

Méthodes numériques en contrôle stochastique

Jean-Philippe Chancelier et Agnès Sulem

INRIA

Domaine de Voluceau Rocquencourt

78153 Le Chesnay Cedex

Tel: 01 39 63 55 69

CERMICS

École Nationale des Ponts et Chaussées

6 et 8 Av Blaise Pascal, Cite Descartes

Champs sur Marne, 77455 Marne La vallee cedex 02

22 février 2005

Table des matières

1	Chaînes de Markov	3
1.1	Chaînes de Markov	3
1.2	Calculs d'espérance et Équations de Kolmogorov	6
1.2.1	Évaluation d'espérance sur un horizon fini.	6
1.2.2	Évaluation d'espérance sur un horizon fini avec temps d'arrêt	8
1.2.3	Évaluation d'espérance sur un horizon infini.	9
1.2.4	Le problème de Dirichlet.	11
1.3	Principe du Maximum discret	12
1.4	Étude analytique des matrices	14
1.4.1	Développement de la résolvante d'un opérateur autour d'une valeur propre	14
1.4.2	Application aux matrices stochastiques	15
1.4.3	Une autre caractérisation de P	15
1.5	Problèmes ergodiques	16
1.6	Propriétés combinatoires	17
1.7	Recherche de composantes fortement connexes	18
1.7.1	Un exemple numérique	19
1.8	Mesures de probabilité invariantes	20
1.9	Calcul de la matrice P sur la forme canonique	21
1.10	Interprétation probabiliste d'une base de $\mathcal{N}(A)$	23
1.11	États transients et états récurrents	25
1.12	Corrigé des exercices	26
2	Contrôle de chaînes de Markov	31
2.1	Problème de temps d'arrêt optimal	32
2.1.1	Application numérique	35
2.1.2	Un exemple :Le problème de la secrétaire	36
2.2	Programmation dynamique en horizon fini	38
2.2.1	Calcul de la fonction coût pour une stratégie donnée dans \mathcal{U}^{HR}	38
2.2.2	Équation de la programmation dynamique pour $\mathbb{U} \in \mathcal{U}^{HR}$	39
2.2.3	Liens avec les problèmes de temps d'arrêt	41
2.3	Programmation dynamique en horizon infini	44
2.3.1	Évaluation pour une stratégie stationnaire de la fonction coût	44
2.3.2	Équation d'optimalité	44
2.3.3	Politique optimale	46
2.3.4	Dépendance de la fonction valeur vis à vis de la fonction coût	47

2.3.5	Itération sur les valeurs	48
2.3.6	Itération sur les politiques ou algorithmes de Howard	51
2.3.7	Variantes sur Howard	53
2.4	Inéquations variationnelles	54
2.5	Corrigé des exercices	55
3	Processus de Diffusion	59
3.1	Introduction	59
3.2	Processus stochastiques	59
3.3	Equations différentielles stochastiques	60
3.3.1	Formule d'Ito	62
3.3.2	Générateur infinitésimal d'une diffusion	62
3.4	Equations de Kolmogorov	64
3.4.1	Problèmes à horizon fini	64
3.4.2	Problèmes à horizon infini	66
3.4.3	Conditions aux bords	66
3.4.4	Evaluation d'un cout moyen - le cas ergodique	68
3.5	Les Equations de Fokker-Planck	68
3.5.1	Exemple : Concentration de particules dans un décanteur	68
3.6	Processus de diffusion avec sauts	68
3.6.1	Equation intégro-différentielle	68
3.6.2	Exemple	68
3.7	Exercices	68
4	Etude numérique des équations de Kolmogorov	71
4.1	Introduction	71
4.2	Équations de Kolmogorov elliptiques	71
4.2.1	Réduction à un domaine borné	72
4.2.2	Méthode des différences finies	72
4.2.3	Conditions de stabilité de la méthode d'approximation	76
4.2.4	Discrétisation d'un opérateur elliptique sans dérivées croisées et uniformément elliptique.	79
4.2.5	Discrétisation d'un opérateur elliptique sans dérivées croisées et dégénéré.	85
4.2.6	Discrétisation d'un opérateur elliptique avec dérivées croisées.	86
4.2.7	Interprétation probabiliste d'un problème discret	89
4.3	Discrétisation du processus	90
4.4	Une section à déplacer dans le futur ?	91
4.5	Equations de Kolmogorov paraboliques	92
4.5.1	Méthode d'approximation	93
4.5.2	Calcul du rayon spectral de $(-\mathcal{B}_h)^{-1}\mathcal{C}_h$	96
4.6	Le cas HJB a déplacer dans un chapitre suivant	98
4.7	Résolution de systèmes linéaires	99
4.7.1	Méthodes directes	100
4.7.2	Méthodes itératives	100
4.8	Un exemple numérique	103
4.9	Corrections	104

