est une surmartingale cadlag sous toute probabilité $Q \in \mathcal{M}_e(S) \neq \emptyset$. Alors il existe $\alpha \in L(S)$ et C processus adapté croissant, $C_0 = 0$, tel que :*

$$X = X_0 + \int \alpha dS - C. \tag{6.6}$$

Remarque 6.2.1 Rappelons que dans la décomposition de Doob-Meyer d'une surmartingale X comme différence d'une martingale locale M et d'un processus croissant C : $X = M - C$, le processus C peut être choisi prévisible et dans ce cas la décomposition est unique. La décomposition (6.6) est universelle au sens où le processus positif $M = X_0 + \int \alpha dS$ est une martingale locale pour tout $Q \in \mathcal{M}_e(S)$. De plus, le processus C n'est en général pas prévisible, mais seulement optionnel, et il n'est pas unique.

Examinons à présent comment ce théorème permet de donner une représentation duale du coût de surréplication d'un actif contingent $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. Pour cela, il suffit de considérer une modification cad-lag du processus

$$X_t = \text{ess} \sup_{Q \in \mathcal{M}_e(S)} E^Q \left[X_T \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T, \tag{6.7}$$

(il n'y a pas d'ambiguité de notation à la date T dans la relation précédente, on a bien $X_T = X_T$!), dont on vérifiera ci-dessous que c'est une surmartingale sous tout $Q \in \mathcal{M}_e(S)$, et de lui appliquer le théorème de décomposition optionnelle.

Théorème 6.2.2 *Soit $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. Alors son coût de surréplication est égal à*

$$v_0 = \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T], \tag{6.8}$$

De plus si $\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < +\infty$, i.e. v_0 est fini, alors v_0 atteint l'infimum dans (6.3) *avec une stratégie de surréplication donnée par la décomposition optionnelle* (6.6) *du processus X défini en* (6.7).

Preuve. Notons que pour tout $\alpha \in \mathcal{A}(S)$ et $Q \in \mathcal{M}_e(S)$, l'intégrale stochastique bornée inférieurement $\int \alpha dS$ est une Q-martingale locale, donc une Q-surmartingale. On en déduit que pour tout $x \in \mathbb{R}$ tel que $x + \int_0^T \alpha_t dS_t \geq X_T$ p.s. avec $\alpha \in \mathcal{A}(S)$, on a $E^Q[X_T] \leq x$ pour tout $Q \in \mathcal{M}_e(S)$. Ceci implique par définition de v_0 :

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq v_0. \tag{6.9}$$

Si $\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] = +\infty$, l'égalité (6.8) est donc évidente. On suppose désormais que

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < +\infty. \tag{6.10}$$

1. Montrons d'abord que le processus $(X_t)_{0 \leq t \leq T}$ défini en (6.7) est une surmartingale sous tout $Q \in \mathcal{M}_e(S)$ et qu'il admet une modification cad-lag. Considérons la famille de processus adaptés $\{\Gamma^Q_t : 0 \leq t \leq T,\ Q \in \mathcal{M}_e(S)\}$ où

$$\Gamma^Q_t = E^Q \left[X_T \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T, \quad Q \in \mathcal{M}_e(S),$$

est bien défini d'après (6.10).

(i) Nous vérifions que pour tout $t \in [0, T]$, l'ensemble $\{\Gamma^Q_t : Q \in \mathcal{M}_e(S)\}$ est stable par suprémum, i.e. pour tout $Q_1, Q_2 \in \mathcal{M}_e(S)$, il existe $Q \in \mathcal{M}_e(S)$ tel que $\max(\Gamma^{Q_1}_t, \Gamma^{Q_2}_t) = \Gamma^Q_t$. Pour cela, fixons un élément $Q^0 \in \mathcal{M}_e(S)$ de processus de densité martingale Z^0 et définissons le processus :

$$Z_s = \begin{cases} Z_s^0, & s \leq t \\ Z_t^0 \left(\frac{Z_s^1}{Z_t^1} 1_A + \frac{Z_s^2}{Z_t^2} 1_{\Omega \setminus A} \right), & t < s \leq T, \end{cases}$$

où Z^1 (resp. Z^2) est le processus de densité martingale de Q^1 (resp. Q^2), $A = \{\omega \ : \Gamma_t^{Q_1}(\omega) \geq \Gamma_t^{Q_2}(\omega)\} \in \mathcal{F}_t$. En jouant avec la loi des espérances conditionnelles itérées, il est facile de voir que Z hérite de Z^0, Z^1 et Z^2 la propriété de martingale sous P. De plus, comme Z est strictement positif avec $Z_0 = 1$, on peut lui associer une probabilité $Q \sim P$ telle que Z soit son processus de densité martingale. Par définition de $\mathcal{M}_e(S)$ et par la formule de Bayes, les processus $Z^1 S$ et $Z^2 S$ sont des martingales locales sous P, et alors ZS hérite aussi de cette propriété de martingale locale. Ainsi $Q \in \mathcal{M}_e(S)$. De plus, on a :

$$\begin{aligned} \Gamma_t^Q = E^Q[X_T|\mathcal{F}_t] \ &= \ E\left[\frac{Z_T}{Z_t} X_T \middle| \mathcal{F}_t \right] \\ &= E\left[\frac{Z_T^1}{Z_t^1} X_T 1_A + \frac{Z_T^2}{Z_t^2} X_T 1_{\Omega\setminus A} \middle| \mathcal{F}_t \right] \\ &= 1_A E^{Q^1}[X_T|\mathcal{F}_t] + 1_{\Omega\setminus A} E^{Q^2}[X_T|\mathcal{F}_t] \\ &= 1_A \Gamma_t^{Q_1} + 1_{\Omega\setminus A}\Gamma_t^{Q_2} \ = \ \max(\Gamma_t^{Q_1}, \Gamma_t^{Q_2}), \end{aligned}$$

qui est la propriété de stabilité par supremum. Pour tout $t \in [0, T]$, il existe alors une suite $(Q_k^t)_{k\geq 1}$ de $\mathcal{M}_e(S)$ telle que :

$$X_t := \text{ess} \sup_{Q\in\mathcal{M}_e(S)} \Gamma_t^Q \ = \ \lim_{k\to+\infty} \uparrow \Gamma_t^{Q_k^t}, \tag{6.11}$$

(le symbole $\lim_{k\to+\infty} \uparrow$ signifiant que la limite est croissante, i.e. $\Gamma_t^{Q_k^t} \leq \Gamma_t^{Q_{k+1}^t}$.)

(ii) Montrons alors la propriété universelle de surmartingale. Soit Q_0 quelconque dans $\mathcal{M}_e(S)$ de processus de densité martingale Z^0 et fixons $0 \leq u < t \leq T$. Notons $(Q_k^t)_{k\geq 1}$ la suite donnée par (6.11) et $(Z^{k,t})_{k\geq 1}$ la suite associée des processus de densité martingale. Notons alors que pour tout $k \geq 1$, le processus défini par

$$\tilde{Z}_s^{k,t} = \begin{cases} Z_s^0, & s \leq t \\ Z_t^0 \frac{Z_s^{k,t}}{Z_t^{k,t}}, & t < s \leq T, \end{cases}$$

est une martingale (sous P) strictement positive de valeur initiale $\tilde{Z}_0^{k,t} = 1$, et est donc associée à une probabilité $\tilde{Q}_k^t \sim P$. De plus, $\tilde{Z}^{k,t} S$ est une martingale locale sous P et donc $\tilde{Q}_k^t \in \mathcal{M}_e(S)$. On a alors pour tout $k \geq 1$,

$$E^{Q_0}[\Gamma_t^{Q_k^t}|\mathcal{F}_u] = E\left[\frac{Z_t^0}{Z_u^0} \Gamma_t^{Q_k^t} \middle| \mathcal{F}_u \right] \ = \ E\left[\frac{Z_t^0}{Z_u^0} E\left[\frac{Z_T^{k,t}}{Z_t^{k,t}} X_T \middle| \mathcal{F}_t \right] \middle| \mathcal{F}_u \right]$$

$$= E\left[\frac{Z_t^0}{Z_u^0}\frac{Z_T^{k,t}}{Z_t^{k,t}}X_T\middle|\mathcal{F}_u\right] = E\left[\frac{\tilde{Z}_T^{k,t}}{\tilde{Z}_u^{k,t}}X_T\middle|\mathcal{F}_u\right]$$
$$= E^{\tilde{Q}_k^t}[X_T|\mathcal{F}_u] = \Gamma_u^{\tilde{Q}_k^t}.$$

D'après (6.11), on en déduit par le théorème de convergence monotone :

$$\begin{aligned} E^{Q_0}[X_t|\mathcal{F}_u] &= \lim_{k\to+\infty} \uparrow E^{Q_0}[\Gamma_t^{Q_k}|\mathcal{F}_u] = \lim_{k\to+\infty} \uparrow \Gamma_u^{\tilde{Q}_k^t} \\ &\leq \operatorname{ess}\sup_{Q\in\mathcal{M}_e(S)} \Gamma_u^Q = X_u, \end{aligned} \tag{6.12}$$

ce qui prouve que X est une Q_0-surmartingale.

(iii) Il reste à vérifier que X admet une modification cad-lag qui vérifie encore la propriété de surmartingale pour tout $Q^0 \in \mathcal{M}_e(S)$. On sait d'après le théorème 1.1.7 que c'est le cas si la fonction $t \to E^{Q^0}[X_t]$ est continue à droite. D'après (6.12) avec $u = 0$, on a :

$$E^{Q^0}[X_t] = \lim_{k\to+\infty} \uparrow E^{\tilde{Q}_k^t}[X_T], \quad \forall t \in \mathbb{T}. \tag{6.13}$$

Soit t fixé dans $\mathbb{T}$ et $(t_n)_{n\geq1}$ une suite dans $\mathbb{T}$ qui décroit vers t. Comme X est une Q^0-surmatingale, on a :

$$\lim_{n\to+\infty} E^{Q^0}[X_{t_n}] \leq E^{Q^0}[X_t].$$

D'autre part, pour tout $\varepsilon > 0$, il existe d'après (6.13), $\hat{k} = \hat{k}(\varepsilon) \geq 1$ tel que :

$$E^{Q^0}[X_t] \leq E^{\tilde{Q}_{\hat{k}}^t}[X_T] + \varepsilon. \tag{6.14}$$

Notons que $\tilde{Z}_T^{\hat{k},t_n}$, la densité de Radon-Nikodym de $\tilde{Q}_{\hat{k}}^{t_n}$ tend p.s. vers $\tilde{Z}_T^{\hat{k},t}$, la densité de Radon-Nikodym de $\tilde{Q}_{\hat{k}}^t$ quand n tend vers l'infini. D'après le lemme de Fatou, on en déduit avec (6.14) :

$$\begin{aligned} E^{Q^0}[X_t] &\leq \lim_{n\to+\infty} E^{\tilde{Q}_{\hat{k}}^{t_n}}[X_T] + \varepsilon \\ &\leq \lim_{n\to+\infty} E^{Q^0}[X_{t_n}] + \varepsilon \end{aligned}$$

où la deuxième inégalité vient de (6.13). Puisque ε est arbitraire, ceci prouve finalement que $\lim_{n\to+\infty} E^{Q^0}[X_{t_n}] = E^{Q^0}[X_t]$, i.e. la continuité à droite de $(E^{Q^0}[X_t])_{t\in\mathbb{T}}$.

2. On peut donc appliquer le théorème de décomposition optionnelle à la modification cad-lag encore notée X et obtenir l'existence d'un processus $\hat{\alpha} \in L(S)$ et d'un processus adapté C croissant, $C_0 = 0$ tels que :

$$X_t = X_0 + \int_0^t \hat{\alpha}_s dS_s - C_t, \quad p.s., \quad 0 \leq t \leq T. \tag{6.15}$$

Comme X et C sont positifs, cette dernière relation montre que $\int \hat{\alpha} dS$ est borné inférieurement par $-X_0$ et donc $\hat{\alpha} \in \mathcal{A}(S)$. De plus, cette relation (6.15) pour $t = T$ implique que :

$$X_T \leq X_0 + \int_0^T \hat{\alpha}_s dS_s, \quad p.s.$$

Ceci prouve par définition de v_0 que :

$$v_0 \leq X_0 \ = \ \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T].$$

On conclut la preuve avec (6.9). □

Grâce à cette représentation duale du coût de surréplication, on en déduit immédiatement la caractérisation suivante des ensembles $\mathcal{C}(x)$ définis en (6.4).

Corollaire 6.2.1 *Pour tout $x \in \mathbb{R}$, on a*

$$\mathcal{C}(x) = \left\{ X_T \in L^0_+(\Omega, \mathcal{F}_T, P) \ : \ \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq x \right\}. \tag{6.16}$$

En particulier, $\mathcal{C}(x)$ est fermé pour la topologie de la convergence en mesure, i.e. si $(X^n)_{n \geq 1}$ est une suite de $\mathcal{C}(x)$ convergent p.s. vers $\hat{X}_T$, alors $\hat{X}_T \in \mathcal{C}(x)$.

On a ainsi une caractérisation très pratique de $\mathcal{C}(x)$: pour savoir si un actif contingent peut être surcouvert sans risque à partir d'un capital initial x, il faut et il suffit de tester si son espérance sous toute probabilité martingale est inférieure à x. Mathématiquement, cette caractérisation est à la base de la résolution par dualité du problème d'optimisation de portefeuille. De plus, on obtient immédiatement avec cette caractérisation la fermeture de $\mathcal{C}(x)$ dans $L^0_+(\Omega, \mathcal{F}_T, P)$, ce qui n'est clairement pas évident d'après sa définition originale (primale) (6.4).

6.2.4 Le cadre de processus d'Itô et de filtration Brownienne

On se place dans le cadre où le processus de prix $S = (S^1, \ldots, S^n)$ suit la dynamique :

$$dS_t = \mu_t dt + \sigma_t dW_t, \quad 0 \ \leq t \leq T, \tag{6.17}$$

où W est un mouvement Brownien standard d-dimensionnel sur $(\Omega, \mathcal{F}, \mathbb{F}, P)$ avec $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ filtration naturelle de W, $d \geq n$, μ est un processus progressif à valeurs dans $\mathbb{R}^n$ tel que $\int_0^T |\mu_t|dt < +\infty$ p.s., σ est un processus progressif à valeurs dans $\mathbb{R}^{n\times d}$ tel que $\int_0^T |\sigma_t|^2 dt < +\infty$ p.s. On suppose que pour tout $t \in [0,T]$, la matrice σ_t est de rang plein égal à n, p.s. La matrice carrée $n \times n$, $\sigma_t\sigma_t'$, est donc inversible et on définit le processus progressif à valeurs dans $\mathbb{R}^d$:

$$\lambda_t = \sigma_t'(\sigma_t\sigma_t')^{-1}\mu_t, \quad 0 \le t \le T.$$

Pour simplifier (voir Remarque 6.2.4), on supposera que le processus λ est borné.

Remarque 6.2.2 Dans la littérature, afin de garantir un processus de prix positif, on modélise souvent sa dynamique d'Itô sous la forme :

$$dS_t = \mathrm{diag}(S_t)\left(\tilde{\mu}_t dt + \tilde{\sigma}_t dW_t\right), \quad 0 \ \le t \le T, \tag{6.18}$$

où $\mathrm{diag(S_t)}$ désigne la matrice diagonale $n \times n$ avec éléments diagonaux S_t^i. Le modèle de Black-Scholes et les modèles à volatilité stochastique vus au chapitres précédents sont des exemples particuliers de (6.18). Notons que le modèle (6.17) englobe celui de (6.18) en posant

$$\mu_t \ = \ \mathrm{diag}(S_t)\tilde{\mu}_t, \ \ \sigma_t \ = \ \mathrm{diag}(S_t)\tilde{\sigma}_t.$$

Dans un premier temps, nous allons donner dans ce cadre une description explicite de l'ensemble $\mathcal{M}_e(S)$ des probabilités martingales. On considère l'ensemble

$$K(\sigma) = \left\{\nu \in L^2_{loc}(W) \ : \sigma\nu \ = \ 0, \quad \text{sur} \ \ [0,T]\times\Omega, \ \ dt\times dP \ p.p.\right\}.$$

Pour tout $\nu \in K(\sigma)$, on définit la martingale locale exponentielle

$$Z_t^\nu = \exp\left(-\int_0^t (\lambda_u + \nu_u)' dW_u - \frac{1}{2}\int_0^t |\lambda_u + \nu_u|^2 du\right), \quad 0 \le t \le T.$$

On définit alors l'ensemble

$$K_m(\sigma) = \{\nu \in K(\sigma) \ : Z^\nu \quad \text{est une martingale}\}.$$

Remarque 6.2.3 Rappelons (voir chapitre 1, paragraphe 1.2.5) qu'une condition suffisante assurant que Z^ν soit une martingale, i.e. $E[Z_T^\nu] = 1$, est le critère de Novikov :

$$E\left[\exp\left(\frac{1}{2}\int_0^T |\lambda_u|^2 + |\nu_u|^2 du\right)\right] < +\infty. \tag{6.19}$$

(Notons ici que puisque λ et ν sont orthogonaux, i.e. $\lambda'\nu = 0$, alors $|\lambda+\nu|^2 = |\lambda|^2 + |\nu|^2$.)

Pour tout $\nu \in K_m(\sigma)$, on peut donc définir une probabilité $P^\nu \sim P$ de processus de densité martingale Z^ν. Rappelons aussi par le théorème de Girsanov que le processus

$$W^\nu = W + \int \lambda + \nu \, dt$$

est un mouvement Brownien sous P^ν.

Grâce au théorème de Girsanov et à celui de représentation d'Itô des martingales Browniennes, on a la caractérisation explicite suivante de $\mathcal{M}_e(S)$.

Proposition 6.2.1 *On a*

$$\mathcal{M}_e(S) = \{P^\nu \ : \ \nu \in K_m(\sigma)\} .$$

Preuve. (i) Puisque par définition, on a $\sigma\lambda = \mu$, et pour tout $\nu \in K_m(\sigma)$, $\sigma\nu = 0$, il s'en suit que la dynamique de S sous P^ν s'écrit :

$$dS_t = \sigma_t dW_t^\nu. \tag{6.20}$$

Ceci prouve que S est une P^ν martingale locale, i.e. $P^\nu \in \mathcal{M}_e(S)$.

(ii) Réciproquement, soit $Q \in \mathcal{M}_e(S)$ et Z son processus (strictement positif) de densité de martingale. D'après le théorème de représentation d'Itô, il existe un processus $\rho \in L^2_{loc}(W)$ tel que :

$$Z_t = \exp\left(- \int_0^t \rho'_u dW_u - \frac{1}{2} \int_0^t |\rho_u|^2 du \right), \quad 0 \leq t \leq T.$$

De plus, par le théorème de Girsanov, le processus

$$B^\rho = W + \int \rho \, dt$$

est un mouvement Brownien sous Q. La dynamique de S sous Q s'écrit donc

$$dS_t = (\mu_t - \sigma_t \rho_t) dt + \sigma_t dB_t^\rho, \quad 0 \ \leq t \leq T.$$

Comme S est une martingale locale sous Q, on doit avoir

$$\sigma\rho = \mu, \quad \text{sur } \ [0,T] \times \Omega, \ \ dt \times dP \ p.p.$$

En posant $\nu = \rho - \lambda$ et puisque $\sigma\lambda = \mu$, ceci montre que $\sigma\nu = 0$ et donc que $\nu \in K(\sigma)$. On a donc finalement que $Z = Z^\nu$ martingale, i.e. $\nu \in K_m(\sigma)$, et ainsi $Q = P^\nu$. □

Remarque 6.2.4 **1.** La partie (i) de la preuve ci-dessus prouve que l'inclusion $\{P^\nu \ : \ \nu \in K_m(\sigma)\} \subset \mathcal{M}_e(S)$ est toujours vraie même sans l'hypothèse de filtration Brownienne.

2. Puisque λ est supposée borné, on voit que la condition de Novikov (6.19) est satisfaite pour tout processus ν borné. En fait, c'est vrai dès que λ vérifie lui même la condition de Novikov : $E\left[\exp\left(\frac{1}{2}\int_0^T |\lambda_u|^2 du\right)\right] < +\infty$. En particulier, le processus nul $\nu = 0$ est dans $K_m(\sigma)$. La probabilité martingale associée P^0 est appelée probabilité martingale minimale selon la terminologie de Föllmer-Schweizer.

3. La remarque ci-dessus montre aussi en particulier que dès lors que λ satisfait la condition de Novikov, $\mathcal{M}_e(S)$ est non vide et contient P^0. Dans le cas où Z^0 n'est pas une martingale, l'hypothèse $\mathcal{M}_e(S) \neq \emptyset$ n'est pas forcément satisfaite et se ramène à l'existence d'un élément ν dans $K_m(\sigma)$.

Nous donnons à présent une démonstration du théorème de décomposition optionnelle. En fait dans ce cadre de processus de prix d'Itô avec filtration Brownienne, le processus C qui apparait dans la décomposition est prévisible.

Théorème 6.2.3 *Soit X une surmartingale positive cad-lag sous toute probabilité martingale P^ν, $\nu \in K_m(\sigma)$. Alors X admet une décomposition sous la forme*

$$X = X_0 + \int \alpha dS - C$$

où $\alpha \in L(S)$ et C est un processus croissant prévisible, $C_0 = 0$.

Preuve. D'après le théorème de décomposition de Doob-Meyer appliqué à la surmartingale positive X sous P^ν pour tout $\nu \in K_m(\sigma)$, on a :

$$X_t = X_0 + M_t^\nu - A_t^\nu, \quad 0 \leq t \leq T,$$

où M^ν est une martingale locale sous P^ν avec $M_0^\nu = 0$, et A^ν est un processus prévisible, intégrable (sous P^ν) et croissant avec $A_0^\nu = 0$. Par le théorème de représentation des martingales Browniennes sous P^ν, il existe alors $\psi^\nu \in L^2_{loc}(W^\nu)$ tel que

$$X_t = X_0 + \int_0^t (\psi_u^\nu)' dW_u^\nu - A_t^\nu, \quad 0 \leq t \leq T. \tag{6.21}$$

On fixe un élément dans $\mathcal{M}_e(S)$, par exemple pour simplifier P^0, et on compare les décompositions (6.21) de X sous P^ν et sous P^0. En notant que $W^\nu = W^0 + \int \nu dt$, et en identifiant les parties martingales locales et les parties à variation finie, on obtient p.s. :

$$\psi_t^\nu = \psi_t^0, \quad 0 \leq t \leq T, \tag{6.22}$$

$$A_t^\nu - \int_0^t \nu_u' \psi_u^\nu du = A_t^0, \quad 0 \leq t \leq T, \tag{6.23}$$

pour tout $\nu \in K_m(\sigma)$.

On définit alors le processus progressif α à valeurs dans $\mathbb{R}^n$ par

$$\alpha_t = (\sigma_t \sigma_t')^{-1} \sigma_t \psi_t^0, \quad 0 \leq t \leq T.$$

On voit que $\int_0^T |\alpha_t' \mu_t| dt = \int_0^T |\lambda_t' \psi_t^0| dt < +\infty$ et $\int_0^T |\alpha_t' \sigma_t|^2 dt = \int_0^T |\psi_t^0|^2 dt < \infty$ p.s. et donc $\alpha \in L(S)$. En posant $\eta_t = \psi_t^0 - \sigma_t' \alpha_t$, on a que $\int_0^T |\eta_t|^2 dt < +\infty$ p.s. et $\sigma\eta = 0$. Autrement dit, $\eta \in K(\sigma)$. En fait, on a écrit la décomposition de ψ^0 sur $\mathrm{Im}(\sigma')$ et son orthogonal $K(\sigma)$:

$$\psi_t^0 = \sigma_t' \alpha_t + \eta_t, \quad 0 \leq t \leq T. \tag{6.24}$$

On va montrer que $\eta = 0$ en utilisant (6.22)-(6.23). Pour cela, considérons le processus

$$\tilde{\nu}_t = -n \frac{\eta_t}{|\eta_t|} 1_{\eta_t \neq 0}, \quad 0 \leq t \leq T,$$

ou n est un entier non nul. Alors $\tilde{\nu}$ est borné et est donc dans $K_m(\sigma)$. D'après (6.22)-(6.23) pour $\tilde{\nu}$ et en utilisant aussi (6.24), on obtient :

$$A_T^{\tilde{\nu}} = A_T^0 - n \int_0^T |\eta_t| 1_{\eta_t \neq 0} dt.$$

Comme $E^{P^0}[A_T^0] < +\infty$ et $E^{P^0}[A_T^{\tilde{\nu}}] \geq 0$, on voit en prenant l'espérance sous P^0 et faisant tendre n vers l'infini dans la relation ci-dessus qu'on doit avoir

$$\eta = 0, \quad \text{sur } [0,T] \times \Omega, \quad dt \times dP^0 \ \ p.p.$$

La décomposition (6.21) de X sous P^0 s'écrit alors en se rappelant la dynamique (6.20) de S sous P^0 et en posant $C = A^0$:

$$\begin{aligned} X &= X_0 + \int \alpha' \sigma dW^0 - A^0 \\ &= X_0 + \int \alpha dS - C \end{aligned}$$

et la preuve est terminée. □

6.3 Dualité pour la maximisation d'utilité

6.3.1 Formulation du problème d'optimisation de portefeuille

Dans le cadre du modèle de marché financier décrit au paragraphe 6.2.1, nous formulons maintenant le problème d'optimisation de portefeuille par critère d'utilité.