5	Contrôle de processus de diffusion	107
5.1	Introduction	107
5.2	Contrôle de diffusions	108
5.3	Problème de temps d'arrêt optimal	109
5.4	Programmation Dynamique	110
5.5	Equations d'Hamilton-Jacobi-Bellman	110
5.5.1	Equation parabolique	111
5.5.2	Equation elliptique	111
5.5.3	Interprétation probabiliste	113
5.5.4	Inéquation variationnelle associée au problème de temps d'arrêt optimal	114
5.6	Solutions de viscosité	115
5.6.1	Introduction	115
5.6.2	Exemple d'applications	117
5.7	Application à l'optimisation dynamique de portefeuille	118
5.8	Contrôles singuliers et Inéquations variationnelles	124
5.8.1	Exemple de Gestion de portefeuille avec coûts de transaction propor- tionnels	124
5.8.2	Exemple des flux de dividendes	130
6	Etude numérique des équations d'Hamilton-Jacobi-Bellman	133
6.1	Equations d'Hamilton-Jacobi-Bellman elliptiques	133
6.1.1	Méthode d'approximation	135
6.1.2	Algorithmes de résolution de l'équation discrétisée	137
6.2	Equations d'Hamilton-Jacobi-Bellman paraboliques	141
6.2.1	Méthodes d'approximation	141
6.2.2	Exemples	143
6.3	Résolution numérique de problème de Contrôle ergodique	145
6.4	Résolution numérique de problème de contrôle de diffusion avec sauts	145
6.5	Méthodes de Kushner	145
6.6	Preuve de convergence par les solutions de viscosité	145
6.7	Remarques conclusives	145
6.8	Exercices	146

Introduction

Le but de ce cours est de fournir une introduction aux méthodes numériques utiles en finance. On s'intéresse en particulier à l'évaluation d'option et à l'optimisation dynamique de portefeuilles lorsque les prix des actifs sont modélisés par des processus stochastique markoviens.

Le prix d'une option européenne est dans certains cas solution d'une équation aux dérivées partielles parabolique (équation de Kolmogorov).

L'évaluation des option américaines peut se formuler comme un problème d'arrêt optimal : c'est le problème de contrôle le plus simple dans lequel la variable de décision est l'instant d'exercice de l'option.

Le problème type de gestion de portefeuille est le suivant : déterminer la stratégie, en avenir incertain, qui maximise un certain critère, par exemple, une fonction d'utilité de la richesse ou de la consommation, les variables de décision étant quand, combien et où investir.

La Méthode de la *Programmation Dynamique* conduit, dans le cas continu, à des équations aux dérivées partielles non linéaires du deuxième ordre pour la fonction valeur. Ce sont les équations d'Hamilton-Jacobi-Bellman (HJB) ou les inéquations variationnelles dont la solution permet de déterminer la fonction valeur et d'obtenir la stratégie optimale sous forme de feedback de l'état.

Le contrôle (c'est à dire la variable de décision) peut être régulier (comme la consommation), singulier ou impulsionnel (cas des coûts de transaction et des dividendes).

Souvent, la fonction valeur n'est pas de classe \mathcal{D}^2 et ne vérifie pas l'équation de la Programmation Dynamique au sens classique. Il est alors utile d'introduire le concept de *solutions de viscosité* qui permet de définir une notion de solution faible particulièrement bien adaptée à ces équations non linéaires souvent dégénérées.

Sauf pour quelques cas particuliers comme le problème d'évaluation d'option européenne dans le modèle de Black et Scholes ou le problème de gestion de portefeuille sans coûts de transaction, dit "problème de Merton", la résolution explicite des équations de Kolmogorov ou de la Programmation Dynamique n'est pas possible en général. Il faut alors avoir recours aux méthodes numériques.

Il y a deux grands types de méthodes :

- les méthodes probabilistes basées sur l'approximation des processus continus modélisant les actifs par des chaînes de Markov convenablement choisies.
- les méthodes d'analyse numérique qui consistent à résoudre numériquement les équations aux dérivées partielles associées au problème. Dans ce cas les conditions de stabilité et consistance numérique permettent d'obtenir une interprétation probabiliste de l'équation discrétisée et faire ainsi le lien avec la première méthode.

On traitera essentiellement dans ce cours les méthodes d'analyse numérique.

Chapitre 1

Chaînes de Markov

1.1 Chaînes de Markov

La notion de chaîne de Markov est apparue en 1907. Elle est due au mathématicien A. Markov puis a été développée en particulier par Kolmogorov et Feller. Le lecteur pourra se référer à [13], [15], [29] pour un point complet sur le sujet.

Il est suffisant, pour les approximations des problèmes en temps continu, de ne considérer que des chaînes à *états finis* et en *temps discret*. On considère un espace de probabilité $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$, c'est à dire un ensemble Ω équipé d'une σ -algèbre \mathcal{A} , d'une mesure de probabilité \mathbb{P} et d'une filtration $\mathcal{F} = (\mathcal{F}_n, n \in \mathbb{N})$. Rappelons qu'une filtration est une suite de sous σ -algèbres de \mathcal{A} telles que $\mathcal{F}_m \subset \mathcal{F}_n$ pour $m \leq n$.