On se donne une fonction $U(x)$ modélisant l'utilité d'un agent ayant une richesse x et on fait les hypothèses classiques suivantes sur U. La fonction U :

$\mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ est continue sur son domaine $\mathrm{dom}(U) = \{x \in \mathbb{R} \ : U(x) > -\infty\}$, dérivable, strictement croissante et strictement concave sur l'intérieur de son domaine. Sans perte de généralité, quite à rajouter une constante, on peut supposer que $U(+\infty) > 0$. Une telle fonction vérifiant ses hypothèses sera appelée fonction d'utilité. On se placera dans le cas où

$$\mathrm{int}(\mathrm{dom}(U)) =]0, +\infty[\tag{6.25}$$

ce qui signifie que les richesses négatives ne sont pas autorisées.

Le problème de maximisation d'utilité de richesse terminale se formule alors ainsi :

$$v(x) = \sup_{\alpha \in \mathcal{A}(S)} E\left[U\left(x + \int_0^T \alpha_t dS_t\right)\right], \quad x > 0. \tag{6.26}$$

Finalement, pour exclure les cas triviaux, on suppose que la fonction valeur v n'est pas dégénérée :

$$v(x) < +\infty, \quad \text{pour un } x > 0. \tag{6.27}$$

En fait, d'après les propriétés de croissance et de concavité de U sur son domaine, qui se transmettent à v, cette hypothèse (6.27) est équivalente à

$$v(x) < +\infty, \quad \text{pour tout } x > 0. \tag{6.28}$$

6.3.2 Résultat général d'existence

Dans cette section, on montre directement l'existence d'une solution au problème de maximisation d'utilité (6.26).

Observons d'abord que puisque $U(x) = -\infty$ pour $x < 0$, il suffit de considérer dans le supremum dans (6.26) les $\alpha \in \mathcal{A}(S)$ qui conduisent à une richesse terminale positive $x + \int_0^T \alpha_t dS_t \geq 0$ p.s. De plus, d'après la (stricte) croissance de U sur $]0, +\infty[$, il est clair que

$$v(x) = \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)], \quad x > 0, \tag{6.29}$$

où on rappelle que $\mathcal{C}(x)$ a été défini dans (6.4). Il est clair aussi que si $\hat{X}_T^x \in \mathcal{C}(x)$ est une solution de (6.29) alors il existe $\hat{\alpha} \in \mathcal{A}(S)$ tel que $\hat{X}_T^x = x + \int_0^T \hat{\alpha}_t dS_t$ et $\hat{\alpha}$ est solution de (6.26).

Nous allons donc montrer l'existence d'une solution à (6.29) grâce à la caractérisation duale de $\mathcal{C}(x)$ et en fait à sa propriété de fermeture dans $L^0_+(\Omega, \mathcal{F}_T, P)$, voir Corollaire 6.2.1. L'idée est de considérer une suite maximisante $(X^n)_{n\geq 1}$ de (6.29), d'utiliser un théorème de compacité dans

$L^0_+(\Omega, \mathcal{F}_T, P)$ qui permet à une combinaison convexe près d'obtenir une limite p.s. $\hat{X}_T$ de X^n, puis de passer à la limite dans $E[U(X^n)]$. Le point technique est d'avoir de l'uniforme intégrabilité sur la suite $U_+(X^n)$. On fait alors l'hypothèse suivante :

$$\limsup_{x \to +\infty} \frac{v(x)}{x} \leq 0. \tag{6.30}$$

Cette condition peut sembler à priori curieuse et difficilement vérifiable en pratique puisqu'elle porte sur la fonction valeur $v(x)$ qu'on cherche à déterminer. En fait, on verra d'une part dans la preuve ci-dessous que c'est exactement la condition nécessaire pour avoir la convergence de $E[U(X^n)]$ vers $E[U(\hat{X}_T)]$. D'autre part, on donnera dans le paragraphe suivant des conditions pratiques portant directement sur U qui garantissent (6.30).

Théorème 6.3.4 *Soit U une fonction d'utilité vérifiant* (6.25), (6.27) *et* (6.30). *Alors pour tout $x > 0$, il existe une unique solution $\hat{X}^x_T$ au problème $v(x)$ en* (6.29).

Preuve. Soit $x > 0$ et $(X^n)_{n \geq 1}$ une suite maximisante dans $\mathcal{C}(x)$ pour $v(x) < +\infty$, i.e. :

$$\lim_{n \to +\infty} E[U(X^n)] = v(x) \ < \ +\infty. \tag{6.31}$$

D'après le théorème de compacité A.3.5 dans $L^0_+(\Omega, \mathcal{F}_T, P)$, on peut trouver une combinaison convexe $\hat{X}^n \in \text{conv}(X^n, X^{n+1}, \ldots)$ qui est encore dans l'ensemble convexe $\mathcal{C}(x)$ et telle que $\hat{X}^n$ converge p.s. vers une variable aléatoire positive $\hat{X}^x_T$. Comme $\mathcal{C}(x)$ est fermé pour la convergence en mesure, on a $\hat{X}^x_T \in \mathcal{C}(x)$. Par concavité de U et d'après (6.31), on a encore

$$\lim_{n \to +\infty} E[U(\hat{X}^n)] = v(x) \ < \ +\infty. \tag{6.32}$$

Notons U^+ et U^- les parties positives et négatives de U et remarquons que d'après (6.32), on a $\sup_n E[U^-(\hat{X}^n)] < +\infty$ et $\sup_n E[U^+(\hat{X}^n)] < +\infty$. D'autre part, avec le lemme de Fatou, on a

$$\liminf_{n \to +\infty} E[U^-(\hat{X}^n)] \geq E[U^-(\hat{X}^x_T)].$$

L'optimalité de $\hat{X}^x_T$, i.e. $v(x) = E[U(\hat{X}^x_T)]$, est alors obtenue ssi on montre que

$$\lim_{n \to +\infty} E[U^+(\hat{X}^n)] = E[U^+(\hat{X}^x_T)], \tag{6.33}$$

i.e. l'uniforme intégrabilité de la suite $(U^+(X^n))_{n \geq 1}$.

Si $U(+\infty) \leq 0$, i.e. $U^+ \equiv 0$, il n'y a rien à prouver. On rappelle que $U(+\infty) > 0$ et on pose

$$x_0 = \inf\{x > 0 \ : U(x) \geq 0\} \ < \ +\infty.$$

On raisonne par l'absurde en supposant au contraire que la suite $(U^+(X^n))_n$ n'est pas uniformément intégrable. Alors il existe $\delta > 0$ tel que

$$\lim_{n\to+\infty} E[U^+(\hat{X}^n)] = E[U^+(\hat{X}^x_T)] + 2\delta.$$

D'après le corollaire A.1.1, en passant éventuellement à une sous-suite encore notée $(\hat{X}^n)_{n\geq 1}$, on peut trouver des ensembles disjoints $(B^n)_{n\geq 1}$ de $(\Omega, \mathcal{F}_T)$ tels que :

$$E[U^+(\hat{X}^n)1_{B^n}] \geq \delta, \quad \forall n \geq 1.$$

On considère alors la suite de variables aléatoires dans $L^0_+(\Omega, \mathcal{F}_T, P)$:

$$H^n = x_0 + \sum_{k=1}^n \hat{X}^k 1_{B^k}.$$

Pour tout $Q \in \mathcal{M}_e(S)$, on a :

$$E^Q[H^n] \leq x_0 + \sum_{k=1}^n E^Q[\hat{X}^k] \ \leq \ x_0 + nx,$$

car $\hat{X}^k \in \mathcal{C}(x)$. La caractérisation (6.16) montre ainsi que $H^n \in \mathcal{C}(x_0 + nx)$. D'autre part, on a :

$$\begin{aligned} E[U(H^n)] &= E\left[U^+\left(x_0 + \sum_{k=1}^n \hat{X}^k 1_{B^k}\right)\right] \\ &\geq E\left[U^+\left(\sum_{k=1}^n \hat{X}^k 1_{B^k}\right)\right] \ = \ \sum_{k=1}^n E[U^+(\hat{X}^k)1_{B^k}] \ \geq \ \delta n. \end{aligned}$$

On en déduit que :

$$\limsup_{x\to+\infty} \frac{v(x)}{x} \geq \limsup_{n\to+\infty} \frac{E[U(H^n)]}{x_0 + nx} \ \geq \ \limsup_{n\to+\infty} \frac{\delta n}{x_0 + nx} \ = \ \delta > 0,$$

ce qui contredit (6.30). Ainsi, on a (6.33) et $\hat{X}^x_T$ est solution de $v(x)$. L'unicité découle de la stricte concavité de U sur $]0, +\infty[$. □

6.3.3 Résolution via la formulation duale

Le problème d'optimisation $v(x)$ en (6.29) se formule comme un problème de maximisation concave en dimension infinie dans $L^0_+(\Omega, \mathcal{F}_T, P)$ avec une infinité de contraintes linéaires données par la caractérisation duale (6.16) de $\mathcal{C}(x)$: pour $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$, on a

$$X_T \in \mathcal{C}(x) \Longleftrightarrow E[Z_T X_T] \leq x, \quad \forall Z_T \in \mathcal{M}_e. \tag{6.34}$$

Ici et dans la suite, on identifie une probabilité $Q \ll P$ avec sa densité de Radon-Nikodym $Z_T = dQ/dP$ et on note pour simplifier $\mathcal{M}_e = \mathcal{M}_e(S)$.

On peut donc essayer d'appliquer à notre contexte les méthodes de dualité pour les problèmes d'optimisation convexe développés dans un cadre abstrait, par exemple dans le livre de Ekeland et Temam [ET74].

Nous commençons par expliquer le principe de cette méthode à notre cadre, puis on soulignera les difficultés qui vont apparaitre et la façon de les surmonter.

On introduit la fonction convexe conjuguée $\tilde{U}$ définie par :

$$\tilde{U}(y) = \sup_{x>0} [U(x) - xy], \quad y > 0, \tag{6.35}$$

et on note $\text{dom}(\tilde{U}) = \{y > 0 \ : \tilde{U}(y) < +\infty\}$. On imposera les conditions économiques usuelles dites d'Inada :

$$U'(0) := \lim_{x \downarrow 0} U'(x) \ = \ +\infty, \ \text{ et } U'(+\infty) := \lim_{x \to +\infty} U'(x) \ = \ 0, \tag{6.36}$$

et on note par $I :]0, +\infty[\to]0, +\infty[$ la fonction inverse de U' sur $]0, +\infty[$, qui est donc strictement décroissante, et vérifie $I(0) = +\infty$, $I(+\infty) = 0$. On rappelle (voir Proposition B.3.5 dans l'appendice B) que sous (6.36), $\text{int}(\text{dom}(\tilde{U}))$ $=]0, +\infty[$, $\tilde{U}$ est dérivable, décroissante, strictement convexe sur $]0, +\infty[$ avec $\tilde{U}(0) = U(+\infty)$ et

$$\tilde{U}' = -(U')^{-1} \ = \ I.$$

De plus, le supremum dans (6.35) est atteint en $x = I(y) > 0$, i.e.

$$\tilde{U}(y) = U(I(y)) - yI(y), \quad y > 0, \tag{6.37}$$

et on a la relation conjuguée :

$$U(x) = \inf_{y>0} \left[\tilde{U}(x) + xy\right], \quad x > 0,$$

avec un infimum atteint en $y = U'(x)$.

Des exemples typiques de fonctions d'utilité vérifiant (6.36), et leurs fonctions conjuguées sont :

$$U(x) \ = \ \ln x, \ \ \tilde{U}(y) \ = \ -\ln y - 1,$$
$$U(x) \ = \ \frac{x^p}{p}, \ p < 1, \ p \neq 0, \ \ \tilde{U}(y) \ = \ \frac{y^{-q}}{q}, \ q \ = \ \frac{p}{1-p}.$$

Le point de départ de l'approche duale est le suivant. Pour tout $x > 0$, $y > 0$, $X_T \in \mathcal{C}(x)$, $Z_T \in \mathcal{M}_e$, on a d'après la définition de $\tilde{U}$ et la caractérisation duale (6.34) de $\mathcal{C}(x)$:

$$\begin{aligned} E[U(X_T)] &\leq E[\tilde{U}(yZ_T)] + E[yZ_TX_T] \\ &\leq E[\tilde{U}(yZ_T)] + xy. \end{aligned} \tag{6.38}$$

On introduit alors le problème dual à $v(x)$:

$$\tilde{v}(y) = \inf_{Z_T \in \mathcal{M}_e} E[\tilde{U}(yZ_T)], \quad y > 0. \tag{6.39}$$

L'inégalité (6.38) montre donc que pour tout $x > 0$:

$$\begin{aligned} v(x) &= \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)] \\ &\leq \inf_{y>0}[\tilde{v}(y) + xy] \ = \ \inf_{y>0, Z_T \in \mathcal{M}_e} \left\{ E[\tilde{U}(yZ_T)] + xy \right\}. \end{aligned} \tag{6.40}$$

La méthode basique de résolution duale pour le problème primal $v(x)$ consiste alors dans les étapes suivantes : montrer l'existence de $(\hat{y}, \hat{Z}_T)$ (dépendant de x) solution du problème dual dans le terme de droite de (6.40). Ceci se ramène aussi à montrer l'existence de $\hat{y} > 0$ atteignant le minimum de $\tilde{v}(y) + xy$, et à montrer l'existence de $\hat{Z}_T$ au problème dual $\tilde{v}(\hat{y})$. On pose alors :

$$\hat{X}_T^x = I(\hat{y}\hat{Z}_T), \quad i.e. \quad U'(\hat{X}_T^x) \ = \ \hat{y}\hat{Z}_T$$

D'après les conditions d'optimalité du premier ordre sur $\hat{y}$ et $\hat{Z}_T$, on verra que ceci implique :

$$\hat{X}_T^x \ \in \ \mathcal{C}(x) \ \text{ et } \ E[\hat{Z}_T\hat{X}_T^x] \ = \ x.$$

D'après (6.37), on aura donc :

$$E[U(\hat{X}_T^x)] = E\left[\tilde{U}\left(\hat{y}\hat{Z}_T\right)\right] + x\hat{y},$$

ce qui prouvera, en se rappelant (6.40), que

$$v(x) = E[U(\hat{X}_T^x)], \quad i.e. \ \ \hat{X}_T^x \text{ est solution de } v(x).$$

et que la relation de conjugaison sur les fonctions valeurs primal et dual est valide :

$$v(x) = \inf_{y>0}[\tilde{v}(y) + xy] \ = \ \tilde{v}(\hat{y}) + x\hat{y}.$$

Avant d'évoquer les difficultés soulevées dans cette approche duale, on peut déjà à ce niveau donner des conditions suffisantes présentes dans certains travaux sur le sujet pour que l'hypothèse (6.30) soit valable, garantissant ainsi l'existence d'une solution au problème primal $v(x)$.

Remarque 6.3.5 Supposons que

$$\tilde{v}(y) < +\infty, \quad \forall y > 0. \tag{6.41}$$

Alors l'inégalité (6.40) montre immédiatement que la condition (6.30) est vraie. La condition (6.41) est évidemment satisfaite dès que

$$\forall y > 0, \ \exists Z_T \in \mathcal{M}_e \ : \ E\left[\tilde{U}\left(yZ_T\right)\right] \ < \ +\infty. \tag{6.42}$$

Une condition portant directement sur U et assurant (6.42) est : il existe $p \in]0,1[$, des constantes positives k_1, k_2 et $Z_T \in \mathcal{M}_e$ tels que

$$U^+(x) \leq k_1 x^p + k_2, \quad \forall x > 0, \tag{6.43}$$

$$E\left[Z_T^{-q}\right] < +\infty, \quad \text{où } \ q \ = \ \frac{p}{1-p} > 0. \tag{6.44}$$

En effet dans ce cas, on a

$$\tilde{U}(y) \leq \sup_{x>0}[k_1 x^p - xy] + k_2 \ = \ (k_1 p)^{\frac{1}{1-p}} \ \frac{y^{-q}}{q} + k_2, \quad \forall y > 0,$$

et (6.42) est alors évident. On verra plus tard une condition plus faible (en fait minimale) sur U assurant (6.41).

Revenons à présent sur l'approche duale décrite formellement ci-dessus. Le point délicat est l'existence au problème dual $\tilde{v}(y)$, $y > 0$. L'ensemble $\mathcal{M}_e$ sur lequel on optimise est naturellement inclus dans $L^1(\Omega, \mathcal{F}_T, P)$ et il n'y a pas de théorème de compacité dans L^1. On dispose bien du théorème de Komlos qui affirme que de toute suite $(Z^n)_n$ bornée dans L^1, on peut en trouver une combinaison convexe qui converge p.s. vers une variable aléatoire $\hat{Z} \in L^1$. Mais cette convergence n'est en général bien sûr pas vraie dans L^1. Au regard de notre problème, la suite maximisante pour $\tilde{v}(y)$ de probabilités $(Z^n)_{n\geq 1}$ de $\mathcal{M}_e$ (qui vérifie $E[Z^n] = 1$) ne converge pas forcément vers une probabilité $\hat{Z}_T$: en général, on a $E[\hat{Z}_T] < 1$. En fait, l'espace $L^0_+(\Omega, \mathcal{F}_T, P)$ dans lequel évolue les variables primales $X_T \in \mathcal{C}(x)$ n'est pas en dualité adéquate avec $L^1(\Omega, \mathcal{F}_T, P)$. Il est plus naturel de permettre aux variables duales de varier aussi dans $L^0_+(\Omega, \mathcal{F}_T, P)$.

On "agrandit" donc dans $L^0_+(\Omega, \mathcal{F}_T, P)$ l'ensemble $\mathcal{M}_e$ de la manière suivante. On définit $\mathcal{D}$ comme l'enveloppe convexe, solide et fermée de $\mathcal{M}_e$ dans $L^0_+(\Omega, \mathcal{F}_T, P)$, i.e. le plus petit ensemble convexe, solide et fermée dans $L^0_+(\Omega, \mathcal{F}_T, P)$ et contenant $\mathcal{M}_e$. Rappelons qu'un ensemble $\mathcal{S}$ de $L^0_+(\Omega, \mathcal{F}_T, P)$ est dit solide si : $Y'_T \leq Y_T$ p.s. et $Y_T \in \mathcal{S}$ implique que $Y'_T \in \mathcal{S}$. Il est facile de voir que $\mathcal{D}$ s'écrit comme :

$$\mathcal{D} = \left\{ Y_T \in L^0_+(\Omega, \mathcal{F}_T, P) \ : \exists \ (Z^n)_{n\geq 1} \in \mathcal{M}_e, \ Y_T \ \leq \ \lim_{n\to+\infty} Z^n \right\},$$

où la limite $\lim_{n\to+\infty} Z^n$ est comprise au sens de la convergence p.s. D'après (6.34) et le lemme de Fatou, on en déduit que l'ensemble $\mathcal{C}(x)$ s'écrit aussi en polarité avec $\mathcal{D}$: pour $X_T \in L^0_+(\Omega, \mathcal{F}_T)$,

$$X_T \in \mathcal{C}(x) \Longleftrightarrow E[Y_T X_T] \leq x, \quad \forall Y_T \in \mathcal{D}. \tag{6.45}$$

On définit alors le problème dual

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E[\tilde{U}(yY_T)], \quad y > 0. \tag{6.46}$$

On verra plus bas que cette définition est cohérente avec celle en (6.39), i.e. l'infimum dans la formule précédente est identique prise dans $\mathcal{M}_e$ ou $\mathcal{D}$.

On impose enfin une condition dite d'élasticité asymptotique raisonnable :

$$AE(U) \;:=\; \limsup_{x\to+\infty} \frac{xU'(x)}{U(x)} < 1. \tag{6.47}$$

Des exemples (et contre-exemples) typiques de telles fonctions d'utilité sont :

- $U(x) = \ln x$, pour lequel $AE(U) = 0$
- $U(x) = \frac{x^p}{p}$, $p < 1$, $p \neq 0$, pour lequel $AE(U) = p$.
- $U(x) = \frac{x}{\ln x}$, pour x suffisamment grand, pour lequel $AE(U) = 1$.

Le théorème suivant montre que sous ces hypothèses sur la fonction d'utilité U, la théorie de la dualité "marche" bien dans ce contexte. En fait, l'hypothèse d'élasticité asymptotique raisonnable est minimale et ne peut être relachée au sens où on peut trouver des contre-exemples de processus de prix continus S pour lequel la fonction valeur $\tilde{v}(y)$ n'est pas finie pour tout y et il n'existe pas de solution au problème primal $v(x)$, dès lors que $AE(U) = 1$. (voir Kramkov et Schachermayer [KS99]).

Théorème 6.3.5 *Soit U une fonction d'utilité vérifiant* (6.25), (6.27), (6.36) *et* (6.47). *Alors*

(1) La fonction v est finie, dérivable, strictement concave sur $]0, +\infty[$, *et il existe une unique solution $\hat{X}^x_T \in \mathcal{C}(x)$ à $v(x)$ pour tout $x > 0$.*

(2) La fonction $\tilde{v}$ est finie, dérivable, strictement convexe sur $]0, +\infty[$, *et il existe une unique solution $\hat{Y}^y_T \in \mathcal{D}$ à $\tilde{v}(y)$ pour tout $y > 0$.*

(3) (i) Pour tout $x > 0$, on a

$$\hat{X}^x_T = I(\hat{y}\hat{Y}_T), \;\; i.e. \;\; U'(\hat{X}^x_T) \;=\; \hat{y}\hat{Y}_T, \tag{6.48}$$

où $\hat{Y}_T \in \mathcal{D}$ est la solution de $\tilde{v}(\hat{y})$ avec $\hat{y} = v'(x)$ l'unique solution de $\mathrm{argmin}_{y>0}[\tilde{v}(y) + xy]$, *et vérifiant :*

$$E[\hat{Y}_T \hat{X}^x_T] = x. \tag{6.49}$$

(ii) On a les relations de conjugaison

$$\begin{aligned} v(x) &= \min_{y>0}[\tilde{v}(y) + xy], \quad \forall x > 0, \\ \tilde{v}(y) &= \max_{x>0}[v(x) - xy], \quad \forall y > 0. \end{aligned}$$

(4) Si de plus il existe $y > 0$ *tel que* $\inf_{Z_T \in \mathcal{M}_e} E[\tilde{U}(yZ_T)] < +\infty$, *alors*

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E[\tilde{U}(yY_T)] \; = \; \inf_{Z_T \in \mathcal{M}_e} E[\tilde{U}(yZ_T)].$$

Remarque 6.3.6 Notons par $\hat{X}_t^x = x + \int_0^t \hat{\alpha}_u dS_u$, $0 \leq t \leq T$, le processus de richesse optimale associée à $v(x)$. (Il n'y a pas d'ambiguité de notation à la date T puisque $x + \int_0^T \hat{\alpha}_u dS_u = \hat{X}_T^x = I(\hat{y}\hat{Y}_T)$ est bien la solution dans $\mathcal{C}(x)$ de $v(x)$). De manière générale, le processus $\hat{X}^x$ est égal à (la modifications cad-lag de)

$$\hat{X}_t^x = \text{ess} \sup_{Q \in \mathcal{M}_e} E^Q \left[I(\hat{y}\hat{Y}_T) \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T.$$

et le contrôle $\hat{\alpha}$ est déterminé par la décomposition optionnelle associée. Dans le cas où le problème dual $\tilde{v}(\hat{y})$ admet une solution $\hat{Z}_T$ dans $\mathcal{M}_e$ de probabilité associée $\hat{Q}$, alors le processus $\hat{X}^x$ est une $\hat{Q}$ martingale locale positive donc $\hat{Q}$-surmartingale telle que $E^{\hat{Q}}[\hat{X}_T^x] = x$ d'après (6.49). $\hat{X}^x$ est une donc $\hat{Q}$-martingale qui s'écrit comme :

$$\hat{X}_t^x = E^{\hat{Q}} \left[I(\hat{y}\hat{Z}_T) \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T.$$

Remarque 6.3.7 L'assertion (4) peut être démontrée sans supposer que $\inf_{Z_T \in \mathcal{M}_e} E[\tilde{U}(yZ_T)] < +\infty$ (voir Proposition 3.2 dans Kramkov et Schachermayer [KS99]). On donne ici une preuve plus simple due à Bouchard et Mazliak [BM03].

Le reste de ce paragraphe est dévolue à la preuve du théorème 6.3.5. Celle-ci sera découpée en plusieurs propositions et lemmes où nous mettrons en évidence les hypothèses requises à chaque étape.

Lemme 6.3.9 *Soit* U *une fonction d'utilité vérifiant* (6.25) *et* (6.36). *Alors pour tout* $y > 0$, *la famille* $\{\tilde{U}^-(yY_T), Y_T \in \mathcal{D}\}$ *est uniformément intégrable.*

Preuve. La fonction $\tilde{U}$ étant (strictement) décroissante, on se place dans le cas où $\tilde{U}(+\infty) = -\infty$ (sinon il n'y a rien à prouver). Soit ϕ l'inverse de la fonction $-\tilde{U}$: ϕ est une fonction strictement croissante de $]-\tilde{U}(0), +\infty[$ dans $]0, +\infty[$. Rappelons que $\tilde{U}(0) = U(+\infty) > 0$ et ϕ est donc bien défini et fini sur $[0, +\infty[$. Pour tout $y > 0$, on a :

$$\begin{aligned} E[\phi(\tilde{U}^-(yY_T))] &\leq E[\phi(\tilde{U}(yY_T))] + \phi(0) \; = \; yE[Y_T] + \phi(0) \\ &\leq y + \phi(0), \qquad \forall Y_T \in \mathcal{D}, \end{aligned}$$

d'après (6.45) car $X_T = 1 \in \mathcal{C}(1)$. De plus, avec un changement de variable trivial et la règle de l'Hôpital et on a d'après (6.36) :

$$\lim_{x\to+\infty} \frac{\phi(x)}{x} = \lim_{y\to+\infty} \frac{y}{-\tilde{U}(y)} = \lim_{y\to+\infty} \frac{1}{I(y)} = +\infty.$$

On conclut avec le théorème de La-Vallée-Poussin. □

Le résultat suivant montre qu'on a bien la relation de conjugaison entre les fonctions valeurs du problème primal et dual.

Proposition 6.3.2 *(Relation de conjugaison)*
Soit U une fonction d'utilité vérifiant (6.25), (6.27) *et* (6.36). *Alors*

$$v(x) = \inf_{y>0}[\tilde{v}(y) + xy], \quad \forall x > 0, \tag{6.50}$$

$$\tilde{v}(y) = \sup_{x>0}[v(x) - xy], \quad \forall y > 0. \tag{6.51}$$

Preuve. Par le même raisonnement que pour (6.38), on a en utilisant (6.45) :

$$\sup_{x>0}[v(x) - xy] \le \tilde{v}(y), \quad \forall y > 0.$$

On fixe $y > 0$ et pour montrer (6.51), on peut donc supposer que $\sup_{x>0}[v(x)-xy] < +\infty$ (sinon il n'y a rien à prouver). Pour tout $n > 0$, on définit l'ensemble

$$\mathcal{B}_n = \left\{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) \ : X_T \le n, \ \ p.s.\right\}.$$

$\mathcal{B}_n$ est compact dans L^∞ pour la topologie faible $\sigma(L^\infty, L^1)$. D'après sa définition, $\mathcal{D}$ est un ensemble convexe et fermé de $L^1(\Omega, \mathcal{F}_T, P)$ et on peut donc appliquer le théorème min-max B.1.2 :

$$\sup_{X_T\in\mathcal{B}_n} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] = \inf_{Y_T\in\mathcal{D}} \sup_{X_T\in\mathcal{B}_n} E[U(X_T) - yX_TY_T], \tag{6.52}$$

pour tout n et $y > 0$. D'après la relation (6.45) de dualité entre $\mathcal{C}(x)$ et $\mathcal{D}$, on a :

$$\begin{aligned}\lim_{n\to+\infty} \sup_{X_T\in\mathcal{B}_n} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] &= \sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} E[U(X_T) - xy] \\ &= \sup_{x>0}[v(x) - xy]. \end{aligned} \tag{6.53}$$

D'autre part, en posant

$$\tilde{U}_n(y) = \sup_{0<x\le n} [U(x) - xy], \quad y > 0,$$

on a

$$\inf_{Y_T\in\mathcal{D}} \sup_{X_T\in\mathcal{B}_n} E[U(X_T) - yX_TY_T] = \inf_{Y_T\in\mathcal{D}} E[\tilde{U}_n(yY_T)] \ := \ \tilde{v}_n(y),$$

de telle sorte que d'après (6.52) et (6.53) :

$$\lim_{n\to+\infty} \tilde{v}_n(y) = \sup_{x>0}[v(x) - xy] \; < \; +\infty.$$

Ainsi, pour obtenir (6.51), on doit montrer

$$\lim_{n\to+\infty} \tilde{v}_n(y) = \tilde{v}(y). \tag{6.54}$$

Clairement $\tilde{v}_n(y)$ est une suite croissante et $\lim_n \tilde{v}_n(y) \leq \tilde{v}(y)$. Soit $(Y^n)_{n\geq 1}$ une suite minimisante dans $\mathcal{D}$ pour $\lim_n \tilde{v}_n(y)$:

$$\lim_{n\to+\infty} E[\tilde{U}_n(yY_T^n)] = \lim_{n\to+\infty} \tilde{v}_n(y) \; < \; +\infty.$$

Par le théorème de compacité A.3.5 dans $L^0_+(\Omega, \mathcal{F}_T, P)$, on peut trouver une combinaison convexe $\hat{Y}^n \in \mathrm{conv}(Y^n, Y^{n+1}, \ldots)$ qui est encore dans l'ensemble convexe $\mathcal{D}$ et qui converge p.s. vers une variable aléatoire positive Y_T. Comme $\mathcal{D}$ est fermée pour la convergence en mesure, on a $Y_T \in \mathcal{D}$. Notons que $\tilde{U}_n(y) = \tilde{U}(y)$ pour $y \geq I(n)$ ($\to 0$ quand n tend vers l'infini). On en déduit d'une part avec le lemme de Fatou :

$$\liminf_{n\to+\infty} E[\tilde{U}_n^+(y\hat{Y}^n)] \geq E[\tilde{U}^+(yY_T)],$$

et d'autre part avec le lemme 6.3.9 :

$$\lim_{n\to+\infty} E[\tilde{U}_n^-(y\hat{Y}^n)] = E[\tilde{U}^-(yY_T)].$$

D'après la convexité de $\tilde{U}_n$, on obtient alors

$$\begin{aligned}\lim_{n\to+\infty} \tilde{v}_n(y) \; = \; \lim_{n\to+\infty} E[\tilde{U}_n(yY^n)] &\geq \liminf_{n\to+\infty} E[\tilde{U}_n(y\hat{Y}^n)] \\ &\geq E[\tilde{U}(yY_T)] \; \geq \; \tilde{v}(y),\end{aligned}$$

ce qui prouve (6.54) et donc (6.51). Sous l'hypothèse (6.27), la relation (6.50) découle alors de la propriété de bipolarité de la transformée de Fenchel-Legendre pour les fonctions convexes (voir Proposition B.3.5 de L'appendice B). □

Remarque 6.3.8 D'après la relation de conjugaison (6.50), on voit que l'hypothèse (6.27) de finitude de v se traduit de manière équivalente en la finitude de $\tilde{v}$ en un point :

$$\exists \text{ (ou } \forall) \; x > 0, \;\; v(x) \; < \; +\infty \Longleftrightarrow \exists y > 0, \;\; \tilde{v}(y) \; < \; +\infty, \;\; \textit{i.e.} \; \mathrm{dom}(\tilde{v}) \neq \emptyset,$$

où

$$\mathrm{dom}(\tilde{v}) = \{y > 0 \;\; : \tilde{v}(y) \; < \; +\infty\}.$$

On verra plus bas qu'avec l'hypothèse supplémentaire (6.47), on aura $\mathrm{dom}(\tilde{v}) =]0, +\infty[$. (et donc (6.30), voir remarque 6.3.5).

Proposition 6.3.3 *(Existence au problème dual)*
Soit U une fonction d'utilité vérifiant (6.25), (6.27) *et* (6.36). *Alors pour tout $y \in$ dom$(\tilde{v})$, il existe une unique solution $\hat{Y}_T^y \in \mathcal{D}$ solution de $\tilde{v}(y)$. En particulier $\tilde{v}$ est strictement convexe sur dom$(\tilde{v})$.*

Preuve. Pour $y \in \mathrm{dom}(\tilde{v})$, soit $(Y^n)_{n\geq 1}$ une suite minimisante dans $\mathcal{D}$ pour $\tilde{v}(y) < +\infty$:

$$\lim_{n\to+\infty} E[\tilde{U}(yY^n)] = \tilde{v}(y).$$

Par le théorème de compacité A.3.5 dans $L^0_+(\Omega, \mathcal{F}_T, P)$, on peut trouver une combinaison convexe $\hat{Y}^n \in \mathrm{conv}(Y^n, Y^{n+1}, \ldots)$ qui est encore dans l'ensemble convexe $\mathcal{D}$ et qui converge p.s. vers une variable aléatoire $\hat{Y}_T^y$. Comme $\mathcal{D}$ est fermée pour la convergence en mesure, on a $\hat{Y}_T^y \in \mathcal{D}$. Comme dans la preuve de la proposition 6.3.2, on a par la convexité de $\tilde{U}$, le lemme de Fatou et le lemme 6.3.9 :

$$\begin{aligned}\tilde{v}(y) \;&=\; \lim_{n\to+\infty} E[\tilde{U}(yY^n)] \geq \liminf_{n\to+\infty} E[\tilde{U}(y\hat{Y}^n)] \\ &\geq E[\tilde{U}(y\hat{Y}_T^y)] \;\geq\; \tilde{v}(y),\end{aligned}$$

ce qui prouve que $\hat{Y}_T^y$ est solution de $\tilde{v}(y)$. L'unicité de la solution découle de la stricte convexité de $\tilde{U}$ qui implique alors aussi que $\tilde{v}$ est strictement convexe sur son domaine dom$(\tilde{v})$. □

Le résultat suivant donne une caractérisation utile de la condition d'élasticité asymptotique raisonnable en termes de U ou $\tilde{U}$.

Lemme 6.3.10 *Soit U une fonction d'utilité vérifiant* (6.25), (6.36). *Alors on a équivalence entre :*
(i) $AE(U) < 1$.
(ii) il existe $x_0 > 0$ et $\gamma \in]0,1[$ tel que

$$xU'(x) < \gamma U(x), \quad \forall x \geq x_0$$

(iii) il existe $x_0 > 0$ et $\gamma \in]0,1[$ tel que

$$U(\lambda x) < \lambda^\gamma U(x), \quad \forall \lambda > 1, \; \forall x \geq x_0$$

(iv) il existe $y_0 > 0$ et $\gamma \in]0,1[$ tel que

$$\tilde{U}(\mu y) < \mu^{-\frac{\gamma}{1-\gamma}} \tilde{U}(y), \quad \forall 0 < \mu < 1, \; \forall 0 < y \leq y_0$$

(v) il existe $y_0 > 0$, $\gamma \in]0,1[$ tel que

$$-y\tilde{U}'(y) < \frac{\gamma}{1-\gamma}\,\tilde{U}(y), \quad \forall 0 < y \leq y_0.$$

Preuve. L'équivalence $(i) \Leftrightarrow (ii)$ est évidente.

$(ii) \Leftrightarrow (iii)$: Fixons $x \geq x_0$ et considérons les fonctions $F(\lambda) = U(\lambda x)$ et $G(\lambda) = \lambda^\gamma U(x)$ pour $\lambda \in [1, +\infty[$. F et G sont dérivables et on a $F(1) = G(1)$. On note que (iii) est équivalent à

$$F(\lambda) < G(\lambda), \quad \forall \lambda > 1,$$

pour tout $x \geq x_0$. Supposons (ii). Alors $F'(1) < G'(1)$ et on en déduit qu'il existe $\varepsilon > 0$ tel que pour tout $\lambda \in]1, 1+\varepsilon]$, on a $F(\lambda) < G(\lambda)$. On va montrer en fait que c'est vrai pour tout $\lambda > 1$. Par l'absurde, cela signifierait que

$$\hat{\lambda} := \inf\{\lambda > 1 \; : F(\lambda) = G(\lambda)\} \; < \; +\infty.$$

En ce point $\hat{\lambda}$, on aurait donc $F'(\hat{\lambda}) \geq G'(\hat{\lambda})$. Or on a d'après (ii) :

$$F'(\hat{\lambda}) \; = \; xU'(\hat{\lambda}x) \; < \; \frac{\gamma}{\hat{\lambda}} U(\hat{\lambda}x) \; = \; \frac{\gamma}{\hat{\lambda}} F(\hat{\lambda}) \; = \; \frac{\gamma}{\hat{\lambda}} G(\hat{\lambda}) \; = \; G'(\hat{\lambda}),$$

qui est la contradiction voulue. Réciproquement supposons (iii). Alors $F'(1) \leq G'(1)$ et on a

$$U'(x) \; = \; \frac{F'(1)}{x} \leq \frac{G'(1)}{x} \; = \gamma \frac{U(x)}{x},$$

qui implique clairement l'assertion (ii).

$(iv) \Leftrightarrow (v)$: Cette équivalence se montre de manière similaire à $(ii) \Leftrightarrow (iii)$ en fixant $0 < y \leq y_0$ et en considérant les fonctions $F(\mu) = \tilde{U}(\mu y)$ et $G(\mu) = \mu^{-\frac{\gamma}{1-\gamma}} \tilde{U}(y)$.

$(ii) \Leftrightarrow (v)$: Supposons (ii) et posons $y_0 = U'(x_0)$. Alors pour tout $0 < y \leq y_0$, on a $I(y) \geq I(y_0) = x_0$ et donc

$$\tilde{U}(y) = U(I(y)) - yI(y) \; > \; \frac{1}{\gamma} I(y) U'(I(y)) - yI(y) \; = \; \frac{1-\gamma}{\gamma} yI(y)$$

Comme $\tilde{U}'(y) = -I(y)$, ceci prouve (v). Réciproquement, supposons (v) et posons $x_0 = I(y_0) = -\tilde{U}'(y_0)$. Alors pour tout $x \geq x_0$, on a $U'(x) \leq U'(x_0) = y_0$ et donc

$$U(x) = \tilde{U}(U'(x)) + xU'(x) \; > \; -\frac{1-\gamma}{\gamma} U'(x) \tilde{U}'(U'(x)) + xU'(x) = \frac{1}{\gamma} xU'(x),$$

qui est précisément (ii). □

Remarque 6.3.9 **1.** Les caractérisations (iv) et (v) montrent que si $AE(U) < 1$ alors il existe $y_0 > 0$ tel que pour tout $0 < \mu < 1$:

$$yI(\mu y) \leq C\tilde{U}(y), \quad 0 < y \leq y_0, \tag{6.55}$$

où C est une constante positive dépendant de μ.

2. Notons que la caractérisation (ii) de $AE(U) < 1$ implique la condition de croissance (6.43) sur U vue à la remarque 6.3.5.

Le résultat suivant donne une caractérisation de la solution du problème dual en écrivant les conditions d'optimalité du premier ordre.

Proposition 6.3.4 *(Caractérisation de la solution du problème dual)*
Soit U une fonction d'utilité vérifiant (6.25), (6.27), (6.36) *et* (6.47). *Alors $\tilde{v}$ est fini, dérivable, strictement convexe sur* $]0,+\infty[$ *avec*

$$\begin{aligned} -\tilde{v}'(y) &= E\left[\hat{Y}_T^y I(y\hat{Y}_T^y)\right] \\ &= \sup_{Y_T \in \mathcal{D}} E\left[Y_T I(y\hat{Y}_T^y)\right], \quad y > 0, \end{aligned} \tag{6.56}$$

et donc

$$I(y\hat{Y}_T^y) \in \mathcal{C}(-\tilde{v}'(y)), \quad y > 0.$$

Preuve. 1. On a vu à la remarque 6.3.8 que l'hypothèse (6.27) est équivalente à $\mathrm{dom}(\tilde{v}) \neq \emptyset$, i.e. il existe $y_1 > 0$ tel que

$$\tilde{v}(y) < +\infty, \quad \forall y \geq y_1, \tag{6.57}$$

d'après la décroissance de $\tilde{U}$ et donc de $\tilde{v}$. Puisque $\tilde{v}(y_1) < +\infty$, il existe $Y_T \in \mathcal{D}$ tel que $E[\tilde{U}(y_1 Y_T)] < +\infty$. Comme on a aussi $\tilde{U}(y_1 Y_T) \geq U(x_0) - x_0 y_1 Y_T$ avec $E[Y_T] \leq 1$ et $x_0 > 0$ donné, ceci prouve que $\tilde{U}(y_1 Y_T) \in L^1(P)$. De plus, la caractérisation (iv) de $AE(U) < 1$ du lemme 6.3.10 montre qu'il existe $y_0 > 0$ tel que pour tout $0 < y < y_1$:

$$\begin{aligned} \tilde{U}(yY_T) &\leq C(y)\tilde{U}(y_1 Y_T)1_{y_1 Y_T \leq y_0} + \tilde{U}(yY_T)1_{y_1 Y_T > y_0} \\ &\leq C(y)\left|\tilde{U}(y_1 Y_T)\right| + \left|\tilde{U}\left(\frac{y}{y_1}y_0\right)\right|, \end{aligned}$$

où $C(y)$ est une constante positive. Ceci prouve que $\tilde{v}(y) < +\infty$ pour $y < y_1$ et donc $\mathrm{dom}(\tilde{v}) =]0,+\infty[$.

2. Soit $y > 0$ fixé. Alors pour tout $\delta > 0$, on a par définition de $\tilde{v}$:

$$\begin{aligned} \frac{\tilde{v}(y+\delta) - \tilde{v}(y)}{\delta} &\leq E\left[\frac{\tilde{U}((y+\delta)\hat{Y}_T^y)) - \tilde{U}(y\hat{Y}_T^y)}{\delta}\right] \\ &\leq E\left[\hat{Y}_T^y \tilde{U}'((y+\delta)\hat{Y}_T^y))\right], \end{aligned}$$

d'après la convexité de $\tilde{U}$. Puisque $\tilde{U}' = -I \geq 0$, on en déduit par le lemme de Fatou en faisant tendre δ vers zéro :

$$\limsup_{\delta \downarrow 0} \frac{\tilde{v}(y+\delta) - \tilde{v}(y)}{\delta} \leq -E\left[\hat{Y}_T^y I(y\hat{Y}_T^y)\right]. \tag{6.58}$$

D'autre part, pour tout $\delta > 0$ tel que $y - \delta > 0$, on a de manière similaire à ci-dessus :

$$\begin{aligned}\frac{\tilde{v}(y)-\tilde{v}(y-\delta)}{\delta} &\geq E\left[\frac{\tilde{U}(y\hat{Y}_T^y))-\tilde{U}((y-\delta)\hat{Y}_T^y)}{\delta}\right]\\ &\geq E\left[\hat{Y}_T^y\tilde{U}'((y-\delta)\hat{Y}_T^y))\right]. \end{aligned}\tag{6.59}$$

Notons (comme au point **1.**) que puisque $E[\tilde{U}(y\hat{Y}_T^y)] < +\infty$, alors $\tilde{U}(y\hat{Y}_T^y) \in L^1(P)$. D'après (6.55) (conséquence de $AE(U) < 1$), il existe $y_0 > 0$ tel que pour tout $0 < \delta < y/2$, on a :

$$\begin{aligned}0 \leq -\hat{Y}_T^y\tilde{U}'((y-\delta)\hat{Y}_T^y)) \;&=\; \hat{Y}_T^y I((y-\delta)\hat{Y}_T^y))\\ \leq C(y)\tilde{U}(y\hat{Y}_T^y)1_{y\hat{Y}_T^y\leq y_0} &+ \hat{Y}_T^y I\left(\frac{y_0}{2}\right)1_{y\hat{Y}_T^y>y_0}\\ \leq C(y)\left|\tilde{U}(y\hat{Y}_T^y)\right| &+ \hat{Y}_T^y I\left(\frac{y_0}{2}\right),\end{aligned}$$

où $C(y) < +\infty$ et $x_0 > 0$ est quelconque. Le terme de droite de cette dernière inégalité étant intégrable, on peut appliquer le théorème de convergence dominée à (6.59) lorsque δ tend vers zéro :

$$\liminf_{\delta\downarrow 0}\frac{\tilde{v}(y)-\tilde{v}(y-\delta)}{\delta} \geq -E\left[\hat{Y}_T^y I(y\hat{Y}_T^y))\right]. \tag{6.60}$$

En combinant avec (6.58) et avec la convexité de $\tilde{v}$, on obtient que $\tilde{v}$ est dérivable en tout $y \in]0,+\infty[$ avec :

$$\tilde{v}'(y) = -E\left[\hat{Y}_T^y I(y\hat{Y}_T^y))\right].$$

3. Etant donné un élément arbitraire $Y_T \in \mathcal{D}$, on pose

$$Y_T^\varepsilon = (1-\varepsilon)\hat{Y}_T^y + \varepsilon Y_T \;\in\; \mathcal{D}, \quad 0 < \varepsilon < 1.$$

Alors par définition de $\tilde{v}(y)$ et la convexité de $\tilde{U}$, on a :

$$\begin{aligned}0 &\leq E[\tilde{U}(yY_T^\varepsilon)] - E[\tilde{U}(y\hat{Y}_T^y)]\\ &\leq yE[\tilde{U}'(yY_T^\varepsilon)(Y_T^\varepsilon - \hat{Y}_T^y)] \;=\; \varepsilon yE[I(yY_T^\varepsilon)(\hat{Y}_T^y - Y_T)],\end{aligned}$$

d'où l'on déduit par la décroissance de I :

$$E[Y_T I(yY_T^\varepsilon)] \leq E[\hat{Y}_T^y I(y(1-\varepsilon)\hat{Y}_T^y)].$$

Comme dans le point **2.**, on peut appliquer le théorème de convergence dominée (sous $AE(U) < 1$) au terme de droite et le lemme de Fatou au terme de gauche pour obtenir quand ε tend vers zéro :

$$E[Y_T I(yY_T)] \leq E[\hat{Y}_T^y I(y\hat{Y}_T^y)],$$

et ceci pour tout $Y_T \in \mathcal{D}$. C'est la relation (6.56). La propriété $I(y\hat{Y}_T^y) \in \mathcal{C}(-\tilde{v}'(y))$ découle alors finalement de la caractérisation duale (6.45). □

Preuve du théorème 6.3.5

• L'existence d'une solution à $v(x)$ pour tout $x > 0$ vient du fait que dom$(\tilde{v})$ $= \,]0, +\infty[$ qui assure l'hypothèse (6.30) (voir remarque 6.3.5). La stricte convexité de $\tilde{v}$ sur $]0, +\infty[$ et la relation de conjugaison (6.50) montre (voir Proposition B.3.5 de l'appendice B) que v est dérivable sur $]0, +\infty[$. La stricte concavité de v sur $]0, +\infty[$ est une conséquence de la stricte concavité de U et de l'unicité de la solution à $v(x)$. Cela implique aussi en retour que $\tilde{v}$ est dérivable sur $]0, +\infty[$.

• L'assertion (2) du théorème découle de la proposition 6.3.3 et du fait que dom$(\tilde{v}) = \,]0, +\infty[$.

• Vérifions d'une part que $\tilde{v}'(+\infty) := \lim_{y\to+\infty} \tilde{v}'(y) = 0$. Comme la fonction $-\tilde{U}$ est croissante et $-\tilde{U}'(y) = I(y)$ tend vers 0 quand y tend vers l'infini, alors pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ tel que

$$-\tilde{U}(y) \leq C_\varepsilon + \varepsilon y, \quad y > 0.$$

D'après la règle de l'Hôpital, on en déduit que :

$$\begin{aligned}
0 \;\leq\; -\tilde{v}'(+\infty) = \lim_{y\to+\infty} \frac{-\tilde{v}(y)}{y} \;&=\; \lim_{y\to+\infty} \sup_{Y_T\in\mathcal{D}} E\left[\frac{-U(yY_T)}{y}\right] \\
&\leq \lim_{y\to+\infty} \sup_{Y_T\in\mathcal{D}} E\left[\frac{C_\varepsilon}{y} + \varepsilon Y_T\right] \\
&\leq \lim_{y\to+\infty} E\left[\frac{C_\varepsilon}{y} + \varepsilon\right] \;=\; \varepsilon,
\end{aligned}$$

où on a utilisé le fait que $E[Y_T] \leq 1$ pour tout $Y_T \in \mathcal{D}$, d'après (6.45) avec $X_T = 1 \in \mathcal{C}(1)$. Ceci prouve

$$\tilde{v}'(+\infty) = 0. \tag{6.61}$$

D'autre part, d'après (6.56), on a

$$-\tilde{v}'(y) \geq E[Z_T I(y\hat{Y}_T^y)], \quad \forall y > 0, \tag{6.62}$$

où $Z_T > 0$ p.s. est un élément fixé dans $\mathcal{M}_e$. Notons que puisque $E[\hat{Y}_T^y] \leq 1$ pour tout $y > 0$, on a par le lemme de Fatou, $E[\hat{Y}_T^0] \leq 1$ où $\hat{Y}_T^0 = \liminf_{y\downarrow 0} \hat{Y}_T^y$. En particulier $\hat{Y}_T^0 < +\infty$ p.s. En faisant tendre y vers zéro dans (6.62) et en se rappelant que $I(0) = +\infty$, on obtient alors par le lemme de Fatou :

$$\tilde{v}'(0) \;=\; \lim_{y\downarrow 0} \tilde{v}'(y) = +\infty. \tag{6.63}$$

D'après (6.61) et (6.63), on en déduit que pour tout $x > 0$, la fonction strictement convexe $y \in \,]0, +\infty[\, \to \tilde{v}(y) + xy$ admet un unique minimum $\hat{y}$ caractérisé par $\tilde{v}'(\hat{y}) = -x$ ou de manière équivalente $\hat{y} = v'(x)$ car $\tilde{v}' = -(v')-1$ d'après la relation de conjugaison (6.51).