Definition 1.1.1. Soit $(X_n)_{n \in \mathbb{N}}$ un processus stochastique sur un espace d'états fini $\mathcal{E} = \{1, 2, \dots, E\}$, défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Le processus $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov si pour tout $n \geq 0$,

$$\begin{aligned} \mathbb{P} \{X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0\} \\ = \mathbb{P} \{X_{n+1} = x_{n+1} \mid X_n = x_n\} \equiv M_{x_n x_{n+1}}^{(n)}. \end{aligned}$$

Cette propriété signifie que la probabilité pour que X_{n+1} se trouve en l'état x_{n+1} ne dépend que de X_n et non pas de ce qui s'est passé avant l'instant n (à une petite restriction prêt toutefois, il faut que l'événement $\{X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0\}$ soit de probabilité non nulle). C'est la propriété de Markov énoncée ici pour les processus à temps et espaces discrets.

Exercice 1.1.1. Montrer que si $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov et σ une fonction strictement croissante de \mathbb{N} dans \mathbb{N} alors :

$$\begin{aligned} \mathbb{P} \{X_{\sigma(n+1)} = x_{\sigma(n+1)} \mid X_{\sigma(n)} = x_{\sigma(n)}, X_{\sigma(n-1)} = x_{\sigma(n-1)}, \dots, X_{\sigma(1)} = x_{\sigma(1)}\} \\ = \mathbb{P} \{X_{\sigma(n+1)} = x_{\sigma(n+1)} \mid X_{\sigma(n)} = x_{\sigma(n)}\}. \end{aligned}$$

Exercice 1.1.2. En utilisant l'exercice précédent, soient k, n, p trois entiers positifs, montrez la propriété suivante :

$$\mathbb{E} [f(X_{n+k+p}) \mid X_n] = \mathbb{E} [\mathbb{E} [f(X_{n+k+p}) \mid X_{n+k}] \mid X_n]$$

Les coefficients $M_{x,y}^{(n)}$ sont appelés probabilités de transition de l'état x à l'état y à la date n . Ils vérifient :

$$\left\{ \begin{array}{l} M_{xy}^{(n)} \geq 0, \quad \forall n \in \mathbb{N}, \quad \forall x, y \in \mathcal{E}, \\ \sum_{y \in \mathcal{E}} M_{xy}^{(n)} = 1, \quad \forall n \in \mathbb{N}, \quad \forall x \in \mathcal{E}. \end{array} \right. \quad (1.1)$$

Les matrices $M^{(n)}$, $n \in \mathbb{N}$ (matrices $E \times E$) sont appelées *matrices de transition* de la chaîne de Markov. Une matrice vérifiant les propriétés (1.1) sera dite *matrice stochastique*.

Posons ici pour $p \geq n$, $P_{x,y}^{(n,p)} = \mathbb{P}\{X_p = y | X_n = x\}$, on obtient facilement, en utilisant la propriété de Markov, la relation matricielle suivante :

$$P^{(n,n+k+p)} = P^{(n,n+k)} P^{(n+k,n+k+p)} \quad (1.2)$$

avec $P^{(n,n)} = \mathbb{I}_{E \times E}$ et $P^{(n,n+1)} = M^{(n)}$.

Exercice 1.1.3. *Montrer la propriété (1.2).*

La fonction p^0 définie pour $x \in \mathcal{E}$ par $p^0(x) = \mathbb{P}\{X_0 = x\}$ est appelée *loi initiale*. Se donner une loi initiale sur un espace de dimension fini revient à se donner un vecteur $p^0 \in \mathbb{R}^E$. On représentera les lois par des vecteurs lignes et on notera indifféremment $p(x)$ où p_x selon que l'on privilégie l'interprétation fonctionnelle ou vectorielle. De la même façon p désignera une fonction ou un vecteur ligne de \mathbb{R}^E suivant le contexte.

La donnée d'une loi initiale et de la famille de probabilité de transition $M^{(n)}$ permet de calculer la probabilité de la trajectoire de la chaîne :

$$\mathbb{P}\{X_{n+1} = x_{n+1}, \dots, X_0 = x_0\} = p^0(x_0) \prod_{k=0}^n M_{x_k x_{k+1}}^{(k)}. \quad (1.3)$$

Ce qui caractérise complètement la chaîne.