• Montrons que $I(\hat{y}\hat{Y}_T)$ est la solution de $v(x)$. D'après la proposition 6.3.4, on a

$$I(\hat{y}\hat{Y}_T) \ \in \ \mathcal{C}(x) \ \text{ et } \ E[\hat{Y}_T I(\hat{y}\hat{Y}_T)] \ = \ x.$$

On a alors :

$$\begin{aligned} v(x) \ \geq \ E[U(I(\hat{y}\hat{Y}_T))] &= E[\tilde{U}(\hat{y}\hat{Y}_T)] + E[\hat{y}\hat{Y}_T I(\hat{y}\hat{Y}_T)] \\ &= E[\tilde{U}(\hat{y}\hat{Y}_T)] + x\hat{y} \\ &\geq \tilde{v}(\hat{y}) + x\hat{y}. \end{aligned}$$

En combinant avec la relation de conjugaison (6.50), ceci prouve qu'on a égalité dans les inégalités ci-dessus et donc que $I(\hat{y}\hat{Y}_T)$ est la solution de $v(x)$.

• Il reste à montrer l'assertion (4) du théorème. Sous l'hypothèse $\inf_{Z_T\in\mathcal{M}_e} E[\tilde{U}(yZ_T)] < +\infty$ pour $y > 0$ donné, on peut trouver $Z_T^0 \in \mathcal{M}_e$ tel que $\tilde{U}(yZ_T^0) \in L^1(P)$. Considérons un élément quelconque $Y_T \in \mathcal{D}$ et soit $(Z^n)_{n\geq 1}$ la suite dans $\mathcal{M}_e$ tel que $Y_T \leq \lim_n Z^n$ p.s. Pour $\varepsilon \in]0,1[$ et $n \geq 1$, on pose

$$\bar{Z}^{n,\varepsilon} = (1-\varepsilon)Z^n + \varepsilon Z_T^0 \ \in \ \mathcal{M}_e.$$

Alors d'après la décroissance de $\tilde{U}$ et la caractérisation (iv) de $AE(U) < 1$ du lemme 6.3.10, il existe $y_0 > 0$ tel que :

$$\tilde{U}(y\bar{Z}^{n,\varepsilon}) \ \leq \ \tilde{U}(y\varepsilon Z_T^0) \leq C_\varepsilon \tilde{U}(yZ_T^0)1_{yZ_T^0\leq y_0} + \tilde{U}(\varepsilon y_0)1_{yZ_T^0 > y_0},$$

où C_ε est une constant positive. On a alors :

$$\tilde{U}^+(y\bar{Z}^{n,\varepsilon}) \leq C_\varepsilon|\tilde{U}(yZ_T^0)| + |\tilde{U}(\varepsilon y_0)|, \quad \forall n \geq 1,$$

ce qui prouve que la suite $\{\tilde{U}^+(y\bar{Z}^{n,\varepsilon}),\, n \geq 1\}$ est uniformément intégrable. On en déduit en utilisant aussi le lemme de Fatou et la décroissance de $\tilde{U}$ que :

$$\begin{aligned} \inf_{Z_T\in\mathcal{M}_e} E[\tilde{U}(yZ_T)] \ \leq \ \limsup_{n\to+\infty} E[\tilde{U}(y\bar{Z}^{n,\varepsilon})] &\leq E[\tilde{U}(y(1-\varepsilon)Y_T + \varepsilon Z_T^0)] \\ &\leq E[\tilde{U}(y(1-\varepsilon)Y_T)]. \end{aligned}$$

En utilisant de nouveau la caractérisation (iv) de $AE(U) < 1$ du lemme 6.3.10, on a l'uniforme intégrabilité de $\{\tilde{U}^+(y(1-\varepsilon)Y_T),\, \varepsilon \in]0,1[\}$. En faisant tendre ε vers zéro dans l'inégalité précédente, on obtient :

$$\inf_{Z_T\in\mathcal{M}_e} E[\tilde{U}(yZ_T)] \leq E[\tilde{U}(yY_T)],$$

et ceci pour tout $Y_T \in \mathcal{D}$. Ceci montre $\inf_{Z_T\in\mathcal{M}_e} E[\tilde{U}(yZ_T)] \leq \tilde{v}(y)$. L'inégalité inverse est évidente car $\mathcal{M}_e \subset \mathcal{D}$.

6.3.4 Le cas des marchés complets

Dans ce paragraphe, on se place dans le cas où le marché financier est complet, c'est à dire que l'ensemble des probabilités martingales est réduit à un singleton :

$$\mathcal{M}_e(S) = \{P^0\}.$$

On note Z^0 le processus de densité martingale P^0. Dans ce contexte, le problème dual est dégénéré

$$\tilde{v}(y) = E[\tilde{U}(yZ_T^0)], \quad y > 0.$$

et la solution au problème dual est évidemment Z_T^0. La solution au problème primal $v(x)$ est

$$\hat{X}_T^x = I(\hat{y}Z_T^0)$$

où $\hat{y} > 0$ est la solution de

$$E[Z_T^0 I(\hat{y}Z_T^0)] \ = \ E^{P^0}[I(\hat{y}Z_T^0)] = x.$$

Le processus de richesse $\hat{X}^x$ et le portefeuille optimal $\hat{\alpha}$ associé sont donnés par :

$$\hat{X}_t^x \ = \ x + \int_0^t \hat{\alpha}_u dS_u = E^{P^0}\left[I(\hat{y}Z_T^0) \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T.$$

Exemple : Modèle de Merton

Prenons l'exemple typique du modèle de Black-Scholes-Merton :

$$dS_t = S_t \left(\mu dt + \sigma dW_t\right),$$

où W est un mouvement Brownien standard réel sur $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ avec $\mathbb{F}$ filtration naturelle de W, $\mathcal{F}_0$ trivial et μ, σ sont deux constantes avec $\sigma > 0$. Rappelons (voir paragraphe 6.2.4) que l'unique probabilité martingale est donnée par

$$Z_T^0 = \exp\left(-\lambda W_T - \frac{1}{2}|\lambda|^2 T\right), \quad \text{où } \lambda \ = \ \frac{\mu}{\sigma},$$

et la dynamique de S sous P^0 est

$$dS_t = S_t \sigma dW_t^0,$$

où $W_t^0 = W_t + \lambda t$, $0 \leq t \leq T$, est un P^0 mouvement Brownien. Considérons l'exemple d'une fonction d'utilité puissance :

$$U(x) \;=\; \frac{x^p}{p},\;\; p<1,\; p\neq 0,\;\; \text{pour lequel}\;\; I(y) \;=\; y^{-r},\;\; r \;=\; \frac{1}{1-p}.$$

On calcule alors explicitement le processus de richesse optimal pour $v(x)$:

$$\begin{aligned}\hat{X}_t^x &= E^{P^0}\left[(\hat{y}Z_T^0)^{-r} \middle| \mathcal{F}_t\right] \;=\; \hat{y}^{-r} E^{P^0}\left[\exp\left(\lambda r W_T^0 - \frac{1}{2}|\lambda|^2 rT\right)\middle| \mathcal{F}_t\right] \\ &= \hat{y}^{-r}\exp\left[\frac{1}{2}\left(|\lambda r|^2 - |\lambda|^2 r\right)T\right]\,\exp\left(\lambda r W_t^0 - \frac{1}{2}|\lambda r|^2 t\right).\end{aligned}$$

Comme $\hat{y}$ est déterminé par l'équation $\hat{X}_0^x = x$, on obtient :

$$\hat{X}_t^x = x\exp\left(\lambda r W_t^0 - \frac{1}{2}|\lambda r|^2 t\right), \quad 0 \le t \le T.$$

Pour déterminer le contrôle optimal $\hat{\alpha}$, on applique la formule d'Itô à $\hat{X}^x$:

$$d\hat{X}_t^x \;=\; \hat{X}_t^x \lambda r dW_t^0$$

et on identifie avec

$$d\hat{X}_t^x = \hat{\alpha}_t dS_t \;=\; \hat{\alpha}_t \sigma S_t dW_t^0,$$

pour obtenir la proportion optimale de richesse investie dans S :

$$\frac{\hat{\alpha}_t S_t}{\hat{X}_t^x} = \frac{\lambda r}{\sigma} \;=\; \frac{\mu}{\sigma^2(1-p)}.$$

On retrouve bien le même résultat que par l'approche de programmation dynamique de Bellman dans le paragraphe 3.6.1. Le calcul de la fonction valeur $v(x) = E[U(\hat{X}_T^x)]$ est aisé et donne bien entendu la même valeur que celle obtenue au paragraphe 3.6.1 :

$$v(x) = \frac{x^p}{p}\;\exp\left(\frac{1}{2}\frac{\mu^2}{\sigma^2}\frac{p}{1-p}T\right).$$

6.3.5 Exemples en marché incomplet

Dans le cadre des marchés incomplets, i.e. $\mathcal{M}_e(S)$ n'est pas réduit à un singleton, et alors en fait est de cardinalité infinie, on ne sait pas calculer en général explicitement la solution du problème dual. Toutefois, des calculs peuvent être menés plus ou moins explicitement dans certains modèles particuliers. Considérons ainsi le cadre du modèle de prix d'Itô décrit au paragraphe 6.2.4. Avec les notations de ce paragraphe, considérons l'ensemble contenant $\mathcal{M}_e(S)$:

$$\mathcal{M}_{loc} = \{Z_T^\nu \;: \nu \in K(\sigma)\} \;\supset\; \mathcal{M}_e(S) \;=\; \{Z_T^\nu \;: \nu \in K_m(\sigma)\}.$$

Il est facile de vérifier par la formule d'Itô que pour tout processus de richesse $X^x = x + \int \alpha dS$, $\alpha \in \mathcal{A}(S)$, et tout $\nu \in K(\sigma)$, le processus $Z^\nu X^x$ est une P-martingale locale. Notons aussi que pour tout $\nu \in K(\sigma)$, le processus borné $\nu^n = \nu 1_{|\nu| \leq n}$ est dans $K_m(\sigma)$ et on a $Z_T^{\nu^n}$ converge p.s. vers Z_T^ν. Donc $\mathcal{M}_{loc} \subset \mathcal{D}$ et d'après l'assertion (4) du théorème 6.3.5 (voir aussi la remarque 6.3.7), on a :

$$\tilde{v}(y) = \inf_{\nu \in K(\sigma)} E[\tilde{U}(yZ_T^\nu)], \quad y > 0. \tag{6.64}$$

L'intérêt d'avoir introduit l'ensemble $\mathcal{M}_{loc}$ est qu'il est explicite (au contraire de $\mathcal{D}$), complètement paramétré par l'espace des contrôles $\nu \in K(\sigma)$, et relachée des contraintes fortes d'intégrabilité martingale de $K_m(\sigma)$, de sorte qu'on peut espérer trouver une solution $\hat{\nu}^y$ dans $K(\sigma)$ à $\tilde{v}(y)$ en (6.64) par des méthodes de contrôle stochastique vues aux chapitres précédents. En fait, si on suppose que la fonction

$$\xi \in \mathbb{R} \longmapsto \tilde{U}(e^\xi) \quad \text{est convexe },$$

ce qui est vérifié dès que $x \in]0, +\infty[\mapsto xU'(x)$ est croissante (c'est le cas des fonctions d'utilité logarithmique et puissance), alors il est prouvé dans Karatzas et al. [KLSX91] que pour tout $y > 0$, le problème dual $\tilde{v}(y)$ admet une solution $Z_T^{\hat{\nu}^y} \in \mathcal{M}_{loc}$. On montre aussi que pour tout $\nu \in K(\sigma)$ tel que $E[\int_0^T |\nu_t|^2 dt] = +\infty$, on a $E[\tilde{U}(yZ_T^\nu)] = +\infty$, de sorte qu'on peut se limiter dans (6.64) à prendre l'infimum sur $K_2(\sigma) = \{\nu \in K(\sigma) \ : E[\int_0^T |\nu_t|^2 dt] < +\infty\}$, et donc $\hat{\nu}^y \in K_2(\sigma)$. Il est à noter qu'en général cette solution $Z_T^{\hat{\nu}^y}$ n'est pas dans $\mathcal{M}_e(S)$. La solution du problème primal $v(x)$ est donnée par :

$$\hat{X}_T^x = I\left(\hat{y} Z_T^{\hat{\nu}^{\hat{y}}}\right),$$

où $\hat{y} > 0$ est la solution de $\mathrm{argmin}_{y>0}[\tilde{v}(y) + xy]$ et vérifiant :

$$E\left[Z_T^{\hat{\nu}^{\hat{y}}} I\left(\hat{y} Z_T^{\hat{\nu}^{\hat{y}}}\right)\right] = x.$$

Pour déterminer le processus de richesse (positif) $\hat{X}^x$, on note que le processus $Z^{\hat{\nu}^{\hat{y}}} \hat{X}^x$ est une martingale locale positive, donc une surmartingale, et qui vérifie de plus $E[Z_T^{\hat{\nu}^{\hat{y}}} \hat{X}_T^x] = x$. C'est donc une martingale, et on a :

$$\hat{X}_t^x = E\left[\frac{Z_T^{\hat{\nu}^{\hat{y}}}}{Z_t^{\hat{\nu}^{\hat{y}}}} I\left(\hat{y} Z_T^{\hat{\nu}^{\hat{y}}}\right) \Big| \mathcal{F}_t\right], \quad 0 \leq t \leq T.$$

Fonction d'utilité logarithmique

Considérons l'exemple d'une fonction d'utilité logarithmique $U(x) = \ln x$, $x > 0$, pour lequel

$$I(y) \;=\; \frac{1}{y} \;\text{ et }\; \tilde{U}(y) \;=\; -\ln y - 1, \quad y > 0.$$

Pour tout $\nu \in K_2(\sigma)$, on a donc :

$$E[\tilde{U}(yZ_T^\nu)] = -\ln y - 1 - \frac{1}{2}\int_0^T |\lambda_t|^2 + [\nu_t|^2 dt, \quad y > 0.$$

Ceci prouve que la solution du problème dual $\tilde{v}(y)$ est atteint pour $\nu = 0$ (en particulier indépendant de y), correspondant à Z_T^0 et la richesse optimale pour $v(x)$ est explicitée par :

$$\hat{X}_t^x = \frac{1}{Z_t^0}, \quad 0 \le t \le T.$$

Le contrôle optimal est déterminé en appliquant la formule d'Itô à l'expression ci-dessus et en identifiant avec $d\hat{X}_t^x = \hat{\alpha}_t dS_t$. Dans un modèle de prix écrit sous forme "géométrique"

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t),$$

on trouve la proportion optimale de richesse investie dans S :

$$\frac{\hat{\alpha}_t S_t}{\hat{X}_t^x} = \frac{\mu_t}{\sigma_t^2}.$$

Fonction d'utilité puissance

Considérons l'exemple d'une fonction d'utilité puissance $U(x) = x^p/p$, $x > 0$, $p < 1$, $p \neq 0$, pour lequel

$$I(y) \;=\; y^{\frac{1}{p-1}} \;\text{ et }\; \tilde{U}(y) \;=\; \frac{y^{-q}}{q}, \quad y > 0,\; q \;=\; \frac{p}{1-p}.$$

Pour tout $\nu \in K(\sigma)$, on a donc :

$$E[\tilde{U}(yZ_T^\nu)] = \frac{y^{-q}}{q}\, E[(Z_T^\nu)^{-q}], \quad y > 0.$$

Ainsi, la solution du problème dual $\tilde{v}(y)$ ne dépend pas de y et est solution du problème

$$\inf_{\nu \in K(\sigma)} E[(Z_T^\nu)^{-q}]. \tag{6.65}$$

Ce problème de contrôle stochastique peut être résolu dans un cadre Markovien, typiquement un modèle de marché incomplet à volatilité stochastique, par la méthode de la programmation dynamique. On peut aussi dans un cadre plus général de processus d'Itô, utiliser des méthodes d'EDSR. En notant par $\hat{\nu}$ la solution de (6.65), la richesse optimale de $v(x)$ est donnée par :

$$\hat{X}_t^x = \frac{x}{E[(Z_T^{\hat{\nu}})^{-q}]} E\left[\frac{(Z_T^{\hat{\nu}})^{-q}}{Z_t^{\hat{\nu}}}\middle|\mathcal{F}_t\right], \quad 0 \le t \le T.$$

6.4 Problème de couverture quadratique

6.4.1 Formulation du problème

On se place dans le cadre du modèle général de processus de prix semimartingale continue décrit au paragraphe 6.2.1. On se donne un actif contingent représenté par une variable aléatoire $H \in L^2(P) = L^2(\Omega, \mathcal{F}_T, P)$, i.e. H $\mathcal{F}_T$-mesurable et $E[H]^2 < +\infty$. Le critère de couverture quadratique consiste à minimiser en norme $L^2(P)$ la différence entre H et la richesse terminale $X_T = x + \int_0^T \alpha_t dS_t$ d'une stratégie de portefeuille $\alpha \in L(S)$.

Dans le problème de maximisation d'utilité, nous avions considéré des fonctions d'utilité de domaine $]0, +\infty[$. Il était donc naturel de définir des stratégies de portefeuille admissible conduisant à un processus de richesse borné inférieurement. Dans le problème de couverture quadratique, la fonction de coût à minimiser $U(x) = (H - x)^2$ est défini sur tout $\mathbb{R}$ et en général, on ne peut espérer trouver une solution $\hat{X} = x + \int \hat{\alpha} dS$ qui soit bornée inférieurement. D'un autre coté, on se doit d'exclure les stratégies d'arbitrage "doubling". Dans notre cadre de minimisation en norme quadratique, on introduit la condition d'admissibilité suivante :

$$\mathcal{A}_2 = \Big\{ \alpha \in L(S) \ : \ \int_0^T \alpha_t dS_t \in L^2(P) \quad \text{et} \\ \int \alpha dS \ \text{ est une } \ Q - \text{martingale pour tout } Q \in \mathcal{M}_e^2 \Big\},$$

où

$$\mathcal{M}_e^2 = \Big\{ Q \in \mathcal{M}_e \ : \frac{dQ}{dP} \in L^2(P) \Big\}$$

est supposé non vide : $\mathcal{M}_e^2 \neq \emptyset$. Cette condition d'admissibilité permet des richesses non bornées inférieurement tout en excluant les possibilités d'arbitrage. En effet, en fixant un élément $Q \in \mathcal{M}_e^2$, on a pour tout $\alpha \in \mathcal{A}_2$, $E^Q[\int_0^T \alpha_t dS_t] = 0$ et donc

$$\nexists \ \alpha \in \mathcal{A}_2, \quad \int_0^T \alpha_t dS_t \geq 0, \ p.s. \quad \text{et} \quad P\left[\int_0^T \alpha_t dS_t > 0 \right] > 0.$$

D'autre part, avec cette condition d'intégrabilité dans $\mathcal{A}_2$, on a que l'espace des intégrales stochastiques $\{\int_0^T \alpha_t dS_t \ : \alpha \in \mathcal{A}_2\}$ est fermé dans $L^2(P)$, ce qui assure l'existence d'une solution au problème de minimisation quadratique.

Proposition 6.4.5 *L'ensemble* $\mathcal{G}_T = \{\int_0^T \alpha_t dS_t \ : \alpha \in \mathcal{A}_2\}$ *est fermé dans* $L^2(P)$.

Preuve. Soit $X^n = \int_0^T \alpha_t^n dS_t$, $n \in \mathbb{N}$, une suite dans $\mathcal{G}_T$ convergent vers X_T dans $L^2(P)$. Fixons un élément arbitraire $Q \in \mathcal{M}_e^2$. Alors d'après l'inégalité de Cauchy-Schwarz, $\int_0^T \alpha_t^n dS_t$ converge vers X_T dans $L^1(Q)$. Grâce au théorème 1.2.10, on en déduit que $X_T = \int_0^T \alpha_t dS_t$ où $\alpha \in L(S)$ est tel que $\int \alpha dS$ est une Q-martingale. Puisque Q est arbitraire dans $\mathcal{M}_e^2$, on conclut que $\alpha \in \mathcal{A}_2$. □

Le problème de minimisation quadratique (appelé aussi parfois problème moyenne-variance) d'un actif contingent $H \in L^2(P)$ se formule ainsi :

$$v_H(x) = \inf_{\alpha \in \mathcal{A}_2} E\left[H - x - \int_0^T \alpha_t dS_t\right]^2, \quad x \in \mathbb{R}. \tag{6.66}$$

Autrement dit, on projette dans l'espace de Hilbert $L^2(P)$, l'élément $H - x$ sur le sous-espace vectoriel fermé $\mathcal{G}_T$. On sait donc déjà l'existence d'une solution à $v_H(x)$. Notre objectif est désormais de caractériser cette solution. La méthode de résolution est basée sur une combinaison du théorème de projection de Kunita-Watanabe, des méthodes de dualité convexe et de changement de numéraire.

6.4.2 Le cas martingale

Dans cette section, nous étudions le cas particulier où S est une martingale locale sous P, i.e. $P \in \mathcal{M}_e^2$. Dans ce cas, la résolution est directe grâce au théorème de projection de Kunita-Watanabe. En effet, en projetant la martingale de carré intégrable $H_t = E[H|\mathcal{F}_t]$, $0 \leq t \leq T$, sur la martingale locale continue S sous P, on a la décomposition :

$$E[H|\mathcal{F}_t] = E[H] + \int_0^t \alpha_u^H dS_u + R_t^H, \quad 0 \leq t \leq T, \tag{6.67}$$

où $\alpha^H \in L(S)$ vérifie la condition d'intégrabilité

$$E\left[\int_0^T \alpha_t^H dS_t\right]^2 = E\left[\int_0^T (\alpha_t^H)' d < S >_t \alpha_t^H\right] < +\infty, \tag{6.68}$$

et $(R_t^H)_t$ est une martingale de carré intégrable, orthogonale à S, i.e. $< R^H, S > = 0$.

Théorème 6.4.6 *Pour tout $x \in \mathbb{R}$ et $H \in L^2(P)$, la solution $\alpha^{mv} \in \mathcal{A}_2$ de $v_H(x)$ est égale à α^H. De plus, on a*

$$v_H(x) = (E[H] - x)^2 + E[R_T^H]^2.$$

Preuve. Vérifions d'abord que $\alpha^H \in \mathcal{A}_2$. D'après la condition (6.68), on a $\int_0^T \alpha_t^H dS_t \in L^2(P)$. D'autre part, avec les inégalites de Cauchy-Schwarz et de Doob, on a pour tout $Q \in \mathcal{M}_e^2$:

$$E^Q\left[\sup_{0\le t\le T}\left|\int_0^t \alpha_u^H dS_u\right|\right] \le 2\left(E\left[\frac{dQ}{dP}\right]^2\right)^{\frac12}\left(E\left[\int_0^T \alpha_t^H dS_t\right]^2\right)^{\frac12} < +\infty.$$

Ceci prouve que la Q-martingale locale $\int \alpha^H dS$ est une Q-martingale uniformément intégrable. Ainsi, on a bien $\alpha^H \in \mathcal{A}_2$.