Preuve : La preuve se fait par récurrence. Soit

$$B_n \stackrel{\text{def}}{=} \mathbb{P}\{X_n = x_n, \dots, X_0 = x_0\}$$

et supposons la propriété vérifiée pour B_n

$$\begin{aligned} \mathbb{P}\{B_{n+1}\} &= \mathbb{P}\{X_{n+1} = x_{n+1} | B_n\} \mathbb{P}\{B_n\} \\ &= \mathbb{P}\{X_{n+1} = x_{n+1} | X_n = x_n\} p^0(x_0) \prod_{k=0}^{n-1} M_{x_k x_{k+1}}^{(k)}. \end{aligned}$$

d'où le résultat. □

A partir de la donnée de la transition de la chaîne et de la loi initiale, on peut construire une version canonique de ce processus de Markov en prenant $(\Omega, \mathcal{F}) = (\mathcal{E}, \mathcal{P}(\mathcal{E}))^{\mathbb{N}}$, X_n la n^e coordonnée et la loi du n -uplet (X_1, \dots, X_n) donnée par (1.3). Il existe alors une probabilité \mathbb{P}_{p^0} sur $(\mathcal{E}, \mathcal{P}(\mathcal{E}))^{\mathbb{N}}$ telle que $(\Omega, \mathcal{F}, \mathbb{P}_{p^0}, (X_n)_n \geq 0)$ soit une chaîne de Markov de transition $M^{(n)}$ et de loi initiale p^0 (Voir [22] p.148). On utilisera souvent la version canonique et pour la donnée de la transition on construira une famille de chaîne de Markov pour la famille de lois initiales $(\delta_x)_{x \in \mathcal{E}}$.

Proposition 1.1.1. Équation de Fokker-Planck (ou Chapman-Kolmogorov) On considère une chaîne de Markov (X_n) de matrice de transition $M^{(n)}$ et de loi initiale p^0 . La loi marginale de X_n définie par $\mathbb{P}\{X_n = x\}$ et notée $p^n(x)$ est solution de l'équation récurrente :

$$\begin{cases} p^n = p^{n-1}M^{(n-1)}, \\ p^0 \text{ donnée.} \end{cases} \quad (1.4)$$

Preuve :

$$p^{n+1}(x) = \mathbb{P}\{X_{n+1} = x\} = \sum_{y \in \mathcal{E}} \mathbb{P}\{X_{n+1} = x | X_n = y\} \mathbb{P}\{X_n = y\} \quad (1.5)$$

d'où le résultat avec les notations matricielles. \square

Soit v une fonction de \mathcal{E} dans \mathbb{R} . Se donner une fonction de \mathcal{E} dans \mathbb{R} revient à se donner un vecteur $v = \begin{pmatrix} v_1 \\ \vdots \\ v_E \end{pmatrix} \in \mathbb{R}^E$ tel que $v_i \equiv v(i)$. On utilisera les mêmes conventions d'écriture que pour les lois, v désignera une fonction ou un vecteur de \mathbb{R}^E suivant le contexte. Remarquer que pour les fonctions on utilise des vecteurs colonne et pour les lois des vecteurs ligne.

Soit p un entier ($p \geq 1$), on obtient facilement que :

$$\begin{aligned} \mathbb{E}[v(X_{n+p}) | X_n = x_n, \dots, X_0 = x_0] &= \sum_{x \in \mathcal{E}} v_x \mathbb{P}\{X_{n+p} = x | X_n = x_n, \dots, X_0 = x_0\} \\ &= \sum_{x \in \mathcal{E}} v_x P_{x_n, x}^{(n, n+p)} = (P^{(n, n+p)}v)_{x_n} \end{aligned}$$

On notera que la démonstration précédente a aussi montré que :

$$\mathbb{E}[v(X_{n+p}) | X_n = x_n, \dots, X_0 = x_0] = \mathbb{E}[v(X_{n+p}) | X_n = x_n]$$

Ainsi que le résultat suivant pour $p = 1$

$$\mathbb{E}[v(X_{n+1}) | X_n = x_n] = (M^{(n)}v)_{x_n}$$

Si $M^{(n)} = M, \forall n \in \mathbb{N}$, on dit que la chaîne est *homogène* en temps. Les matrices $P^{(n, n+p)}$ ne dépendent alors que de p et on omettra n de la notation. L'équation (1.2) s'écrit alors comme une récurrence matricielle :

$$P^{(n+1)} = P^{(n)}M \quad \text{et} \quad P^{(0)} = \mathbb{I}_{E \times E} \quad (1.6)$$

On peut aussi remarquer la propriété suivante. Soit $(R_v^n)_{n \in \mathbb{N}}$ le processus défini par :

$$R_v^n \equiv v(X_n) - \sum_{k=0}^{n-1} ((M^{(k)} - I)v)(X_k) \quad (1.7)$$

Il vérifie $R_v^{n+1} - R_v^n = v(X_{n+1}) - (M^{(n)}v)(X_n)$ et le calcul précédent montre que R_v^n est une martingale discrète partant de $R_v^0 = v(X_0)$ pour la filtration $\mathcal{F}_n \stackrel{\text{def}}{=} \sigma(X_k, k \leq n)$. ($\mathbb{E}[R_v^m | \mathcal{F}_n] = R_v^n$ pour tout couple (m, n) vérifiant $m \geq n$).