Notons que pour tout $\alpha \in \mathcal{A}_2$, l'intégrale stochastique $\int \alpha dS$ est une P-martingale de carré intégrable. Par les inégalités de Doob et de Cauchy-Schwarz, on en déduit que $E[\sup_t |R_t^H \int_0^t \alpha_u dS_u|] < +\infty$. La condition d'orthogonalité de R^H et S implique alors que $R^H \int \alpha dS$ est une P-martingale uniformément intégrable. On a donc

$$E[R_T^H \int_0^T \alpha_t dS_t] = 0, \quad \forall \alpha \in \mathcal{A}_2.$$

En écrivant d'après (6.67) que $H = E[H] + \int_0^T \alpha_t^H dS_t + R_T^H$, on obtient alors que pour tout $\alpha \in \mathcal{A}_2$:

$$E\left[H - x - \int_0^T \alpha_t dS_t\right]^2 = (E[H]-x)^2 + E\left[\int_0^T (\alpha_t^H - \alpha_t)dS_t\right]^2 + E[R_T^H]^2,$$

ce qui montre le résultat voulu. □

Exemple

Considérons le modèle à volatilité stochastique :

$$\begin{aligned} dS_t &= \sigma(t, S_t, Y_t)S_t dW_t^1 \\ dY_t &= \eta(t, S_t, Y_t)dt + \gamma(t, S_t, Y_t)dW_t^2, \end{aligned}$$

où W^1 et W^2 sont deux mouvements Browniens, supposés non corrélés pour simplifier. On fait les hypothèses standard sur les coefficients σ, η, γ pour avoir l'existence et l'unicité d'une solution (S, Y) à valeurs dans $\mathbb{R}_+ \times \mathbb{R}$ à l'EDS ci-dessus étant donné une condition initiale. On prend un actif contingent de la forme $H = g(S_T)$ où g est une fonction mesurable et on considère la fonction

$$h(t, s, y) = E[g(S_T)|(S_t, Y_t) = (s, y)], \quad (t, s, y) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}.$$

Sous des conditions appropriées sur les coefficients σ, η, γ et g, la fonction $h \in C^{1,2}([0, T[\times\mathbb{R}_+ \times \mathbb{R})$ et est alors solution du problème de Cauchy :

$$\begin{aligned} \frac{\partial h}{\partial t} + \eta\frac{\partial h}{\partial y} + \frac12\sigma^2\frac{\partial^2 h}{\partial s^2} + \frac12\gamma^2\frac{\partial^2 h}{\partial y^2} &= 0, \quad \text{sur} \quad [0, T[\times\mathbb{R}_+ \times \mathbb{R} \\ h(T, ., .) &= g, \quad \text{sur} \quad \mathbb{R}_+ \times \mathbb{R}. \end{aligned}$$

La décomposition de Kunita-Watanabe (6.67) est simplement obtenue par la formule d'Itô appliqué processus martingale $h(t, S_t, Y_t) = E[g(S_T)|\mathcal{F}_t]$, $0 \leq t \leq T$:

$$E[g(S_T)|\mathcal{F}_t] = E\left[g(S_T)\right] + \int_0^t \frac{\partial h}{\partial s}(u, S_u, Y_u)dS_u + \int_0^t \gamma \frac{\partial h}{\partial y}(u, S_u, Y_u)dW_u^2.$$

La solution du problème de minimisation quadratique est donc donnée par :

$$\alpha_t^{mv} = \frac{\partial h}{\partial s}(t, S_t, Y_t), \quad 0 \leq t \leq T,$$

et la fonction valeur est

$$v_H(x) = (E[g(S_T)] - x)^2 + E\left[\int_0^T \left|\gamma \frac{\partial h}{\partial y}(t, S_t, Y_t)\right|^2 dt\right].$$

Dans la suite, on considère le cas général où S est une semimartingale (continue). Le principe de résolution est le suivant : On choisit un processus de richesse approprié qui sera lié à une probabilité martingale et obtenue par une méthode de dualité. On utilisera ensuite ce processus de richesse comme numéraire et par une méthode de changement de numéraire, on se ramène au cas martingale.

6.4.3 Probabilité martingale variance-optimale et numéraire quadratique

Considérons le problème de minimisation quadratique correspondant à $H = 1$ et $x = 0$:

$$v_1 = \min_{\alpha \in \mathcal{A}_2} E\left[1 - \int_0^T \alpha_t dS_t\right]^2. \tag{6.69}$$

Notons par $\alpha^{var} \in \mathcal{A}_2$ la solution de v_1 et X^{var} le processus de richesse :

$$X_t^{var} = 1 - \int_0^t \alpha_u^{var} dS_u, \quad 0 \leq t \leq T.$$

L'objectif de cette section est de montrer que X^{var} est relié de manière duale à une probabilité martingale et qu'il est alors strictement positif. On va utiliser la relation de dualité entre $\mathcal{G}_T$ et $\mathcal{M}_e^2$ et adopter une approche duale comme pour le problème de maximisation d'utilité mais adaptée au cadre $L^2(P)$.

Puisque $E^Q[1 - \int_0^T \alpha_t dS_t] = 1$ pour tout $\alpha \in \mathcal{A}_2$ et $Q \in \mathcal{M}_e^2$, on a par l'inégalité de Cauchy-Schwarz :

$$1 = \left(E\left[\frac{dQ}{dP}\left(1 - \int_0^T \alpha_t dS_t\right)\right]\right)^2$$
$$\leq E\left[\frac{dQ}{dP}\right]^2 E\left[1 - \int_0^T \alpha_t dS_t\right]^2.$$

On introduit alors le problème quadratique dual de v_1 :

$$\tilde{v}_1 = \inf_{Q \in \mathcal{M}_e^2} E\left[\frac{dQ}{dP}\right]^2, \tag{6.70}$$

de sorte que

$$1 \leq \inf_{Q \in \mathcal{M}_e^2} E\left[\frac{dQ}{dP}\right]^2 \min_{\alpha \in \mathcal{A}_2} E\left[1 - \int_0^T \alpha_t dS_t\right]^2 = \tilde{v}_1 \, v_1.$$

Nous allons montrer l'existence d'une solution P^{var} au problème dual $\tilde{v}_1$ et voir qu'elle est liée à la solution du problème primal v_1 par $\frac{dP^{var}}{dP} = \mathrm{Cte} X_T^{var}$. On aura alors égalité dans les inégalités précédentes : $1 = \tilde{v}_1 \, v_1$.

Théorème 6.4.7 *Il existe une solution unique à $\tilde{v}_1$, notée P^{var} et appelée probabilité martingale variance-optimale. La solution du problème primal v_1 est liée à celle du problème dual $\tilde{v}_1$ par :*

$$X_T^{var} = \frac{\frac{dP^{var}}{dP}}{E\left[\frac{dP^{var}}{dP}\right]^2} \quad i.e. \quad \frac{dP^{var}}{dP} = \frac{X_T^{var}}{E[X_T^{var}]}. \tag{6.71}$$

En particulier, on a

$$X_t^{var} > 0, \quad 0 \leq t \leq T, \; P \; p.s. \tag{6.72}$$

Remarque 6.4.10 Puisque X^{var} est une Q-martingale sous tout $Q \in \mathcal{M}_e^2$, en particulier sous P^{var}, on a d'après (6.71) et en notant par Z^{var} le processus de densité martingale de P^{var} :

$$X_t^{var} = \frac{E^Q[Z_T^{var}|\mathcal{F}_t]}{E[Z_T^{var}]^2} = \frac{E[(Z_T^{var})^2|\mathcal{F}_t]}{Z_t^{var} E[Z_T^{var}]^2}, \quad 0 \leq t \leq T, \quad \forall Q \in \mathcal{M}_e^2. \tag{6.73}$$

La relation (6.71) montre aussi que :

$$E[X_T^{var}]^2 \, E\left[\frac{dP^{var}}{dP}\right]^2 = 1.$$

Pour démontrer le théorème 6.4.7, on ne peut pas appliquer directement la méthode de dualité vue au paragraphe précédent pour la maximisation d'utilité concave définie sur $]0, +\infty[$ car ici la fonction d'utilité est $U(x) = -x^2$

définie sur tout $\mathbb{R}$. On va utiliser de façon cruciale la propriété quadratique du critère et les caractérisations par projection dans un espace de Hilbert. On note par $\mathcal{G}_T^\perp$ l'orthogonal de $\mathcal{G}_T$ dans $L^2(P)$:

$$\mathcal{G}_T^\perp = \left\{ Z_T \in L^2(P) \ : E\left[Z_T \int_0^T \alpha_t dS_t \right] = 0, \ \forall \alpha \in \mathcal{A}_2 \right\}.$$

Lemme 6.4.11 *Le processus de richesse associé à la solution $\alpha^{var} \in \mathcal{A}_2$ du problème v_1 est positif :*

$$X_t^{var} = 1 - \int_0^t \alpha_u^{var} dS_u \ \geq \ 0, \quad 0 \leq t \leq T, \ \ P \ p.s.$$

De plus, on a

$$X_T^{var} \ \in \ \mathcal{G}_T^\perp \ \ et \ \ E[X_T^{var}] \ = \ E[X_T^{var}]^2 \ > \ 0.$$

Preuve. Définissons le temps d'arrêt $\tau = \inf\{t \in [0,T] : X_t^{var} \leq 0\}$, avec la convention que $\inf \emptyset = +\infty$. Puisque S et donc X^{var} est continu, et $X_0^{var} = 1$, on a :

$$\begin{aligned} X_{\tau\wedge T}^{var} \ &= \ 0 \text{ sur } A := \{\tau \leq T\}, \\ X^{var} \ &> \ 0 \text{ sur } A^c = \{\tau = +\infty\}. \end{aligned}$$

Définissons le processus $\bar{\alpha}$ par $\bar{\alpha}_t = \alpha_t^{var}$ si $0 \leq t \leq \tau \wedge T$ et 0 sinon. Puisque $\alpha^{var} \in \mathcal{A}_2$, il est clair que $\bar{\alpha} \in \mathcal{A}_2$ et :

$$1 - \int_0^T \bar{\alpha}_t dS_t \ = \ X_{\tau\wedge T}^{var} \ = \ X_T^{var} 1_{A^c} \ \geq 0, \quad p.s.$$

On en déduit que $E[1 - \int_0^T \bar{\alpha}_t dS_t]^2 \leq E[X_T^{var}]^2$. Comme α^{var} est solution de v_1, on doit alors avoir $X_T^{var} = 1 - \int_0^T \bar{\alpha}_t dS_t \geq 0$. En notant que X^{var} est une martingale sous Q quelconque dans $\mathcal{M}_e^2$, on a :

$$X_t^{var} = E^Q\left[X_T^{var} | \mathcal{F}_t\right] \ \geq 0, \quad 0 \leq t \leq T, \ p.s.$$

D'autre part, la solution du problème v_1, qui est obtenue par projection de l'élément 1 sur le sous-espace vectoriel fermé $\mathcal{G}_T$ est caractérisée par le fait que $X_T^{var} \in \mathcal{G}_T^\perp$, i.e. :

$$E\left[X_T^{var} \int_0^T \alpha_t dS_t \right] = 0, \quad \forall \alpha \in \mathcal{A}_2. \tag{6.74}$$

Pour $\alpha = \alpha^{var}$, ceci implique en particulier que

$$E[X_T^{var}] = E[X_T^{var}]^2 \ > \ 0$$

puisque X_T^{var} ne peut être égal à zéro P p.s. d'après le fait que $E^Q[X_T^{var}] = 1$ pour $Q \in \mathcal{M}_e^2$. $\Box$

Dans la suite, on identifie de manière usuelle une probabilité absolument continue et sa densité de Radon-Nikodym.

Lemme 6.4.12

$$\mathcal{M}_e^2 = \left\{ Z_T \in L^2(P) \ : \ Z_T > 0, \ p.s., \ \ E[Z_T] = 1 \right\} \ \cap \ \mathcal{G}_T^{\perp}. \qquad (6.75)$$

Preuve. Notons par $\tilde{\mathcal{M}}_e^2$ le terme de droite dans (6.75). Par définition de $\mathcal{A}_2$, il est clair que $\mathcal{M}_e^2 \subset \tilde{\mathcal{M}}_e^2$. Soit $Z_T \in \tilde{\mathcal{M}}_e^2$ et $Q \sim P$ la probabilité associée de densité de Radon-Nikodym Z_T. Puisque S est continu (donc localement bornée), il existe une suite de temps d'arrêts τ_n tendant vers l'infini tel que le processus arrêté $S^{\tau_n} = (S_{t\wedge\tau_n})_{0\leq t\leq T}$ soit borné. (Par exemple $\tau_n = \inf\{t \geq 0 \ : |S_t| \geq n\}$). $\mathcal{A}_2$ contient donc les intégrands simples de la forme $\alpha_t = \theta 1_{]s\wedge\tau_n, u\wedge\tau_n]}(t)$ où $0 \leq s \leq u \leq T$ et θ, $\mathcal{F}_{s\wedge\tau_n}$-mesurable à valeurs dans $\mathbb{R}^n$. Puisque $Z_T \in \mathcal{G}_T^{\perp}$, on a :

$$0 \ = \ E[Z_T \int_0^T \theta 1_{]s\wedge\tau_n, u\wedge\tau_n]}(t) dS_t] = E^Q[\theta.(S_u^{\tau_n} - S_s^{\tau_n})], \qquad (6.76)$$

pour tous $0 \leq s \leq u \leq T$ et θ, $\mathcal{F}_{s\wedge\tau_n}$-mesurable. Avec la caractérisation des variables aléatoires $\mathcal{F}_{s\wedge\tau_n}$-mesurables (voir Proposition 1.1.2), la relation (6.76) est vraie aussi pour tout θ, $\mathcal{F}_s$-mesurable. Ceci prouve que S^{τ_n} est une Q-martingale, i.e. S est une Q-martingale locale, et donc $Q \in \mathcal{M}_e^2$. $\Box$

On introduit l'adhérence $\mathcal{M}^2$ de $\mathcal{M}_e^2$ dans $L^2(P)$ qui est donc donnée d'après le lemme précédent par :

$$\begin{aligned} \mathcal{M}^2 &= \left\{ Q \ll P \ : \frac{dQ}{dP} \in L^2(P) \ \text{ et } \ S \ \text{ est une } Q - \text{martingale locale} \right\} \\ &= \left\{ Z_T \in L^2(P) \ : \ Z_T \geq 0, \ \text{p.s.}, \ \ E[Z_T] = 1 \right\} \ \cap \ \mathcal{G}_T^{\perp}. \end{aligned}$$

Le problème dual s'écrit donc aussi :

$$\tilde{v}_1 = \inf_{Z_T \in \mathcal{M}^2} E[Z_T]^2. \qquad (6.77)$$

Lemme 6.4.13 *Il existe une solution unique $Z_T^{var} \in \mathcal{M}^2$ au problème dual* (6.77)*. Cette solution est liée au problème v_1 par :*

$$Z_T^{var} = E[Z_T^{var}]^2 X_T^{var}. \qquad (6.78)$$

Preuve. L'ensemble non vide $\mathcal{M}^2$ est clairement convexe et fermé dans $L^2(P)$. Le problème (6.77) qui est un problème de projection dans $L^2(P)$ de l'élément zéro sur $\mathcal{M}^2$ admet donc une solution unique Z_T^{var}. Rappelons que cette solution est caractérisée par le fait que $Z_T^{var} \in \mathcal{M}^2$ et

$$E[Z_T^{var}(Z_T^{var} - Z_T)] \leq 0, \quad \forall Z_T \in \mathcal{M}^2. \tag{6.79}$$

Considérons la variable aléatoire

$$\bar{Z}_T = \frac{X_T^{var}}{E[X_T^{var}]}. \tag{6.80}$$

D'après le lemme 6.4.11, $\bar{Z}_T \in \mathcal{M}^2$. De plus, par définition de $\mathcal{A}_2$ et $\bar{Z}_T$, on a pour tout $Z_T \in \mathcal{M}_e^2$:

$$E[Z_T \bar{Z}_T] = \frac{1}{E[X_T^{var}]} \ = \ E[\bar{Z}_T]^2.$$

Par densité de $\mathcal{M}_e^2$ dans $\mathcal{M}^2$, la relation ci-dessus est vraie pour tout $Z_T \in \mathcal{M}^2$. D'après la caractérisation (6.79) de Z_T^{var}, ceci prouve que $Z_T^{var} = \bar{Z}_T$, et conclut la preuve. □

Preuve du théorème 6.4.7.

Au vu des trois lemmes précédents, il reste à montrer que $Z_T^{var} > 0$ p.s. définissant ainsi une probabilité $P^{var} \in \mathcal{M}_e^2$. C'est le point délicat et technique dont la preuve peut être omise en première lecture.

Soit Z^{var} le processus positif de densité martingale de $P^{var} \in \mathcal{M}^2$: $Z_t^{var} = E[Z_T^{var}|\mathcal{F}_t] = E\left[\frac{dP^{var}}{dP}\Big|\mathcal{F}_t\right]$, $0 \leq t \leq T$. On veut montrer que P p.s., $Z_t^{var} > 0$ pour tout $t \in [0, T]$. Considérons alors le temps d'arrêt

$$\tau = \inf\{0 \leq t \leq T \ : \ Z_t^{var} = 0\}$$

avec la convention $\inf \emptyset = +\infty$. Considérons aussi le temps d'arrêt

$$\sigma = \inf\{0 \leq t \leq T \ : \ X_t^{var} = 0\}.$$

Fixons un élément $Q^0 \in \mathcal{M}_2^e$ et notons Z^0 son processus de densité martingale.

Sur l'ensemble $\{\sigma < \tau\} \subset \{\sigma < +\infty\}$, on a par la propriété de martingale du processus positif X^{var} sous Q^0 :

$$0 \ = \ X_\sigma^{var} = E^{Q^0}[X_T^{var}|\mathcal{F}_\sigma].$$

Puisque $X_T^{var} \geq 0$ d'après le lemme (6.4.11), ceci prouve que $X_\sigma^{var} = 0$ sur $\{\sigma < \tau\}$. Avec la relation (6.78), on a aussi $Z_T^{var} = 0$ sur $\{\sigma < \tau\}$. Par la propriété de martingale de Z^{var}, on a donc $Z_\sigma^{var} = 0$ sur $\{\sigma < \tau\}$. Ceci est clairement en contradiction avec la définition de τ à moins que $P\{\sigma < \tau\} = 0$.

Sur l'ensemble $\{\tau < \sigma\} \subset \{\tau < +\infty\}$, on a par la propriété de martingale du processus positif Z^{var} sous P :

$$0 \ = \ Z_\tau^{var} = E[Z_T^{var}|\mathcal{F}_\tau].$$

Puisque $Z_T^{var} \geq 0$, ceci prouve que $Z_T^{var} = 0$ sur $\{\tau < \sigma\}$ et donc aussi d'après (6.78) que $X_T^{var} = 0$ sur $\{\tau < \sigma\}$. Par la propriété de martingale de X^{var} sous

Q^0, on a $X_\tau^{var} = 0$ sur $\{\tau < \sigma\}$. Ceci est clairement en contradiction avec la définition de τ à moins que $P\{\tau < \sigma\} = 0$.

On en déduit donc que $\tau = \sigma$ p.s. et par continuité du processus positif X^{var}, τ est un temps d'arrêt prévisible : il existe une suite de temps d'arrêts $(\tau_n)_{n\geq 1}$ croissante avec $\tau_n < \tau$ sur $\{\tau > 0\}$, τ_n convergent vers τ (on dit que τ est annoncé par $(\tau_n)_n$). Il suffit en effet de prendre $\tau_n = \inf\{t \geq 0 \ : X_t^{var} \leq 1/n\} \wedge n$. Par la propriété de martingale du processus positif Z^{var} et puisque $Z_{\tau_n}^{var} > 0$, on a par l'inégalité de Cauchy-Schwarz :

$$\begin{aligned} 1 = E\left[\frac{Z_T^{var}}{Z_{\tau_n}^{var}} \middle| \mathcal{F}_{\tau_n} \right] &= E\left[\frac{Z_T^{var}}{Z_{\tau_n}^{var}} 1_{Z_\tau^{var} \neq 0} \middle| \mathcal{F}_{\tau_n} \right] \\ &\leq E\left[\left(\frac{Z_T^{var}}{Z_{\tau_n}^{var}} \right)^2 \middle| \mathcal{F}_{\tau_n} \right]^{\frac{1}{2}} E[1_{Z_\tau^{var} \neq 0} | \mathcal{F}_{\tau_n}]^{\frac{1}{2}}. \end{aligned} \tag{6.81}$$

Comme $E[1_{Z_\tau^{var} \neq 0} | \mathcal{F}_{\tau_n}]$ converge vers 0 sur l'ensemble $\{Z_\tau^{var} = 0\}$, l'inégalité (6.81) prouve que

$$E\left[\left(\frac{Z_T^{var}}{Z_{\tau_n}^{var}} \right)^2 \middle| \mathcal{F}_{\tau_n} \right] \to +\infty, \quad \text{sur} \quad \{Z_\tau^{var} = 0\}. \tag{6.82}$$

Puisque Z^0 est une P-martingale strictement positive telle que $Z_T^0 \in L^2(P)$, on a :

$$\sup_{0 \leq t \leq T} E\left[\left(\frac{Z_T^0}{Z_t^0} \right)^2 \middle| \mathcal{F}_t \right] < +\infty, \qquad p.s.$$

Supposons que $P[Z_\tau^{var} = 0] > 0$. D'après (6.82), on a alors que pour n assez grand, l'ensemble

$$A_n = \left\{ E\left[\left(\frac{Z_T^0}{Z_{\tau_n}} \right)^2 \middle| \mathcal{F}_{\tau_n} \right] \ < \ E\left[\left(\frac{Z_T^{var}}{Z_{\tau_n}^{var}} \right)^2 \middle| \mathcal{F}_{\tau_n} \right] \right\}$$

est non vide dans $\mathcal{F}_{\tau_n}$. Définissons alors la martingale :

$$Z_t = \begin{cases} Z_t^{var} & t < \tau_n \\ Z_t^0 \frac{Z_{\tau_n}^{var}}{Z_{\tau_n}^0} & t \geq \tau_n \ \text{ sur } A_n \\ Z_t^{var} & t \geq \tau_n \ \text{ en dehors de } A_n. \end{cases}$$

On vérifie aisément que ZS hérite de la propriété de martingale locale de $Z^{var}S$ et Z^0S, et donc $Z_T \in \mathcal{M}^2$. De plus, par construction, on a :

$$E\left[Z_T\right]^2 < E\left[Z_T^{var}\right]^2,$$

ce qui est en contradiction avec la définition de Z_T^{var}. On conclut alors que $P[Z_\tau^{var} = 0] = 0$ et donc que $Z_t^{var} > 0$, pour tout $t \in [0, T]$, P p.s. Finalement la stricte positivité de X^{var} découle de la stricte positivité de X_T^{var} et de la propriété de martingale de X^{var} sous $Q^0 \sim P$. □

Le processus de richesse X^{var} strictement positif est appelé numéraire quadratique et est utilisé dans le paragraphe suivant pour la résolution du problème moyenne-variance.

6.4.4 Résolution du problème par changement de numéraire

On utilise X^{var} comme numéraire, c'est à dire qu'on définit un processus de prix S^{var} à valeurs dans $\mathbb{R}^{n+1}$ par :

$$S^{var,0} \;=\; \frac{1}{X^{var}} \text{ et } S^{var,i} \;=\; \frac{S^i}{X^{var}},\; i = 1, \ldots, n.$$

Rappelons aussi que pour tout $Q \in \mathcal{M}_e^2$, le processus X^{var} est une Q-martingale de valeur initiale 1. On définit alors l'ensemble des probabilités de processus de densité martingale X^{var} par rapport à une probabilité Q dans $\mathcal{M}_e^2$:

$$\mathcal{M}_e^{2,var} = \Big\{ Q^{var} \text{ probabilité sur } (\Omega, \mathcal{F}_T) \;: \exists Q \in \mathcal{M}_e^2 \\ \left.\frac{dQ^{var}}{dQ}\right|_{\mathcal{F}_t} \;=\; X_t^{var},\;\; 0 \leq t \leq T. \Big\}.$$

Puisque par définition, $\mathcal{M}_e^2$ est l'ensemble des probabilités $Q \sim P$ de densité de Radon-Nikodym dans $L^2(P)$ et rendant martingale locale S, on remarque aisément par la formule de Bayes que $\mathcal{M}_e^{2,var}$ s'écrit aussi :

$$\mathcal{M}_e^{2,var} = \Big\{ Q^{var} \sim P \;\; : \frac{1}{X_T^{var}} \frac{dQ^{var}}{dP} \in L^2(P) \text{ et} \\ S^{var} \text{ est une } Q^{var} - \text{martingale locale} \Big\}. \quad (6.83)$$

On introduit alors l'ensemble admissible des intégrands par rapport à S^{var} :

$$\Phi_2^{var} = \Big\{ \phi \in L(S^{var}) \;:\; X_T^{var} \int_0^T \phi_t dS_t^{var} \in L^2(P) \quad \text{et} \\ \int \phi dS^{var} \;\text{ est une }\; Q^{var} - \text{martingale pour tout } Q^{var} \in \mathcal{M}_e^{2,var} \Big\}.$$

Nous établissons d'abord un résultat général d'invariance des intégrales stochastiques par changement de numéraire.

Proposition 6.4.6 *Pour tout $x \in \mathbb{R}$, on a :*

$$\left\{ x + \int_0^T \alpha_t dS_t : \alpha \in \mathcal{A}_2 \right\} = \left\{ X_T^{var} \left(x + \int_0^T \phi_t dS_t^{var} \right) : \phi \in \Phi_2^{var} \right\}. \quad (6.84)$$

De plus, la relation de correspondance entre $\alpha = (\alpha^1, \ldots, \alpha^n) \in \mathcal{A}_2$ *et* $\phi = (\phi^0, \ldots, \phi^n) \in \Phi_2^{var}$ *est donnée par* $\phi = F_x^{var}(\alpha)$ *où* $F_x^{var} : \mathcal{A}_2 \to \Phi_2^{var}$ *est défini par :*

$$\phi^0 \;=\; x + \int \alpha dS - \alpha.S \text{ et } \phi^i \;=\; \alpha^i, \quad i = 1, \ldots, n, \tag{6.85}$$

et $\alpha = F_x^{-1,var}(\phi)$ *où* $F_x^{-1,var} : \Phi_2^{var} \to \mathcal{A}_2$ *est défini par :*

$$\alpha^i = \phi^i - \alpha^{var,i}\left(x + \int \phi dS^{var} - \phi.S^{var}\right), \quad i = 1, \ldots, n. \tag{6.86}$$

Preuve. La preuve est basée sur le formule du produit d'Itô, le point technique concernant les problèmes d'intégrabilité sur les intégrands.
(1) Par le produit d'Itô, on a :

$$d\left(\frac{S}{X^{var}}\right) = Sd\left(\frac{1}{X^{var}}\right) + \frac{1}{X^{var}}dS + d < S, \frac{1}{X^{var}} > . \tag{6.87}$$

Soit $\alpha \in \mathcal{A}_2$ et considérons l'intégrand borné tronqué $\alpha^{(n)} = \alpha 1_{|\alpha| \leq n}$ qui est intégrable par rapport à S/X^{var}, $1/X^{var}$ et $< S, \frac{1}{X^{var}} >$. On a alors :

$$\int \alpha^{(n)} d\left(\frac{S}{X^{var}}\right) = \int \alpha^{(n)} Sd\left(\frac{1}{X^{var}}\right) + \int \alpha^{(n)} \frac{1}{X^{var}} dS + \int \alpha^{(n)} d < S, \frac{1}{X^{var}} > . \tag{6.88}$$

Notons $X^{x,\alpha^{(n)}} = x + \int \alpha^{(n)} dS$. Alors par la formule d'Itô et (6.88), on a :

$$\begin{aligned} d\left(\frac{X^{x,\alpha^{(n)}}}{X^{var}}\right) &= X^{x,\alpha^{(n)}} d\left(\frac{1}{X^{var}}\right) + \frac{1}{X^{var}}\alpha^{(n)} dS + \alpha^{(n)} d < S, \frac{1}{X^{var}} > \\ &= \left(X^{x,\alpha^{(n)}} - \alpha^{(n)}.S\right) d\left(\frac{1}{X^{var}}\right) + \alpha^{(n)} d\left(\frac{S}{X^{var}}\right) \\ &= \phi^{(n)} dS^{var}, \end{aligned} \tag{6.89}$$

avec $\phi^{(n)} = F_x^{var}(\alpha^{(n)}) \in L(S^{var})$. La relation (6.89) montre que :

$$x + \int \alpha^{(n)} dS = X^{var}\left(x + \int \phi^{(n)} dS^{var}\right). \tag{6.90}$$

Puisque α est S-intégrable, i.e $\alpha \in L(S)$, on a que $\int \alpha^{(n)} dS$ converge vers $\int \alpha dS$ pour la topologie des semimartingales quand n tend vers l'infini. Ceci implique que $X^{x,\alpha^{(n)}}/X^{var}$ converge aussi pour la topologie des semimartingales. D'après (6.90), on en déduit que $\int \phi^{(n)} dS^{var}$ converge vers $\int \psi d\tilde{X}$ pour la topologie des semimartingales avec $\psi \in L(S^{var})$, car l'espace $\{\int \psi dS^{var}$:

$\psi \in L(S^{var})$ } est fermé pour la topologie des semimartingales. Comme $\phi^{(n)}$ converge p.s. vers $\phi = F_x^{var}(\alpha)$, on a $\psi = \phi$. On obtient donc en faisant tendre n vers l'infini dans (6.90) :

$$x + \int \alpha dS = X^{var}\left(x + \int \phi dS^{var}\right). \tag{6.91}$$

Puisque X_T^{var} et $\int_0^T \alpha_t dS_t \in L^2(P)$, on a d'après (6.91) : $X_T^{var} \int_0^T \phi_t dS_t^{var} \in L^2(P)$. Puisque $\int \alpha dS$ est une Q-martingale pour tout $Q \in \mathcal{M}_e^2$, on a par définition de $\mathcal{M}_e^{2,var}$, d'après (6.91) et la formule de Bayes que $\int \phi dS^{var}$ est une Q^{var}-martingale pour tout $Q^{var} \in \mathcal{M}_e^{2,var}$. Ainsi $\phi \in \varPhi_2^{var}$. L'inclusion $\subseteq$ dans (6.84) est prouvée.

(2) La preuve de la réciproque est similaire. Par le produit d'Itô, on a :

$$d(X^{var}X) = X^{var}dS^{var} + S^{var}dX^{var} + d < X^{var}, S^{var} > . \tag{6.92}$$

Soit $\phi \in \varPhi_2^{var}$ et considérons l'intégrand tronqué borné $\phi^{(n)} = \phi 1_{|\phi| \leq n}$. Alors d'après (6.92) et les définitions de S^{var} et X^{var}, on a :

$$\begin{aligned}
& d\left(X^{var}(x + \int \phi^{(n)} dS^{var})\right) \\
&= \left(x + \int \phi^{(n)} dS^{var}\right) dX^{var} + X^{var}\phi^{(n)} dS^{var} + \phi^{(n)} d < X^{var}, S^{var} > \\
&= \left(x + \int \phi^{(n)} dS^{var}\right) dX^{var} + \phi^{(n)} d(X^{var}S^{var}) - \phi^{(n)}.S^{var}dX^{var} \\
&= \alpha^{(n)} dS,
\end{aligned}$$

avec $\alpha^{(n)} = F_x^{-1,var}(\phi^{(n)}) \in L(S)$. Par des arguments similaires comme au point (1), on obtient en faisant tendre n vers l'infini que :

$$X^{var}\left(x + \int \phi dS^{var}\right) = x + \int \alpha dS, \tag{6.93}$$

avec $\alpha = F_x^{-1,var}(\phi) \in L(S)$. On vérifie aussi comme en (1) que $\alpha \in \mathcal{A}_2$ puisque $\phi \in \varPhi_2^{var}$. L'inclusion $\supseteq$ dans (6.84) est prouvée, ce qui termine la preuve. □

Remarquons que dans la preuve de la proposition 6.4.6, nous avons seulement utilisé la stricte positivité du processus $X^{var} = 1 - \int \alpha^{var} dS$. Le résultat précédent d'invariance de l'espace des intégrales stochastiques par changement de numéraire est en fait valide pour n'importe quel choix de numéraire $X^{num} = 1 - \int \alpha^{num} dS_t$, $\alpha^{num} \in \mathcal{A}_2$, avec $X_t^{num} > 0$, $0 \leq t \leq T$. Le choix particulier de $X^{num} = X^{var}$ solution du problème (6.69) est maintenant utilisé de façon cruciale pour la résolution du problème de minimisation quadratique. On va montrer grâce à ce changement de numéraire particulier, comment se ramèner à un problème moyenne-variance comme dans le cas martingale du paragraphe 6.4.2.

A la probabilité martingale variance-optimale $P^{var} \in \mathcal{M}_e^2$, on associe $P^{2var} \in \mathcal{M}_e^{2,var}$ définie par :

$$\frac{dP^{2var}}{dP^{var}} = X_T^{var}. \tag{6.94}$$

D'après la relation de dualité (6.71), la densité de Radon-Nikodym de P^{2var} par rapport à P est :

$$\frac{dP^{2var}}{dP} = E\left[\frac{dP^{var}}{dP}\right]^2 (X_T^{var})^2. \tag{6.95}$$

Comme $H \in L^2(P)$, la relation (6.95) implique en particulier que l'actif contingent normalisé $H^{var} = H/X_T^{var} \in L^2(P^{2var})$. Rappelons aussi que S^{var} est une martingale locale (continue) sous P^{2var} d'après la caractérisation (6.83) de $\mathcal{M}_e^{2,var}$. On peut donc appliquer le théorème de projection de Kunita-Watanabe de la P^{2var}-martingale de carré intégrable $H_t^{var} = E^{P^{2var}}[H^{var}|\mathcal{F}_t]$, $0 \leq t \leq T$, sur S^{var}, et on a alors :

$$E^{P^{2var}}\left[\frac{H}{X_T^{var}}\bigg|\mathcal{F}_t\right] = E^{P^{2var}}\left[\frac{H}{X_T^{var}}\right] + \int_0^t \phi_u^H dS_u^{var} + R_t^{var,H}, 0 \leq t \leq T, \tag{6.96}$$

où $\phi^H \in L(S^{var})$ vérifie

$$E^{P^{2var}}\left[\int_0^T \phi_t^H dS_t^{var}\right]^2 = E^{P^{2var}}\left[\int_0^T (\phi_t^H)' d < S^{var} >_t \phi_t^H\right] < +\infty, \tag{6.97}$$

et $R^{var,H}$ est une P^{2var}-martingale de carré intégrable, orthogonale à S^{var}.

Théorème 6.4.8 *Pour tout $x \in \mathbb{R}$ et $H \in L^2(P)$, la solution α^{mv} de $v_H(x)$ est donnée par :*

$$\alpha^{mv} = F_x^{-1,var}(\phi^H), \tag{6.98}$$

où ϕ^H défini en (6.96) est dans Φ_2^{var} et $F_x^{-1,var}$ est défini à la proposition 6.4.6. De plus, on a :

$$v_H(x) = \frac{\left(E^{P^{var}}[H] - x\right)^2}{E\left[\frac{dP^{var}}{dP}\right]^2} + E\left[X_T^{var} R_T^{var,H}\right]^2. \tag{6.99}$$

Preuve. Par des arguments similaires à la preuve du théorème 6.4.6 (en utilisant les inégalités de Doob et de Cauchy-Schwarz), on voit que la condition d'intégrabilité (6.97) sur ϕ^H implique (en fait est équivalent à ce) que $\phi^H \in$

Φ_2^{var}. D'après la proposition 6.4.6 d'invariance par changement de numéraire et la relation (6.95) on a :

$$v_H(x) := \inf_{\alpha \in \mathcal{A}_2} E\left[H - x - \int_0^T \alpha_t dS_t\right]^2 \tag{6.100}$$
$$= \inf_{\phi \in \Phi_2^{var}} E\left[H - X_T^{var}\left(x + \int_0^T \phi_t dS_t^{var}\right)\right]^2$$
$$= \frac{1}{E\left[\frac{dP^{var}}{dP}\right]^2} \inf_{\phi \in \Phi_2^{var}} E^{P^{2var}}\left[\frac{H}{X_T^{var}} - x - \int_0^T \phi_t dS_t^{var}\right]^2, \tag{6.101}$$

et de plus la solution de (6.100) et celle de (6.101) sont reliées par la fonction de correspondance $F_x^{-1,var}$. Or le problème (6.101) est un problème de minimisation quadratique comme dans le cas martingale dont la solution est donnée par la décomposition de Kunita-Watanabe (6.96). Ceci prouve d'une part (6.98) et aussi :

$$v_H(x) = \frac{1}{E\left[\frac{dP^{var}}{dP}\right]^2}\left\{\left(E^{P^{2var}}\left[\frac{H}{X_T^{var}}\right] - x\right)^2 + E^{P^{2var}}[R_T^{var}]^2\right\}.$$

On obtient l'expression (6.99) de $v_H(x)$ avec (6.94) et (6.95). □

Remarque 6.4.11 La solution $x_{mv}(H)$ du problème $\inf_{x \in \mathbb{R}} v_H(x)$, appelée prix d'approximation quadratique de H, est donnée d'après (6.99) par :

$$x_{mv}(J) = E^{P^{var}}[H].$$

Le théorème ci-dessus indique donc que le problème de couverture quadratique peut être résolu en trois étapes :

1) Déterminer la solution du problème v_1 définissant le numéraire quadratique X^{var} ou de manière équivalente la solution de son problème dual $\tilde{v}_1$ définissant la probabilité martingale variance-optimale. Bien entendu, si S est une martingale sous P, la solution est triviale : $X^{var} = 1$ et $P^{var} = P$. Nous donnons dans le paragraphe suivant d'autres exemples de modèles où le calcul de X^{var} et P^{var} est explicite.

2) Changer de numéraire en normalisant le processus de prix S, l'actif contingent H et la probabilité martingale variance-optimale par X^{var}. On définit ainsi le processus de prix $S^{var} = (1/X^{var}, S/X^{var})$, l'actif contingent $H^{var} = H/X_T^{var}$ et la probabilité P^{2var} de densité de Radon-Nikodym par rapport à P^{var} : X_T^{var}. On projette ensuite selon la décomposition de Kunita-Watanabe la P^{2var}-martingale $E^{P^{2var}}[H^{var}|\mathcal{F}_t]$ sur S^{var}. Dans un cadre de modèle Markovien, par exemple de diffusion, cette décomposition est obtenue dans le cas régulier par la formule d'Itô. Dans un cadre plus général, l'intégrand de S^{var} dans la décomposition peut être exprimé à l'aide de la formule de Clark-Ocone ou la dérivée de Malliavin.

3) La solution du problème de couverture quadratique est donnée finalement par la relation de correspondance entre l'espace des intégrales stochastiques par rapport à S et l'espace des intégrales stochastiques par rapport à S^{var}.

6.4.5 Exemple

On se place dans le cadre du modèle et avec les notations du paragraphe 6.2.4. On a alors une description explicite de $\mathcal{M}_e^2$:

$$\mathcal{M}_e^2 = \left\{ P^\nu \; : \frac{dP^\nu}{dP} = Z_T^\nu, \;\; \nu \in K_m^2(\sigma) \right\} \tag{6.102}$$

où on rappelle que

$$Z_t^\nu = \exp\left(-\int_0^t (\lambda_u + \nu_u)' dW_u - \frac{1}{2}\int_0^t |\lambda_u|^2 + |\nu_u|^2 du \right), \quad 0 \leq t \leq T.$$

et $K_m^2(\sigma)$ est l'ensemble des éléments ν dans $K(\sigma)$ tels que Z^ν soit une martingale de carré intégrable.

On suppose dans cet exemple que la quantité

$$\hat{K}_T = \int_0^T |\lambda_t|^2 dt,$$

appelée ratio moyenne-variance, est déterministe. C'est une généralisation du cas S martingale locale sous P pour lequel $\hat{K}_T = 0$.

Considérons pour tout $\nu \in K_m^2(\sigma)$ la martingale locale exponentielle de Doléans-Dade

$$\xi_t^\nu = \exp\left(-2\int_0^t (\lambda_u + \nu_u)' dW_u - 2\int_0^t |\lambda_u|^2 + |\nu_u|^2 du \right), \quad 0 \leq t \leq T.$$

Il est clair que $|\xi_t^\nu| \leq |Z_t^\nu|^2$. Puisque Z^ν est une martingale de carré intégrable, on a $E[\sup_{0\leq t\leq T} |Z_t^\nu|^2] < +\infty$. Il en découle que ξ^ν est uniformément intégrable et est donc une martingale. On peut donc définir une probabilité Q^ν équivalente à P de processus de densité martingale ξ^ν. On a alors pour tout $\nu \in K_m^2(\sigma)$:

$$\begin{aligned}
E\left[\frac{dP^\nu}{dP}\right]^2 &= E\left[\exp\left(-2\int_0^T (\lambda_u + \nu_u)' dW_u - \int_0^T |\lambda_u|^2 + |\nu_u|^2 du \right) \right] \\
&= E^{Q^\nu}\left[\exp\left(\int_0^T |\lambda_u|^2 + |\nu_u|^2 du \right) \right] \\
&= \exp(\hat{K}_T) E^{Q^\nu}\left[\exp\left(\int_0^T |\nu_u|^2 du \right) \right] \\
&\geq \exp(\hat{K}_T),
\end{aligned} \tag{6.103}$$

où la troisième égalité découle du fait que $\hat{K}_T$ est déterministe. Notons que l'égalité dans (6.103) a lieu pour $\nu = 0$, ce qui prouve que la solution du problème définissant la probabilité martingale variance-optimale, défini d'après (6.102) par

$$\tilde{v}_1 = \inf_{\nu \in K_m^2(\sigma)} E\left[\frac{dP^\nu}{dP}\right]^2 \tag{6.104}$$

est atteinte pour $\nu = 0$. On a donc

$$P^{var} \;=\; P^0 \quad \text{et} \quad \tilde{v}_1 \;=\; E\left[\frac{dP^0}{dP}\right]^2 \;=\; \exp(\hat{K}_T).$$

On calcule le numéraire quadratique à l'aide de la l'expression (6.73) :

$$\begin{aligned} X_t^{var} &= \frac{1}{E\left[Z_T^0\right]^2} E^{P^0}[Z_T^0|\mathcal{F}_t] \\ &= E^{P^0}\left[\exp\left(-\int_0^T \lambda_u' dW_u^0 - \frac{1}{2}\int_0^T |\lambda_u|^2 du\right)\middle|\, \mathcal{F}_T\right], \end{aligned}$$

où $W^0 = W + \int \lambda dt$ est un P^0 mouvement Brownien. Notons que la condition de Novikov $E^{P^0}[\exp(\frac{1}{2}\int_0^T |\lambda_t|^2 dt)] = \exp(\hat{K}_T/2) < +\infty$ est satisfaite de sorte que :

$$X_t^{var} = \exp\left(-\int_0^t \lambda_u' dW_u^0 - \frac{1}{2}\int_0^t |\lambda_u|^2 du\right), \quad 0 \le t \le T.$$

Puisque $dX_t^{var} = -(\alpha_t^{var})'\sigma_t dW_t^0$, on a par identification et en se rappelant la définition $\lambda = \sigma'(\sigma\sigma')^{-1}\mu$:

$$\alpha^{var} = (\sigma\sigma')^{-1}\mu X^{var}.$$

Dans le cas général où $\hat{K}_T$ n'est pas déterministe, le problème dual (6.104) définissant la probabilité martingale variance-optimale est un problème de contrôle stochastique qui peut être étudié par les méthodes de la programmation dynamique ou les équations différentielles stochastiques rétrogrades. Nous donnons des références dans le dernier paragraphe de ce chapitre.

6.5 Commentaires bibliographiques

Le théorème de décomposition optionnelle des surmartingales a été démontré initialement dans le cadre de processus de prix d'Itô par El Karoui et Quenez [ElkQ95]. Il a ensuite été généralisé pour des processus semimartingales localement bornées par Kramkov [Kr96]. La version la plus générale pour des semimartingales quelconques est due à Föllmer et Kabanov [FoK98].

L'approche duale pour le problème de maximisation d'utilité a été formulée initialement dans un cadre de marché complet par Pliska [Pli86], Karatzas, Lehoczky et Shreve [KLS87] et Cox et Huang [CH89]. Elle a ensuite été étendue au cas des marchés incomplets pour des processus de prix d'Itô, indépendamment par Karatzas et al [KLSX91] et He et Pearson [HeP91]. L'étude générale pour des processus de prix semimartingale et sous l'hypothèse minimale d'élasticité asymptotique raisonnable est due à Kramkov et Schachermayer [KS99], [KS01]. Notre présentation au paragraphe 6.3 reprend les idées principales de leurs travaux. Le cas de maximisation d'utilité définie sur tout $\mathbb{R}$, typiquement la fonction d'utilité exponentielle, est étudié dans Delbaen et al. [DGRSSS02], Bellini et Frittelli [BF02] et Schachermayer [Scha01].

Le critère de couverture quadratique a été introduit par Föllmer et Sondermann [FoS86] dans le cas martingale. La méthode de résolution dans le cas général de processus de prix semimartingale continue, présentée au paragraphe 6.4, est due à Gouriéroux, Laurent et Pham [GLP98]. Rheinländer et Schweizer [RhS97] ont proposé une autre approche de résolution en se basant sur la décomposition dite de Föllmer-Schweizer. Le résultat d'existence de la probabilité martingale variance-optimale (notion introduite par Schweizer [Schw96]) équivalente à la probabilité initiale est due à Delbaen et Schachermayer [DS96]. L'exemple dans le paragraphe 6.4.5 est inspiré de Pham, Rheinländer et Schweizer [PRS98]. D'autres calculs explicites de la probabilité martingale variance-optimale et du numéraire quadratique dans des modèles à volatilité stochastique sont développés dans Laurent et Pham [LP99] et Biagini, Guasoni et Pratelli [BGP00].

A

Compléments d'intégration

On considère un espace de probabilité $(\Omega, \mathcal{F}, P)$ et $L^1 = L^1(\Omega, \mathcal{F}, P)$ est l'ensemble des variables aléatoires intégrables.

A.1 Uniforme intégrabilité

Définition A.1.1 *(Variables aléatoires uniformément intégrables)*
Soit $(f_i)_{i\in I}$ une famille de variables aléatoires dans L^1. On dit que $(f_i)_{i\in I}$ est uniformément intégrable si

$$\lim_{x\to+\infty} \sup_{i\in I} E[|f_i| 1_{|f_i|\geq x}] = 0.$$

Notons que toute famille de variables aléatoires majorées en module par une fonction intégrable fixe (en particulier, toute famille finie de variables aléatoires dans L^1) est uniformément intégrable.

Le théorème suivant généralise le théorème de convergence dominée.

Théorème A.1.1 *Soit $(f_n)_{n\geq 1}$ une suite de variables aléatoires dans L^1 convergent p.s. vers une variable aléatoire f. Alors f est intégrable et la convergence de (f_n) vers f a lieu dans L^1 si et seulement si la suite $(f_n)_{n\geq 1}$ est uniformément intégrable. Si les variables aléatoires f_n sont positives, il faut et il suffit que*

$$\lim_{n\to+\infty} E[f_n] = E[f].$$

Le corollaire suivant est utilisé dans la preuve du théorème 6.3.4.

Corollaire A.1.1 *Soit $(f_n)_{n\geq 1}$ une suite de variables aléatoires positives bornée dans L^1, i.e. $\sup_n E[f_n] < +\infty$, convergent p.s. vers une variable aléatoire positive f et telle que $\lim_{n\to+\infty} E[f_n] = E[f] + \delta$ avec $\delta > 0$. Il existe alors une sous-suite $(f_{n_k})_{k\geq 1}$ de $(f_n)_{n\geq 1}$ et une suite $(A_k)_{k\geq 1}$ d'ensembles disjoints deux à deux de $(\Omega, \mathcal{F})$ telles que :*

$$E[f_{n_k} 1_{A_k}] \geq \frac{\delta}{2}, \quad \forall k \geq 1.$$

Preuve. On pose $B_n = \{f_n \geq (f+\delta) \vee 1/\delta\}$. La suite $(f_n 1_{\Omega \setminus B_n})_{n \geq 1}$ est uniformément intégrable et converge p.s. vers f. Ceci implique que $E[f_n 1_{\Omega \setminus B_n}]$ converge vers $E[f]$ et donc $E[f_n 1_{B_n}]$ converge vers δ. Il existe donc $N = N(\delta)$ ≥ 1 tel que

$$E[f_n 1_{B_n}] \geq \frac{3\delta}{4}, \quad \forall n \geq N.$$

On pose $n_1 = N$. La suite $(f_{n_1} 1_{B_m})_{m \geq 1}$ est uniformément intégrable et converge p.s. vers 0. Il existe donc $n_2 \geq n_1 + 1$ tel que

$$E[f_{n_1} 1_{B_{n_2}}] \leq \frac{\delta}{4}.$$

On pose alors $A_1 = B_{n_1} \setminus B_{n_2}$ de sorte que

$$E[f_{n_1} 1_{A_1}] \geq E[f_{n_1} 1_{B_{n_1}}] - E[f_{n_1} 1_{B_{n_2}}] \ \geq \ \frac{\delta}{2}.$$

La suite $(f_{n_2} 1_{B_{n_1}} 1_{B_m})_{m \geq 1}$ est uniformément intégrable et converge p.s. vers 0. Il existe donc $n_3 \geq n_2 + 1$ tel que

$$E[f_{n_2} 1_{B_{n_1} \cup B_{n_3}}] \leq \frac{\delta}{4}.$$

On pose $A_2 = B_{n_2} \setminus (B_{n_1} \cup B_{n_3})$ de sorte que A_2 est disjoint de A_1 et

$$E[f_{n_2} 1_{A_2}] \geq \frac{\delta}{2}.$$

On continue ainsi de suite, i.e. à l'étape k : la suite $(f_{n_k} 1_{\cup_{i=1}^{k-1} B_{n_i}} 1_{B_m})_{m \geq 1}$ est uniformément intégrable et converge p.s. vers 0. Il existe donc $n_{k+1} \geq n_k + 1$ tel que

$$E[f_{n_k} 1_{\cup_{i=1}^{k-1} B_{n_i} \cup B_{n_{k+1}}}] \leq \frac{\delta}{4}.$$

On pose alors $A_k = B_{n_k} \setminus \left(\cup_{i=1}^{k-1} B_{n_i} \cup B_{n_{k+1}}\right)$ de sorte que A_k est disjoint de A_i, $i \leq k-1$, et

$$E[f_{n_k} 1_{A_k}] \geq \frac{\delta}{2}.$$

□

Le théorème suivant, dû à la Vallée-Poussin, donne une condition pratique pour montrer l'uniforme intégrabilité.