Exemple 1.1.1. Soit X_0 une variable aléatoire discrète à valeurs dans $\mathcal{E} = \{1, \dots, E\}$, et soient $(Y_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires indépendantes et indépendantes de X_0 et de même loi de densité f sur $[0, 1]$. On se donne aussi une fonction $b : \mathcal{E} \times [0, 1] \mapsto \mathcal{E}$ et on définit par récurrence $X_{n+1} = b(X_n, Y_n)$. Il est alors facile et laissé en exercice au lecteur de montrer que la suite $(X_n)_{n \in \mathbb{N}}$ définit une chaîne de Markov homogène d'espace d'état \mathcal{E} et de matrice de transition M donnée par :

$$M_{i,j} = \mathbb{E} [\mathbb{I}_{b(i, Y_0)=j}] = \int_0^1 \mathbb{I}_{b(i,u)=j} f(u) du.$$

Le résultat est bien sur conservé si l'on remplace dans ce qui précède la loi de densité f par une loi discrète. Beaucoup d'exemples de chaînes de Markov sont en fait construits de cette façon.

Exercice 1.1.4. Inversement, pour une matrice stochastique M donnée pouvez vous construire une chaîne de Markov de matrice de transition M sous la forme précédente $X_{n+1} = b(X_n, Y_n)$.

Exemple 1.1.2. La marche aléatoire de Cox Ross qui est très utilisée en finance est un cas particulier de ce qui précède. Elle est définie ainsi : soient a et b deux nombres réels et soit $(U_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes de loi commune $\mathbb{P}\{U_n = 1 + a\} = p$, $\mathbb{P}\{U_n = 1 + b\} = 1 - p$ avec $0 < p < 1$. On pose $S_0 = s_0$ (cours initial) et $S_{n+1} = S_n U_{n+1}$. On a donc $S_n = s_0 \prod_{i=1}^n U_i$ (cours à l'instant n). Supposons que l'on s'intéresse au temps $n \in [0, N]$ l'espace d'état de la chaîne S_n est alors fini, $\mathcal{E} = \{\gamma_{i,j} \equiv s_0(1+a)^i(1+b)^j, \quad 0 \leq i+j \leq N\}$. On vérifie que S_n est une chaîne de Markov homogène de matrice de transition :

$$M_{\gamma_{i,j}, \gamma_{i+1,j}} = p, \quad M_{\gamma_{i,j}, \gamma_{i,j+1}} = 1 - p, \quad M_{\gamma_{i,j}, \gamma_{k,l}} = 0 \quad \text{sinon.} \quad (1.8)$$

Ce qui précède pourrait s'étendre dans le cas d'un horizon infini, l'espace d'état \mathcal{E} devient alors un ensemble dénombrable et la matrice de transition est une matrice infinie.

1.2 Calculs d'espérance et Équations de Kolmogorov

1.2.1 Évaluation d'espérance sur un horizon fini.

On veut calculer l'espérance mathématique d'une fonctionnelle additive (ou multiplicative) de la trajectoire sur un horizon fini c.a.d. sur un nombre fini de périodes de temps.

Proposition 1.2.1. Soient \diamond un opérateur qui peut être soit l'addition soit la multiplication ¹ et $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice de transition $M^{(n)}$ définie sur un espace \mathcal{E} , la fonctionnelle :

$$v^n(x) = \mathbb{E} [c^n(X_n) \diamond c^{n+1}(X_{n+1}) \diamond \dots \diamond c^{N-1}(X_{N-1}) \diamond f(N, X_N) | X_n = x],$$

où f est une fonction de $[0, N] \times \mathcal{E}$ dans \mathbb{R} et les c^i des fonctions de \mathcal{E} dans \mathbb{R} , est solution de l'équation récurrente de Kolmogorov rétrograde :

$$\begin{cases} v^n &= c^n \diamond M^{(n)} v^{n+1}, \quad \forall n = 0, \dots, N-1, \\ v^N &= f(N, \cdot). \end{cases} \quad (1.9)$$

¹Noter qu'il s'agit d'une multiplication termes à termes dans le cas de deux vecteurs $(v \diamond w)_i = v_i \diamond w_i$

Preuve : Pour $n = N$, la définition de v donne immédiatement $v^N(x) = f(N, x)$, pour $n < N$ en utilisant les propriétés de l'espérance conditionnelle et la propriété de Markov, on peut écrire :

$$\begin{aligned} v^n(x) &= c^n(x) \diamond \mathbb{E} [\mathbb{E} [c^{n+1}(X_{n+1}) \diamond \cdots \diamond + f(X_N) | X_{n+1}] | X_n = x] \\ &= c^n(x) \diamond \mathbb{E} [v^{n+1}(X_{n+1}) | X_n = x] = c^n(x) \diamond M^{(n)} v^{n+1}(x) \end{aligned}$$

pour $n < N$ et pour l'opérateur $\diamond = +$, une deuxième preuve qui utilise les probabilités conditionnelles à travers la relation (1.2) et la définition de v est la suivante :

$$\begin{aligned} v^n &= \sum_{j=n}^{N-1} P^{(n,j)} c^j + P^{(n,N)} f = c^n + \sum_{j=n+1}^{N-1} P^{(n,j)} c^j + P^{(n,N)} f \\ &= c^n + \sum_{j=n+1}^{N-1} P^{(n,n+1)} P^{(n+1,j)} c^j + P^{(n,n+1)} P^{(n+1,N)} f = c^n + M^{(n)} v^{n+1} \end{aligned}$$

□

Prix d'option européenne.

Dans le cas de modèles discrets, le prix à l'instant n d'une option européenne promettant $f(X_N)$ à l'instant N , s'exprime dans certains cas sous la forme :

$$v_n = \mathbb{E} \left[\frac{1}{(1+r)^{N-n}} f(X_N) | \mathcal{F}_n \right]. \quad (1.10)$$

où $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov, r le taux d'intérêt sans risque sur une période et $\mathcal{F}_n = \sigma\{X_p, p \leq n\}$ [?].

En appliquant la proposition précédente avec $c^j(x) = 1/(1+r)$ et pour l'opérateur $\diamond = \star$, on déduit le calcul du prix d'une option européenne (1.10). La fonctionnelle :

$$v_n(x) = \mathbb{E} \left[\frac{1}{(1+r)^{N-n}} f(X_N) | X_n = x \right]$$

satisfait l'équation :

$$\begin{cases} v_n &= \frac{1}{1+r} M^{(n)} v_{n+1} \quad \forall n = 0, \dots, N-1, \\ v_N &= f. \end{cases} \quad (1.11)$$

En particulierisant le résultat précédent pour la chaîne du modèle de Cox-Ross (exemple 1.1.2), le prix à l'instant n qui satisfait l'équation (1.10) est donnée par la solution de l'équation :

$$\begin{cases} v_{n-1}(y) &= \frac{p}{1+r} v_n(y(1+a)) + \frac{1-p}{1+r} v_n(y(1+b)), \quad \forall n = 1, \dots, N, \quad \forall y \in \mathcal{E} \\ v_N(y) &= f(y). \end{cases} \quad (1.12)$$

Un exemple de calcul numérique :

```
-->function y=f(x); y=max(0,x-K) ; endfunction ← fonction f
-->pa=0.5; pb=1-pa; X1= 1; a = -0.05 ; b = 0.05; K=1; r=0.01; N=10
Les points accessibles en i étapes à partir de X1 stockés dans X(i,1:i)
```

```

X = zeros(N,N) ;
X(:,1) = X1*(1+a).^ (0:N-1)';           ← première colonne X(1,i)= X_1(1+a)^(i-1)
for i=2:N                                 ← calcul récursif de chaque lignes de X
    X(i,2:i) = X(i-1,1:i-1)*(1+b);
end
calcul de v_n(x) aux points donnés par X: V(n,i) ≡ v_n(X(n,i))
V = zeros(N,N) ; V(N,:) = f(X(N,:));    ← initialisation

for i=N-1:-1:1
    V(i,1:i) = V(i+1,1:i)*pa/(1+r);
    V(i,1:i) = V(i,1:i) + (pb/(1+r))* V(i+1,2:i+1); ← utilisation de la formule (1.12)
end

for i=N-3:N
    plot2d(X(i,1:i),V(i,1:i),N-i+1,"111","",[0.8,0.0,1.2,0.2]) ← graphique
end

```

Exercice 1.2.1. *Montrer que la fonctionnelle :*

$$v_n(x) = \mathbb{E} \left[\sum_{i=n}^{N-1} \left(\prod_{k=n}^{i-1} \xi^k(X_k) \right) c^i(X_i) + \left(\prod_{k=n}^{N-1} \xi^k(X_k) \right) f(N, X_N) \middle| X_n = x \right]$$

satisfait l'équation de Kolmogorov suivante :

$$\begin{cases} v^n(x) &= c^n(x) + \xi^n(x)M^{(n)}v^{n+1}(x), \forall n = 0, \dots, N-1, \\ v^N(x) &= f(N, x). \end{cases}$$

Exercice 1.2.2. *Soit X_n une chaîne de Markov et L_n un processus aléatoire donné par $L_n = \sum_{i=0}^n X_i$. On cherche ici à calculer :*

$$v_n(x) = \mathbb{E}[\phi(L_N) | X_n = x]$$

Montrer que ce problème se ramène à un calcul d'espérance en horizon fini pour une chaîne de Markov à préciser et écrire l'équation récurrente vérifiée par $v_n(x)$. Même exercice pour $L_n = \max_{i=0}^n X_i$ et $L_n = \prod_{i=0}^n X_i$.

1.2.2 Évaluation d'espérance sur un horizon fini avec temps d'arrêt

On considère à présent le cas où l'on peut s'arrêter avant la date N selon une procédure d'arrêt aléatoire.