Théorème A.1.2 *(La Vallée-Poussin)*
Soit $(f_i)_{i\in I}$ une famille de variables aléatoires. On a équivalence entre
(1) $(f_i)_{i\in I}$ est uniformément intégrable
(2) Il existe une fonction φ définie sur $\mathbb{R}_+$, positive, $\lim_{x\to+\infty}\varphi(x)/x = +\infty$ et telle que

$$\sup_{i\in I} E[\varphi(|f_i|)] < +\infty.$$

En pratique, on utilise surtout l'implication (2) $\Longrightarrow$ (1). Par exemple, en prenant $\varphi(x) = x^2$, on a que toute famille de variables aléatoires bornée dans L^2 est uniformément intégrable. On peut trouver les preuves des théorèmes A.1.1 et A.1.2 par exemple dans le livre de Doob [Do94].

A.2 Essentiel supremum d'une famille de variables aléatoires

Définition A.2.2 *(Essentiel supremum)*
Soit $(f_i)_{i\in I}$ une famille de variables aléatoires réelles. L'essentiel supremum de cette famille, notée $\operatorname{ess\,sup}_{i\in I} f_i$ est une variable aléatoires $\hat{f}$ telle que :
(a) $f_i \leq \hat{f}$ p.s., pour tout $i \in I$
(b) Si g est une variable aléatoire vérifiant $f_i \leq g$ p.s., pour tout $i \in I$, alors $\hat{f} \leq g$ p.s.

Le résultat suivant est prouvé dans Neveu [Nev75].

Théorème A.2.3 *Soit $(f_i)_{i\in I}$ une famille de variables aléatoires réelles. Alors $\hat{f} = \operatorname{ess\,sup}_{i\in I} f_i$ existe et est unique. De plus, si la famille $(f_i)_{i\in I}$ est stable par supremum, i.e. pour tout i, j dans I, il existe k dans I tel que $f_i \vee f_j = f_k$, alors il existe une suite croissante $(f_{i_n})_{n\geq 1}$ dans $(f_i)_{i\in I}$ vérifiant*

$$\hat{f} = \lim_{n\to+\infty} \uparrow f_{i_n} \quad p.s.$$

On définit l'essentiel infimum d'une famille de variables aléatoires réelles $(f_i)_{i\in I}$ par : $\operatorname{ess\,inf}_{i\in I} f_i = -\operatorname{ess\,sup}_{i\in I}(-f_i)$.

A.3 Quelques théorèmes de compacité en probabilité

Ce premier résultat de compacité bien connu est dû à Komlos [Kom67].

Théorème A.3.4 *(Komlos)*
Soit $(f_n)_{n\geq 1}$ une suite de variables aléatoires bornée dans L^1. Alors il existe une sous-suite $(f_{n_k})_{k\geq 1}$ de $(f_n)_{n\in\mathbb{N}}$ et une variable aléatoire f dans L^1 telle que

$$\frac{1}{k}\sum_{j=1}^{k} f_{n_j} \to f \quad \textit{p.s. quand } k \textit{ tend vers l'infini.}$$

Le théorème suivant de compacité dans $L^0_+(\Omega, \mathcal{F}, P)$ est très utile pour obtenir des résultats d'existence dans les problèmes d'optimisation en finance. Il a été prouvé par Schachermayer, voir aussi l'appendice de Delbaen et Schachermayer [DS94].

Théorème A.3.5 *Soit $(f_n)_{n\geq 1}$ une suite de variables aléatoires dans L^0_+ $(\Omega, \mathcal{F}, P)$. Alors il existe une suite $g_n \in \mathrm{conv}(f_n, f_{n+1}, \ldots)$, i.e. $g_n = \sum_{k=n}^{N_n} \lambda_k f_k$, $\lambda_k \in [0,1]$ et $\sum_{k=n}^{N_n} \lambda_k = 1$, telle que la suite $(g_n)_{n\geq 1}$ converge p.s. vers une variable aléatoire g à valeurs dans $[0, +\infty]$.*

B

Considérations d'analyse convexe

Les références standard pour les rappels d'analyse convexe qui suivent sont les livres de Rockafellar [Ro70] et Ekeland et Temam [ET74]. Pour les besoins de ce livre, on se limitera essentiellement (sauf mention contraire) au cas dans $\mathbb{R}^d$. On note $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

B.1 Fonctions semicontinues, convexes

Etant donnée une fonction f de $\mathcal{O}$ ouvert de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$, on définit les fonctions f_* et $f^* : \mathcal{O} \to \bar{\mathbb{R}}$ par :

$$
\begin{aligned}
f_*(x) &= \liminf_{y \to x} f(y) \;:=\; \lim_{\varepsilon \to 0} \inf\{f(y) \;: y \in \mathcal{O}, |y-x| \leq \varepsilon\} \\
f^*(x) &= \limsup_{y \to x} f(y) \;:=\; \lim_{\varepsilon \to 0} \sup\{f(y) \;: y \in \mathcal{O}, |y-x| \leq \varepsilon\}.
\end{aligned}
$$

Définition B.1.1 *(Semicontinuité)*
Soit f une fonction de $\mathcal{O}$ ouvert de $\mathbb{R}^d$ dans $\bar{R}$. On dit que f est semicontinue inférieurement (s.c.i.) si l'une des conditions équivalentes suivantes est satisfaite :

(i) $\forall\, x \in \mathcal{O}$, $f(x) \leq \liminf_{n \to +\infty} f(x_n)$, *pour toute suite* $(x_n)_{n \geq 1}$ *convergent vers* x.

(ii) $\forall\, x \in \mathcal{O}$, $f(x) = f_*(x)$.

(iii) $\{x \in \mathcal{O} \;: f(x) \leq \lambda\}$ *est fermé pour tout* $\lambda \in \mathbb{R}$.

On dit que f est semicontinue supérieurement (s.c.s.) si $-f$ est semicontinue inférieurement.

Notons que f est continue sur $\mathcal{O}$ si et seulement si f est semicontinue inférieurement et supérieurement. La fonction f_* est appelée enveloppe semicontinue inférieure de f : c'est la plus grande fonction s.c.i. minorant f.

La fonction f^* est appelée enveloppe semicontinue supérieure de f : c'est la plus petite fonction s.c.s. majorant f.

Théorème B.1.1 *Une fonction s.c.i. (resp. s.c.s.) atteint son minimum (resp. maximum) sur tout compact.*

Etant donné un sous-ensemble convexe C de E espace vectoriel, on rappelle qu'une fonction f de C dans $\bar{\mathbb{R}}$ est dite convexe si pour tout $x, y \in C$, $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$. On dit que f est strictement convexe sur C si pour tous $x, y \in C$, $x \neq y$, $\lambda \in]0,1[$, $f(\lambda x + (1-\lambda)y) < \lambda f(x)$ $+ (1-\lambda) f(y)$. On dit que f est (strictement) concave si $-f$ est (strictement) convexe.

Le théorème minimax suivant est démontré dans Strasser [Str85], théorème 45.8.

Théorème B.1.2 *(Minimax)*
Soit $\mathcal{X}$ un sous-ensemble convexe de E espace vectoriel normé et compact pour la topologie faible $\sigma(E, E')$ et $\mathcal{Y}$ un sous-ensemble convexe d'un espace vectoriel. Soit $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ une fonction satisfaisant :

(1) $x \to f(x,y)$ est continue et concave sur $\mathcal{X}$ pour tout $y \in \mathcal{Y}$

(2) $y \to f(x,y)$ est convexe sur $\mathcal{Y}$ pour tout $x \in \mathcal{Y}$.

Alors on a

$$\sup_{x \in \mathcal{X}} \inf_{y \in \mathcal{Y}} f(x,y) = \inf_{y \in \mathcal{Y}} \sup_{x \in \mathcal{X}} f(x,y).$$

Dans la suite, on se limitera au cas $E = \mathbb{R}^d$. Etant donnée une fonction convexe f de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$, on définit son domaine

$$\mathrm{dom}(f) = \left\{ x \in \mathbb{R}^d \ : f(x) < +\infty \right\},$$

qui est un ensemble convexe de $\mathbb{R}^d$. On dit qu'une fonction convexe f de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ est *propre* si elle ne prend jamais la valeur $-\infty$ et si $\mathrm{dom}(f) \neq \emptyset$.

On a le résultat suivant de continuité des fonctions convexes.

Proposition B.1.1 *Une fonction convexe propre de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ est continue sur l'intérieur de son domaine.*

On s'intéresse à la différentiabilité des fonctions convexes.

Définition B.1.2 *(Sous-différentiel)*
Etant donnée une fonction convexe f de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$, on appelle sous-différentiel de f en $x \in \mathbb{R}^d$, et on note $\partial f(x)$, l'ensemble des vecteurs y de $\mathbb{R}^d$ tel que :

$$f(x) + y.(z - x) \leq f(z), \quad \forall z \in \mathbb{R}^d.$$

Proposition B.1.2 *Soit f une fonction convexe de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$.*
(1) Si f est finie et continue en $x \in \mathbb{R}^d$ alors $\partial f(x) \neq \emptyset$.
(2) f est finie et différentiable en $x \in \mathbb{R}^d$ de gradient $Df(x)$ si et seulement si $\partial f(x)$ est réduit à un singleton et dans ce cas $\partial f(x) = \{Df(x)\}$.

B.2 Transformée de Fenchel-Legendre

Définition B.2.3 *(Fonctions polaires)*
Etant donnée une fonction f de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$, on appelle fonction polaire (ou conjuguée) de f la fonction $\tilde{f}$ de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ définie par :

$$\tilde{f}(y) = \sup_{x \in \mathbb{R}^d} [x.y - f(x)], \quad y \in \mathbb{R}^d.$$

Lorsque f est convexe, on dit aussi que $\tilde{f}$ est la transformée de Fenchel-Legendre de f. Il est clair que dans la définition de $\tilde{f}$, on peut se limiter dans le supremum aux x dans le domaine de f. La fonction polaire $\tilde{f}$ est défini comme le supremum point par point x des fonctions affines $y \to x.y - f(x)$. C'est donc une fonction convexe sur $\mathbb{R}^d$.

On peut également définir la fonction polaire d'une fonction polaire. On a le résultat de bipolarité suivant :

Théorème B.2.3 *(Fenchel-Moreau)*
Soit f une fonction convexe propre s.c.i. de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ et $\tilde{f}$ sa transformée de Fenchel-Legendre. Alors

$$f(x) = \sup_{y \in \mathbb{R}^d} \left[x.y - \tilde{f}(y)\right], \quad x \in \mathbb{R}^d.$$

On établit le lien entre différentiabilité et fonctions polaires.

Proposition B.2.3 *Soit f une fonction convexe propre s.c.i. de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ et $\tilde{f}$ sa transformée de Fenchel-Legendre. Alors pour tout $x, y \in \mathbb{R}^d$, on a les équivalences suivantes :*

$$y \in \partial f(x) \iff x \in \partial \tilde{f}(y) \iff f(x) = x.y - \tilde{f}(y).$$

Proposition B.2.4 *Soit f une fonction convexe propre s.c.i. de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$, strictement convexe sur $\mathrm{int}(\mathrm{dom}(f))$. Alors sa transformée de Fenchel-Legendre $\tilde{f}$ est différentiable sur $\mathrm{int}(\mathrm{dom}(\tilde{f}))$. Si de plus, f est différentiable sur $\mathrm{int}(\mathrm{dom}(f))$ alors le gradient de f, Df, est une bijection de $\mathrm{int}(\mathrm{dom}(f))$ sur $\mathrm{int}(\mathrm{dom}(f))$ avec $Df = (D\tilde{f})^{-1}$ et $\tilde{f}$ est strictement convexe sur $\mathrm{int}(\mathrm{dom}(\tilde{f}))$.*

B.3 Exemple dans $\mathbb{R}$

Dans le chapitre 6, section 6.3, on rencontre la situation suivante. On a une fonction $u :]0,+\infty[\rightarrow \mathbb{R}$, croissante, concave sur $]0,+\infty[$ et on considère la fonction $\tilde{u} :]0,+\infty[\rightarrow R \cup \{+\infty\}$ définie par

$$\tilde{u}(y) = \sup_{x>0}[u(x) - xy], \quad y > 0.$$

$\tilde{u}$ est une fonction décroissante, convexe sur $]0,+\infty[$ et on note $\mathrm{dom}(\tilde{u}) = \{y > 0 \ : \tilde{v}(y) < +\infty\}$. La proposition suivante collecte certains résultats utilisés dans la section 6.3.

Proposition B.3.5 *On a la relation de conjugaison*

$$u(x) = \inf_{y>0}[\tilde{u}(y) + xy], \quad x > 0, \tag{B.1}$$

et

$$\tilde{u}(0) := \lim_{y\downarrow 0} \tilde{u}(y) = u(+\infty) := \lim_{x\rightarrow+\infty} u(x).$$

Supposons que u est strictement concave sur $]0,+\infty[$. Alors $\tilde{u}$ est dérivable sur int(dom($\tilde{u}$)). Si de plus, l'une des deux conditions équivalente suivante :

(i) u est dérivable sur $]0,+\infty[$

(ii) $\tilde{u}$ est strictement convexe sur int(dom($\tilde{u}$))

est satisfaite, alors la différentielle u' est une bijection de $]0,+\infty[$ sur int (dom($\tilde{u}$)) $\neq \emptyset$. avec $I := (u')^{-1} = -\tilde{u}'$ et on a

$$\tilde{u}(y) = u(I(y)) - yI(y), \quad \forall y \in \ \mathrm{int}(\mathrm{dom}(\tilde{u})).$$

Finalement, sous les conditions supplémentaires

$$u'(0) \ = \ +\infty \ \ \textit{et} \ \ u'(+\infty) = 0, \tag{B.2}$$

on a int(dom($\tilde{u}$)) = dom($\tilde{u}$) = $]0,+\infty[$.

Preuve. La fonction u étant concave et finie sur $]0,+\infty[$, elle est à croissance linéaire. On en déduit clairement que $\mathrm{dom}(\tilde{u}) \neq \emptyset$ et que son intérieur est de la forme :

$$\mathrm{int}(\mathrm{dom}(\tilde{u})) =]y_0,+\infty[,$$

où $y_0 = \inf\{y > 0 \ : \tilde{u}(y) < +\infty\}$.

Notons que $u(+\infty)$ existe dans $\bar{\mathbb{R}}$ par croissance de u sur $]0,+\infty[$. De même $\tilde{u}(0)$ existe dans $\bar{\mathbb{R}}$. D'après la définition de $\tilde{u}$, on a $\tilde{u}(y) \geq u(x) - xy$ pour tous $x, y > 0$, d'où l'on déduit $\tilde{u}(0) \geq u(+\infty)$. D'autre part, on a pour tout $y > 0$, $\tilde{u}(y) \leq \sup_{x>0} u(x) = u(+\infty)$ par croissance de u. Ceci prouve que $\tilde{u}(0) = u(+\infty)$.

Considérons la fonction f de $\mathbb{R}$ dans $\bar{\mathbb{R}}$ définie par :

$$f(x) = \begin{cases} -u(x), & x \geq 0 \\ +\infty, & x < 0. \end{cases}$$

f est une fonction convexe s.c.i., propre sur $\mathbb{R}$ et $\mathrm{int}(\mathrm{dom}(f)) =]0, +\infty[$. Sa transformée de Fenchel-Legendre est donnée par

$$\tilde{f}(y) \;=\; \sup_{x \in \mathbb{R}}[xy - f(x)] = \sup_{x>0}[xy + u(x)], \quad y \in \mathbb{R}.$$

Lorsque $y < 0$, on a par définition de $\tilde{u}$, $\tilde{f}(y) = \tilde{u}(-y)$. Lorsque $y > 0$, on a par la croissance de u, $\tilde{f}(y) \geq \lambda x_0 y + u(x_0)$ pour tout $\lambda > 1$ et $x_0 > 0$ fixé. Ceci prouve que $\tilde{f}(y) = +\infty$ pour $y > 0$. Pour $y = 0$, on a $\tilde{f}(0) = \sup_{x>0} u(x) = u(+\infty) = \tilde{u}(0)$. On a donc

$$\tilde{f}(y) = \begin{cases} \tilde{u}(-y), & y \leq 0 \\ +\infty, & y > 0, \end{cases}$$

et $\mathrm{int}(\mathrm{dom}(\tilde{f})) = -\mathrm{int}(\mathrm{dom}(\tilde{u})) =]-\infty, -y_0[$. D'après le théorème B.2.3 de bipolarité, on a alors pour tout $x \in \mathbb{R}$:

$$\begin{aligned} f(x) = \sup_{y \in \mathbb{R}}[xy - \tilde{f}(y)] \;&=\; \sup_{y<0}[xy - \tilde{u}(-y)] \\ = \sup_{y>0}[-xy - \tilde{u}(y)] \;&=\; -\inf_{y>0}[xy + \tilde{u}(y)]. \end{aligned}$$

On en déduit en particulier pour $x > 0$ la relation (B.1).

De plus, si u est strictement concave sur $]0, +\infty[$ alors f est strictement convexe sur $\mathrm{int}(\mathrm{dom}(f))$. D'après la proposition B.2.4, $\tilde{f}$ est alors dérivable sur $\mathrm{int}(\mathrm{dom}(\tilde{f}))$, i.e. $\tilde{u}$ est dérivable sur $\mathrm{int}(\mathrm{dom}(\tilde{u}))$.

L'équivalence des conditions (i) et (ii) de la proposition vient de l'équivalence entre

(i') f est dérivable sur $\mathrm{int}(\mathrm{dom}(f))$

(ii') $\tilde{f}$ est strictement convexe sur $\mathrm{int}(\mathrm{dom}(\tilde{f}))$.

Ceci est en effet une conséquence de la proposition B.2.4 appliquée d'une part à f et à $\tilde{f}$ qui est aussi convexe, propre et s.c.i. sur $\mathbb{R}$. Sous l'une de ces conditions, on a alors que f' est une bijection de $\mathrm{int}(\mathrm{dom}(f))$ sur $\mathrm{int}(\mathrm{dom}(\tilde{f}))$ avec $(f')^{-1} = \tilde{f}'$ et d'après la proposition B.2.3, on a pour tout $y \in \mathrm{int}(\mathrm{dom}(\tilde{f}))$:

$$\tilde{f}(y) = xy - f(x) \quad \text{où} \;\; x = \tilde{f}'(y), \text{ i.e. } y = f'(x).$$

Ceci montre les relations voulues sur u et $\tilde{u}$.

Finalement sous les conditions (B.2), l'image de $]0, +\infty[$ par f' est $]-\infty, 0[= \mathrm{int}(\mathrm{dom}(\tilde{f}))$ ce qui signifie que $\mathrm{int}(\mathrm{dom}(\tilde{u})) =]0, +\infty[= \mathrm{dom}(\tilde{u})$. □

Références

[ADEH99] Artzner P., Delbaen F., Eber J.M. et D. Heath (1999) : “Coherent measures of risk”, *Mathematical Finance*, 9, 203–228.

[ALP95] Avellaneda M., Levy A. et A. Paras (1995) : “Pricing and hedging derivative securities in markets with uncertain volatilities”, *Applied Mathematical Finance*, 2, 73–88.

[Ba95] Barles G. (1995) : Solutions de viscosité des équations d’Hamilton-Jacobi, Springer Verlag, Mathématiques et Applications.

[BElk04] Barrieu P. et N. El Karoui (2004) : “Optimal design of derivatives under dynamic risk measures”, Proceedings of the AMS, Spring 2004.

[BF02] Bellini F. et M. Frittelli (2002) : “On the existence of minimax martingale measures”, *Mathematical Finance*, 12, 1–21.

[Be57] Bellman R. (1957) : Dynamic programming, Princeton University Press.

[Ben81] Bensoussan A. (1981) : Lectures on Stochastic Control in Non linear filtering and stochastic control, Proceedings of the 3rd 1981 session, CIME, Lect. Notes in Maths. 972.

[Ben92] Bensoussan A. (1992) : Stochastic control of partially observable systems, Cambridge University Press.

[BL78] Bensoussan A. et J.L. Lions (1978) : Applications des inéquations variationnelles en contrôle stochastique, Dunod.

[BL82] Bensoussan A. et J.L. Lions (1982) : Contrôle impulsionnel et inéquations qausi-variationnelles contrôle stochastique, Dunod.

[BN91] Bensoussan A. et H. Nagai (1991) : “An ergodic control problem arising from the principal eigenfunction of an elliptic operator”, *J. Math. Soc. Japan*, 43, 49–65.

[BGP00] Biagini F., Guasoni P. and M. Pratelli (2000) : “Mean-Variance Hedging for Stochastic Volatility Models”, *Mathematical finance*, 10, 109–123.

[BP99] Bielecki T. et S. Pliska (1999) : “Risk-sensitive dynamic asset management”, *Applied Math. Optim.*, 39, 337–360.

[Bis76] Bismut J.M. (1976) : "Théorie probabiliste du contrôle des diffusions", *Mem. Amer. Math. Soc.*, 4, n^o 167.

[Bis78] Bismut J.M. (1978) : "Contrôle des systèmes linéaires quadratiques : applications de l'intégrale stochastique", Sem. Prob. XII, Lect. Notes in Math., 649, 180–264.

[Bor89] Borkar V. (1989) : Optimal control of diffusion processes, Pitman Research Notes in Math., 203.

[BM03] Bouchard B. et L. Mazliak (2003) : "A multidimensional bipolar theorem in $L^0(\mathbb{R}^d, \Omega, \mathcal{F}, P)$", *Stoch. Proc. Appli.*, 107, 213–231.

[CS93] Chatelain M. et C. Stricker (1995) : "Componentwise and vector stochastic integration with respect to certain multi-dimensional continuous local martingales", *Seminar on Stochastic Analysis, Random fields and Applications*, Ascona 1993, Prog. Proba., 36, 319–325.

[CST03] Cheridito P., Soner M. et N. Touzi (2003) : "The multi-dimensional super-replication problem under Gamma constraints", à paraıitre dans *Annales Inst. H. Poincaré, Anal. non linéaire.*

[CTZ03] Choulli T., M. Taksar et X.Y. Zhou (2003) : "A diffusion model for optimal dividend distribution with constraints on risk control", *SIAM J. Cont. Optim.*, 41, 1946–1979.

[CH89] Cox J. et C.F. Huang (1989) : "Optimal consumption and portfolio policies when asset prices follow a diffusion process", *Journal of Economic Theory*, 49, 33–83.

[CIL92] Crandall M., Ishii. H et P.L. Lions (1992) : "User's Guide to Viscosity Solutions of Second Order Partial Differential Equations", *Bull. Amer. Math. Soc.*, 27, 1–67.

[CPT99] Cvitanic J., H. Pham et N. Touzi (1999) : "Superreplication in stochastic volatility models under portfolio constraints", *Journal of Applied Probability*, 36, 523–545.

[Da77] Davis M. (1977) : Linear estimation and stochastic control, Chapman and Hall.

[DN90] Davis M. et A. Norman (1990) : "Portfolio selection with transaction costs", *Math. of Oper. Research*, 15, 676–713.

[DGRSSS02] Delbaen F., P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer et C. Stricker (2002) : "Exponential hedging and entropic penalties", *Mathematical Finance*, 12, 99–123.

[DS94] Delbaen F. et W. Schachermayer (1994) : "A general version of the fundamental theorem of asset pricing", *Math. Annalen*, 300, 463–520.

[DS96] Delbaen F. and W. Schachermayer (1996) : "The Variance-Optimal Martingale Measure for Continuous Processes", *Bernoulli*, Vol 2, 1, 81–105.

[DS98] Delbaen F. et W. Schachermayer (1998) : "The fundamental theorem of asset pricing for unbounded stochastic processes", *Math. Annalen*, 312, 215–250.