Definition 1.2.1. *Une variable aléatoire $\nu : \Omega \rightarrow \mathbb{R}^+$ est un temps d'arrêt pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ si $\{\nu = k\}$ est \mathcal{F}_k -mesurable $\forall k \in \mathbb{N}$.*

Exemple : Étant donné un ensemble B inclus dans \mathcal{E} , $\nu_B = \inf\{n \geq 0 : X_n \notin B\}$ est un temps d'arrêt pour la filtration $\mathcal{F}_n = \sigma\{X_p, p \leq n\}$ (le montrer est laissé en exercice) appelé *temps de sortie* de l'ensemble B .

Proposition 1.2.2. *Étant donnée une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ de matrice de transition $M^{(n)}$ et démarrant du point x à la date n , la fonctionnelle :*

$$v^n(x) = \mathbb{E}_{n,x} \left[\sum_{j=n}^{N \wedge \nu_B^n - 1} c^j(X_j) \mathbb{I}_{N \wedge \nu_B^n > n} + f(X_N) \mathbb{I}_{N < \nu_B^n} + g(X_{\nu_B^n}) \mathbb{I}_{N \geq \nu_B^n} \right] \quad (1.13)$$

est solution de l'équation récurrente de Kolmogorov rétrograde :

$$\begin{cases} v^n(x) &= c^n(x) + M^{(n)} v^{n+1}(x), \forall n = 0, \dots, N-1, & \text{sur } B, \\ v^N(x) &= f(x) & \text{sur } B, \\ v^n(x) &= g(x) \forall n = 0, \dots, N & \text{sur } C_B. \end{cases} \quad (1.14)$$

où $\nu_B^n = \inf\{k \geq n : X_k \notin B\}$ est le temps de sortie de l'ensemble B pour la chaîne démarrant en x à la date n et la notation $\mathbb{E}_{n,x}[\cdot] = \mathbb{E}[\cdot | X_n = x]$.

Preuve : Notons tout d'abord que de façon évidente, il existe une unique solution $(v_k(x))_{k \in [0, N]}$ à l'équation (1.14). Nous allons montrer que cette solution vérifie l'équation (1.13). En définissant un prolongement à tout \mathcal{E} de la fonction $c^k(x)$ par l'expression suivante $\bar{c}^k(x) = v^k(x) - M^{(k)} v^{k+1}(x)$, on vérifie facilement que $\mathcal{M}_k = v^k(X_k) + \sum_{i=n}^{k-1} \bar{c}^i(X_i) \mathbb{I}_{\{k > n\}}$ définit bien une martingale pour $k \in [n, N]$. $\mathcal{M}_{k \wedge \nu_B^n}$ est alors aussi une martingale. On le vérifie en utilisant le fait que $\mathbb{I}_{\{\nu_B^n \geq k\}} = 1 - \mathbb{I}_{\{\nu_B^n \leq k-1\}}$ est \mathcal{F}_{k-1} mesurable et qu'en conséquence :

$$\mathbb{E}[\mathcal{M}_{k \wedge \nu_B^n} | \mathcal{F}_{k-1}] = \sum_{i=n}^{k-1} \mathbb{I}_{\{\nu_B^n = i\}} \mathcal{M}_i + \mathbb{I}_{\{\nu_B^n \geq k\}} \mathbb{E}[\mathcal{M}_k | \mathcal{F}_{k-1}] = \mathcal{M}_{k-1 \wedge \nu_B^n}.$$

On a donc $v^n X_n = \mathcal{M}_n = \mathcal{M}_{n \wedge \nu_B^n} = \mathbb{E}[\mathcal{M}_{N \wedge \nu_B^n} | \mathcal{F}_n]$, soit aussi :

$$v^n(x) = \mathbb{E} \left[v^{N \wedge \nu_B^n}(X_{N \wedge \nu_B^n}) + \sum_{i=n}^{N \wedge \nu_B^n - 1} \bar{c}^i(X_i) \mathbb{I}_{N \wedge \nu_B^n > n} | X_n = x \right]$$

Pour conclure on remarque que d'une part en utilisant (1.14) on obtient :

$$v^{N \wedge \nu_B^n}(X_{N \wedge \nu_B^n}) = \mathbb{I}_{\nu_B^n < N} g(X_{\nu_B^n}) + \mathbb{I}_{\nu_B^n \geq N} f(X_N)$$

et d'autre part sur l'ensemble $\nu_B^n > n$ on a $\bar{c}^k(X_k) = c^k(X_k)$ pour k dans l'ensemble aléatoire $k \in [n, \nu_B^n[$. \square

Exercice 1.2.3. (option barrière) Soit X_n une chaîne de Markov d'espace d'état fini inclus dans \mathbb{R} . Écrire l'équation récurrente vérifiée par $v_n(x)$:

$$v_n(x) \equiv \mathbb{E} \left[\frac{1}{(1+r)^{N-n}} \phi(X_N) \prod_{i=n}^N \mathbb{I}_{X_i \leq K} | X_n = x \right]$$

1.2.3 Évaluation d'espérance sur un horizon infini.