[DM75] Dellacherie C. et P.A. Meyer (1975) : Probabilités et Potentiel, ch. I à IV, Théorie des Martingales, Hermann.

[DM80] Dellacherie C. et P.A. Meyer (1980) : Probabilités et Potentiel, ch. V à VIII, Théorie des Martingales, Hermann.

[DeMa03] Denis L. et C. Martini (2003) : "A theoretical framework for the pricing of contingent claims in the presence of model uncertainty", prépublication de l'université du Maine.

[DP94] Dixit A. et R. Pindick (1994) : Investment under uncertainty, Princeton University Press.

[Do94] Doob J.L. (1994) : Measure theory, Springer Verlag.

[DZ00] Duckworth K. et M. Zervos (2000) : "An investment model with entry and exit decisions", *J. Applied Prob.*, 37, 547–559.

[Dyn63] Dynkin E. (1963) : "The optimal choice of the instant for stopping a Markov process", *Dolk. Acad. Nauk USSR*, 150, 238–240.

[ET74] Ekeland I. et R. Temam (1974) : Analyse convexe et problèmes variationnels, Dunod.

[Elk81] El Karoui N. (1981) : Les Aspects Probabilistes du Contrôle Stochastique, Lect. Notes in Math., 816, Springer Verlag.

[ElkNJ87] El Karoui N., D. Huu Nguyen et M. Jeanblanc-Picqué (1987) : "Compactification methods in the control of degenerate diffusions : existence of optimal controls", *Stochastics and Stochastics Reports*, 20, 169–219.

[ElkM97] El Karoui N. et L. Mazliak (editors) (1997) : Backward stochastic differential equations, Pitman research notes in mathematics series.

[ElkPQ97] El Karoui N., S. Peng et M.C. Quenez (1997) : "Backward stochastic differential equations in finance", *Mathematical Finance*, 7, 1–71.

[ElkQ95] El Karoui N. et M.C. Quenez (1995) : "Dynamic programming and pricing contingent claims in incomplete markets", *SIAM J. Cont. Optim.*, 33, 29–66.

[ElkR00] El Karoui N. et R. Rouge (2000) : "Pricing via utility maximization and entropy", *Mathematical Finance*, 7, 1–71.

[Em79] Emery M. (1979) : "Une topologie sur l'espace des semimartingales", *Sem. de Prob.*, XIII, vol. 721, 152–160, Lect. Notes in Math., Springer Verlag.

[FM95] Fleming W. et W. McEneaney (1995) : "Risk-sensitive control on an infinite horizon", *SIAM J. Cont. and Optim.*, 33, 1881–1915.

[FR75] Fleming W. et R. Rishel (1975) : Deterministic and stochastic optimal control, Springer Verlag.

[FP04] Fleming W. et T. Pang (2004) : "A stochastic control model of investment, production and consumption", Preprint, 2004.

[FS00] Fleming W. et S. Sheu (2000) : "Risk sensitive control and an optimal investment model", *Math. Finance*, 10, 197–213.

[FSo93] Fleming W. et M. Soner (1993) : Controlled Markov Processes and Viscosity Solutions, Springer Verlag.

[FoK98] Föllmer H. et Y. Kabanov (1998) : "Optional decomposition and Lagrange multipliers ", *Finance and Stochastics*, 1, 69–81.

[FoS02] Föllmer H. et A. Schied (2002) : Stochastic finance. An introduction in discrete-time., Berlin de Gruyter Studies in Mathematics.

[FoS86] Föllmer H. et D. Sondermann (1986) : "Hedging of non-redundant contingent claims", *Contributions to Mathematical Economics*, eds. A. Mas-Colell et W. Hildenbrand, North-Holland, 205–223.

[FoS91] Föllmer H. et M. Schweizer (1991) : "Hedging of Contingent Claims under Incomplete Information", *Applied Stochastic Analysis*, eds. M.H.A. Davis and R.J. Elliott, Stochastics Monographs vol. 5, Gordon and Breach, London/New York, 389–414.

[Fr75] Friedman A. (1975) : Stochastic Differential Equations and Applications, Vol. 1, Academic Press.

[FG04] Frittelli M. et M. Rosazza Gianin (2004) : "Dynamic convex risk measures", New risk measures in the 21th century, G. Szego ed., Wiley.

[Ga76] Galtchouk L. (1976) : "Représentation des martingales engendrées par un processus à accroissement indépendants", *Ann. Inst. H. Poincaré*, 12, 199–211.

[GS72] Gihman I. et A. Skorohod (1972) : Stochastic differential equations, Springer Verlag.

[GT85] Gilbarg D. et N. Trudinger (1985) : Elliptic differential equations of second order, Springer Verlag.

[GS89] Gilboa I. et D. Schmeidler (1989) : "Maxmin expected utility with non-unique prior", *J. Math. Econ.*, 18, 141–153.

[GLP98] Gouriéroux C., Laurent J.P et H. Pham (1998) : "Mean-Variance Hedging and Numéraire", *Mathematical Finance*, 8, 179–200.

[GV02] Gozzi F. et T. Vargiolu (2002) : "Superreplication of European multiasset derivatives with bounded stochastic volatility", *Mathematical Methods of Operations Research*, 55 (1), 69–91.

[Gu04] Gundel A. (2004) : "Robust utility maximization for complete and incomplete market models", Preprint Humbold-Universiät Berlin.

[GP05] Guo X. et H. Pham (2005) : "Optimal partially reversible investment with entry decision and general production function", *Stoc. Proc. Appli.*, vol. 115, 705–736.

[HL95] Hamadène S. et J.-P. Lepeltier (1995) : "Backward equations, stochastic control and zero-sum stochastic differential games", *Stochastics and stochastic Reports*, vol. 54, p.221–231.

[HP81] Harrison M. et S. Pliska (1981) : "Martingales and stochastic integrals in the theory of continuous trading", *Stoch. Proc. Appli.*, 11, 215–260.

[Ha04] Hata H. (2004) : "A risk-sensitive stochastic control approach to an optimal investment problem with partial information", Preprint, Osaka University.

[HeP91] He H. et N.D. Pearson (1991) : "Consumption and Portfolio Policies with Incomplete Markets and Short Sale Constraints", *Journal of Economic Theory*, 54, 259–305.

[Hi80] Hida T. (1980) : Brownian Motion, Springer Verlag.

[HIM04] Hu Y., P. Imkeller et M. Müller (2004) : "Utility maximization in incomplete markets", Preprint Humboldt Berlin.

[IW81] Ikeda N. et S. Watanabe (1981) : Stochastic differential equations and diffusion processes, North-Holland.

[Ish89] Ishii H. (1989) : "On uniqueness and existence of viscosity solutions of fully nonlineqr second order elliptic PDE's", Comm. Pure. Appl. Math., 42, 15–45.

[Ja93] Jacka S. (1993) : "Local times, optimal stopping and semimartingales", Annals of Probability, 21, 329–339.

[Jac79] Jacod J. (1979) : Calcul Stochastique et Problèmes de Martingales, Lect. Notes in Math., 714, Springer Verlag.

[JS95] Jeanblanc-Picqué M. et A. Shiryaev (1995) : "Optimization of the flow of dividends", *Russian Math. Surveys*, vol. 50, 257–277.

[Je88] Jensen R. (1988) : "The maximum principle for viscosity solutions of second order fully nonlinear partial differential equations", Arch. Rat. Mech. Anal., 101, 1–27.

[KS81] Kamien M. et N. Schwartz (1981) : Dynamic optimization, North Holland.

[Kar80] Karatzas I. (1980) : "On a stochastic representation for the principal eigenvalue of a second order differential equation", *Stochastics and Stochastics Reports*, 3, 305–321.

[KLS87] Karatzas I., Lehoczky J. et S. Shreve (1987) : "Optimal portfolio and consumption decisions for a small investor on a finite horizon", *SIAM J. Cont. Optim.*, 25, 297–323.

[KLSX91] Karatzas I., Lehoczky J., Shreve S. et G. Xu (1991) : "Martingale and duality methods for utility maximization in incomplete market", *SIAM J. Cont. Optim.*, 29, 702–730.

[KaSh88] Karatzas I. et S. Shreve (1988) : Brownian Motion and Stochastic Calculus, Springer Verlag.

[KaSh98] Karatzas I. et S. Shreve (1998) : Methods of Mathematical Finance, Springer Verlag.

[Ko00] Kobylanski M. (2000) : "Backward stochastic differential equations and partial differential equations with quadratic growth", *Annals of Probability*, 28, 558–602.

[KT02] Kohlmann M. et S. Tang (2002) : "Global adapted solution of one-dimensional backward stochastic differential Riccati equations, with application to the mean-variance hedging", *Stochastic Process. Appl.*, 97, 255–288.

[KZ00] Kohlmann M. et X.Y. Zhou (2000) : "Relationship between backward stochastic differential equations and stochastic controls : a linear-quadratic approach", *SIAM Journal on Control and Optimization*, 38, 1392–1407.

[Kom67] Komlos J. (1967) : "A generalisation of a theorem of Steinhaus", *Acta Math. Acad. Sci. Hung.*, 18, 217–229.

[Kor97] Korn R. (1997) : Stochastic Models for Optimal Investment and Risk Management in Continuous Time, World Scientific.

[Kr96] Kramkov D. (1996) : "Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets", *Prob. Theor. Rel. Fields*, 145, 459–480.

[KS99] Kramkov D. et W. Schachermayer (1999) : "The asymptotic elasticity of utility functions and optimal investment in incomplete markets", *Annals of Applied Probability*, 9, 904–950.

[KS01] Kramkov D. et W. Schachermayer (2001) : "Necessary and sufficient conditions in the problem of optimal investment in incomplete markets", *Annals of Applied Probability*, 13, 1504–1516.

[Kry80] Krylov N. (1980) : Controlled Diffusion Processes, Springer Verlag.

[Kry87] Krylov N. (1987) : Nonlinear Elliptic and Parabolic Equations of Second Order, Boston, D. Reidel.

[KW67] Kunita H. et S. Watanabe (1967) : "On square integrable martingales", *Nagoya Math. J.*, 30, 209–245.

[Ku75] Kushner H. (1975) : "Existence results for optimal stochastic controls", *J. Optim. Theory and Appl.*, 15, 347–359.

[L98] Lamberton D. (1998) : "American options", *Statistics and Finance*, D. Hand, S. Jacka eds. Arnold.

[La74] Lasry J.M. (1974) : Thèse détat, Université Paris Dauphine.

[LP99] Laurent J.P. and H. Pham (1999) : "Dynamic Programming and Mean-Variance Hedging", *Finance and stochastics*, 23, 83–110.

[LeG89] Le Gall J.F. (1989) : "Introduction au mouvement Brownien", *Gazette des mathématiciens*, 40, 43–64. Soc. Math. France.

[Lio83] Lions P.L. (1983) : "Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations"", *Comm. P.D.E*, 8, Part I, 1101-1134, Part II, 1229–1276.

[LS84] Lions P.L. and A. Snitzman (1984) : "Stochastic differential equations with reflecting boundary conditions". *Comm. Pure. Appl. Math.*, Vol. **37**, 511–537.

[MY00] Ma J. et J. Yong (2000) : Forward-Backward stochastic differential equations and their applications, Lect. Notes in Math., 1702.

[Ma52] Markowitz H. (1952) : "Portfolio selection", *J. of Finance*, 7, 77–91.

[Mer69] Merton R. (1969) : "Lifetime portfolio selection under uncertainty : the continuous time case", *Rev. Econ. Stat.*, 51, 239–265.

[Mer73] Merton R. (1973) : "Optimum consumption and portfolio rules in a continuous-time model", *J. Econ. Theory*, 3, 373–413.

[Nev75] Neveu J. (1975) : Martingales à temps discret, Masson.

[Nis81] Nisio M. (1981) : Lectures on Stochastic Control Theory, ISI Lect. Notes, 9, Kaigai Publ. Osaka.

[Oks00] Oksendal B. (2000) : Stochastic differential equations : an introduction with applications, 6th edition, Springer Verlag.

[OS02] Oksendal B. et A. Sulem (2002) : "Optimal Consumption and Portfolio with both fixed and proportional transaction costs : A Combined Stochastic Control and Impulse Control Model", *SIAM J. Control and Optim.*, 40, 1765–1790.

[OS04] Oksendal B. et A. Sulem (2004) : Applied stochastic control of jump diffusion, Springer Verlag.

[Pa98] Pardoux E. (1998) : "Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order", Stochastic analysis and related topics, VI (Geilo, 1996), 79–127, Progr. Probab., 42, Birkhäuser Boston, Boston, MA, 1998.

[PaPe90] Pardoux E. et S. Peng (1990) : "Adapted solutions of a backward stochastic differential equation", *Systems and Control Letters*, 14, 55–61.

[Pha03a] Pham H. (2003a) : "A large deviations approach to optimal long term investment'", *Finance and Stochastics*, 7, 169–195.

[Pha03b] Pham H. (2003b) : "A risk-sensitive control dual approach to a large deviations control problem", *Systems and Control Letters*, 49, 295–309.

[PRS98] Pham H., Rheinländer T. et M. Schweizer (1998) : "Mean-Variance Hedging for Continuous Processes : New Proofs and Examples", *Finance and Stochastics*, 2, 173–198.

[Pli86] Pliska S. (1986) : "A stochastic calculus model of continuous trading : optimal portfolios", *Math. Oper. Res.*, 11, 371–382.

[Pro90] Protter P. (1990) : Stochastic integration and differential equations, Springer Verlag.

[Rev94] Revuz D. (1994) : Mesures et Intégration, Hermann.

[Rev97] Revuz D. (1997) : Probabilités, Hermann.

[ReY91] Revuz D. et M. Yor (1991) : Continuous Martingale and Brownian Motion, Springer-Verlag.

[RhS97] Rheinländer T. et M. Schweizer (1997) : "On L^2-Projections on a Space of Stochastic Integrals", *Annals of Probability*, 25, 1810–1831.

[Ro70] Rockafellar R. (1970) : Convex Analysis, Princeton University Press.

[Scha01] Schachermayer W. (2001) : "Optimal investment in incomplete markets when wealth may become negative", *Annals of Applied Probability*, 11, 694–734.

[Schi03] Schied A. (2003) : "Optimal investment for robust utility functionals in complete markets", Preprint TU Berlin.

[Schw96] Schweizer M. (1996) : Approximation pricing and the variance-optimal martingale measure, *Annals of Probability*, 24, 206–236.

[SS87] Seierstad A. et K. Sydsaeter (1987) : Optimal control theory with economic applications, North Holland.

[Se02] Sekine J. (2002) : "Exponential hedging by solving a backward stochastic differential equation : an alternative approach", preprint Osaka University.

[SZ94] Sethi S.P. et Zhang Q. (1994) : Hierarchical Decision Making in Stochastic Manufacturing Systems, in series Systems and Control : Foundations and Applications, Birkhauser Boston, Cambridge.

[Sh78] Shiryayev A. (1978) : Optimal stopping rules, Springer Verlag.

[ShSo94] Shreve S. et M. Soner (1994) : "Optimal investment and consumption with transaction costs", *Annals of Applied Probability*, 4, 609–692.

[ST00] Soner M. et N. Touzi (2000) : "Super replication under gamma constraints", SIAM Journal on Control and Optimization, 39 (1), 73–96.

[ST02] Soner M. et N. Touzi (2002) : "Stochastic target problems, dynamic programming and viscosity solutions", SIAM Journal on Control and Optimization, 41, 404–424.

[Str85] Strasser H. (1985) : Mathematical theory of statistics : statistical experiments and asymptotic decision theory, W. de Gruyter.

[Van76] Van Moerbecke P. (1976) : "On optimal stopping and free boundary problems", *Arch. Rat. Mech.*.

[YZ00] Yong J. et X.Y. Zhou (2000) : Stochastic controls, Hamiltonian systems and HJB equations, Springer Verlag.

[Yo78] Yor M. (1978) : "Sous-espaces denses dans L^1 ou H^1 et représentation des martingales", Séminaire de Probabilités XII, Lect. Notes in Math., 649, Springer Verlag, 265–309.

[Zar88] Zariphopoulou T. (1988) : Optimal investment-consumption models with constraints, Brown University, Phd.

[ZL00] Zhou X.Y. et D. Li (2000) : "Continuous-time mean-variance portfolio selection : a stochastic LQ framework", *Applied Mathematics and Optimization*, 42, 19–33.

Index

Actif
  contingent, 36
  risqué, 33
Adapté, 2
Adjoint
  équation, 105
  processus, 96

Bayes (formule), 21

Changement
  de numéraire, 152
  de probabilité, 21
Comparaison (principe, théorème), 68, 97
Conditions habituelles, 2
Consommation, 34
Contrôle, 31, 42
  de processus de diffusion, 41
  markovien, 43
  optimal, 44
  singulier, 62
Couverture
  moyenne-variance, 33
  quadratique, 35
Crochet oblique, 10

Doob-Meyer (décomposition), 11
Dualité convexe, 117

Equation aux dérivées partielles, 28
  elliptique, 56
  parabolique, 28
Equations différentielles stochastiques, 24
  rétrogrades, 93
Espace de probabilité, 1
  filtré, 2
Essentiel infimum, 169

Fenchel-Legendre (transformée), 173
Feynman-Kac (formule), 28, 98
Filtration, 2
  brownienne, 5
  naturelle, 2
Fonction
  convexe, 171
  semicontinue, 171
  test, 67
Fonction d'utilité, 33
Fonction valeur, 32, 44

Générateur infinitésimal, 28
Girsanov (théorème), 21

Hamilton-Jacobi-Bellman (équation), 47
Hamiltonien, 48
Horizon
  fini, 42
  infini, 44

Inégalité
  de Burkholder-Davis-Gundy, 11
  de Doob, 8
Intégrale stochastique, 16
Investissement, 35
Itô (formule), 19

Kunita-Watanabe (décomposition), 20

Marché
  complet, 24, 145
  incomplet, 146
Martingale, 6
  locale, 8
Maximisation d'utilité, 108, 128
Merton (problème de gestion de portefeuille), 33, 57
Modèle
  à volatilité incertaine, 36
  de Black-Scholes, 33
Mouvement brownien, 5

Numéraire, 158

Optionnel, 3

Portefeuille, 32
Prévisible, 3
Principe du maximum de Pontryagin, 105
Probabilité martingale, 120
  variance-optimal, 152
Processus, 2
  variation finie, 9
  càd-làg, 1
  croissant, 9
  densité martingale, 21
  Itô, 18
  stochastique, 1
Production, 34
Programmation dynamique (principe), 45
Progressif, 3

Radon-Nikodym (densité), 21
Richesse, 32

Semimartingale, 12
Solution
  classique, 67
  de viscosité, 67
  forte d'EDS, 24
Surmartingale, 6
Surréplication, 36
  coût de, 83
  problème de, 119
  représentation duale, 120

Temps d'arrêt, 4
Théorème
  décomposition optionnelle, 120
  Komlos, 169
  La Vallée-Poussin, 168
  minimax, 172
  représentation d'Itô, 23
  représentation des martingales, 20
Théorème de vérification, 52

Uniforme
  ellipticité, 29
  intégrabilité, 167

Variation
  finie, 9
  quadratique, 10
Viscosité (solution), 67

Déjà parus dans la même collection

1. T. Cazenave, A. Haraux : Introduction aux problèmes d'évolution semi-linéaires. 1990
2. P. Joly : Mise en œuvre de la méthode des éléments finis. 1990
3/4. E. Godlewski, P.-A. Raviart : Hyperbolic systems of conservation laws. 1991
5/6. Ph. Destuynder : Modélisation mécanique des milieux continus. 1991
7. J. C. Nedelec : Notions sur les techniques d'éléments finis. 1992
8. G. Robin : Algorithmique et cryptographie. 1992
9. D. Lamberton, B. Lapeyre : Introduction au calcul stochastique appliqué. 1992
10. C. Bernardi, Y. Maday : Approximations spectrales de problèmes aux limites elliptiques. 1992
11. V. Genon-Catalot, D. Picard : Eléments de statistique asymptotique. 1993
12. P. Dehornoy : Complexité et décidabilité. 1993
13. O. Kavian : Introduction à la théorie des points critiques. 1994
14. A. Bossavit : Électromagnétisme, en vue de la modélisation. 1994
15. R. Kh. Zeytounian : Modélisation asymptotique en mécanique des fluides Newtoniens. 1994
16. D. Bouche, F. Molinet : Méthodes asymptotiques en électromagnétisme. 1994
17. G. Barles : Solutions de viscosité des équations de Hamilton-Jacobi. 1994
18. Q. S. Nguyen : Stabilité des structures élastiques. 1995
19. F. Robert : Les systèmes dynamiques discrets. 1995
20. O. Papini, J. Wolfmann : Algèbre discrète et codes correcteurs. 1995
21. D. Collombier : Plans d'expérience factoriels. 1996
22. G. Gagneux, M. Madaune-Tort : Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière. 1996
23. M. Duflo : Algorithmes stochastiques. 1996
24. P. Destuynder, M. Salaun : Mathematical Analysis of Thin Plate Models. 1996
25. P. Rougee : Mécanique des grandes transformations. 1997
26. L. Hörmander : Lectures on Nonlinear Hyperbolic Differential Equations. 1997
27. J.F. Bonnans, J.C. Gilbert, C. Lemaréchal, C. Sagastizábal : Optimisation numérique. 1997
28. C. Cocozza-Thivent : Processus stochastiques et fiabilité des systèmes. 1997
29. B. Lapeyre, É. Pardoux, R. Sentis : Méthodes de Monte-Carlo pour les équations de transport et de diffusion. 1998
30. P. Sagaut : Introduction à la simulation des grandes échelles pour les écoulements de fluide incompressible. 1998
31. E. Rio : Théorie asymptotique des processus aléatoires faiblement dépendants. 1999
32. J. Moreau, P.-A. Doudin, P. Cazes (Eds.) : L'analyse des correspondances et les techniques connexes. 1999
33. B. Chalmond : Eléments de modélisation pour l'analyse d'images. 1999
34. J. Istas : Introduction aux modélisations mathématiques pour les sciences du vivant. 2000
35. P. Robert : Réseaux et files d'attente : méthodes probabilistes. 2000
36. A. Ern, J.-L. Guermond : Eléments finis : théorie, applications, mise en œuvre. 2001
37. S. Sorin : A First Course on Zero-Sum Repeated Games. 2002
38. J. F. Maurras : Programmation linéaire, complexité. 2002

39. B. Ycart : Modèles et algorithmes Markoviens. 2002
40. B. Bonnard, M. Chyba : Singular Trajectories and their Role in Control Theory. 2003
41. A. Tsybakov : Introdution à l'estimation non-paramétrique. 2003
42. J. Abdeljaoued, H. Lombardi : Méthodes matricielles – Introduction à la complexité algébrique. 2004
43. U. Boscain, B. Piccoli : Optimal Syntheses for Control Systems on 2-D Manifolds. 2004
44. L. Younes : Invariance, déformations et reconnaissance de formes. 2004
45. C. Bernardi, Y. Maday, F. Rapetti : Discrétisations variationnelles de problèmes aux limites elliptiques. 2004
46. J.-P. Françoise : Oscillations en biologie : Analyse qualitative et modèles. 2005
47. C. Le Bris : Systèmes multi-échelles : Modélisation et simulation. 2005
48. A. Henrot, M. Pierre : Variation et optimisation de formes : Une analyse géometric. 2005
49. B. Bidégaray-Fesquet : Hiérarchie de modèles en optique quantique : De Maxwell-Bloch à Schrödinger non-linéaire. 2005
50. R. Dáger, E. Zuazua : Wave Propagation, Observation and Control in $1-d$ Flexible Multi-Structures. 2005
51. B. Bonnard, L. Faubourg, E. Trélat Mécanique céleste et contrôle des véhicules spatiaux. 2005
52. F. Boyer, P. Fabrie : Eléments d'analyse pour l'étude de quelques modèles d'écoulements de fluides visqueux incompressibles. 2005
53. E. Cancès, C. L. Bris, Y. Maday : Méthodes mathématiques en chimie quantique. Une introduction. 2006
54. J-P. Dedieu : Points fixes, zeros et la methode de Newton. 2006
55. P. Lopez, A. S. Nouri : Théorie élémentaire et pratique de la commande par les régimes glissants. 2006
56. J. Cousteix, J. Mauss : Analyse asympotitque et couche limite. 2006
57. J.-F. Delmas, B. Jourdain : Modèles aléatoires. 2006
58. G. Allaire : Conception optimale de structures. 2007
59. M. Elkadi, B. Mourrain : Introduction à la résolution des systèmes polynomiaux. 2007
60. N. Caspard, B. Leclerc, B. Monjardet : Ensembles ordonnés finis : concepts, résultats et usages. 2007
61. H. Pham : Optimisation et contrôle stochastique appliqués à la finance. 2007