On considère le même problème que précédemment mais sur un horizon infini et un coût par période actualisé c'est à dire décroissant de façon exponentielle avec le temps. On suppose

ici également que la chaîne est homogène. Nous commençons par rappeler la propriété de Markov forte.

Sur l'espace canonique $\mathcal{E}^{\mathbb{N}}$ on peut définir l'opérateur de translation $\theta(\omega) \equiv (X_{n+1}(\omega))_{n \geq 0}$ on notera alors θ_p son $p^{\text{ième}}$ itéré et pour une v.a T à valeur dans \mathbb{N} , θ_T est défini par $X_n \circ \theta_T = X_{n+T(\omega)}(\omega)$. Le résultat suivant est la propriété de Markov forte.

Théorème 1.2.1. *Soit X_n une chaîne de Markov homogène, T un \mathcal{F} -temps d'arrêt et Y une v.a mesurable bornée sur (Ω, \mathcal{F}) , pour toute loi initiale p^0 on a :*

$$\mathbb{E}_{p^0} [Y \circ \theta_T \mathbb{I}_{\{T < \infty\}} | \mathcal{F}_T] = \mathbb{I}_{\{T < \infty\}} \mathbb{E}_{X_T} [Y] \quad (1.15)$$

démonstration [5] page 117.

Proposition 1.2.3. *Étant donnée une chaîne de Markov homogène $(X_n)_{n \in \mathbb{N}}$ de matrice de transition M , la fonctionnelle :*

$$v(x) = \mathbb{E} \left[\sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} c(X_n) | X_0 = x \right]$$

avec $\lambda > 0$ est l'unique solution de l'équation de Kolmogorov :

$$(A - \lambda I)v + c = 0, \quad (1.16)$$

où $A = M - I$ est appelé le générateur de la chaîne de Markov.

Preuve :

$$\begin{aligned} v(x) &= \frac{1}{1+\lambda} c(x) + \frac{1}{1+\lambda} \mathbb{E} \left[\sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} c(X_{n+1}) | X_0 = x \right] \\ &= \frac{1}{1+\lambda} c(x) + \frac{1}{1+\lambda} \mathbb{E} [v(X_1) | X_0 = x] \\ &= \frac{1}{1+\lambda} c(x) + \frac{1}{1+\lambda} (Mv)(x) \end{aligned}$$

$v(x)$ est bien solution de l'équation (1.16) et nous verrons plus loin (Proposition 1.3.2) que pour toute matrice stochastique la matrice $M - I - \lambda I$ est inversible pour $\lambda > 0$ et donc que l'équation (1.16) admet une unique solution. \square

Remarque : On peut noter que la clef de la démonstration est la propriété de Markov forte (1.15). En effet si nous détaillons un peu, soit $Y(\omega) = \sum_{n=0}^{\infty} c(X_n)/(1+\lambda)^{n+1}$. C'est une variable aléatoire bornée \mathcal{F} mesurable et elle vérifie $Y = (c(X_0) + Y \circ \theta)/(1+\lambda)$. On calcule

$$v(x) = \mathbb{E} [Y | X_0 = x] = c(x)/(1+\lambda) + \mathbb{E} [Y \circ \theta | X_0 = x] / (1+\lambda)$$

et en utilisant la propriété de Markov forte (avec $T \equiv 1$) on obtient :

$$\mathbb{E} [Y \circ \theta_1 | X_0 = x] = \mathbb{E} [\mathbb{E} [Y \circ \theta_1 | \mathcal{F}_1] | X_0 = x] \quad (1.17)$$

$$= \mathbb{E} [\mathbb{E}_{X_1} [Y] | X_0 = x] = \mathbb{E} [v(X_1) | X_0 = x] \quad (1.18)$$

\square

Remarque : On peut montrer le même résultat par une démonstration matricielle en remarquant que $v(x) = \sum_{y \in \mathcal{E}} \sum_{n=0}^{\infty} 1/(1+\lambda)^{n+1} M_{x,y}^{(n)} c_y$. Les séries précédentes convergent pour $\lambda > 0$ car $M^{(n)}$ est une matrice stochastique (produit de matrices stochastiques) et que l'on a alors $|M^{(n)} c|_{\infty} \leq |c|_{\infty}$. On peut donc intervertir valeur moyenne et série. On notera aussi que $(A - \lambda I)$ est inversible pour $\lambda > 0$ (proposition 1.3.2 et 1.3.3) et que cette deuxième approche nous montre que $(\lambda I - A)^{-1} = \sum_{n=0}^{\infty} 1/(1+\lambda)^{n+1} M^{(n)}$. \square

Remarque : Pour $\lambda = 0$ la matrice $(A = M - I)$ n'est pas inversible (le vecteur $\mathbb{1}_E$ est dans le noyau de A), le problème sera de moins en moins bien conditionné quand λ se rapproche de zéro. Nous montrerons dans des paragraphes ultérieurs que λv^λ , ou dans certains cas v^λ , peuvent avoir une limite quand λ décroît vers zéro et nous caractériserons ces limites. \square

1.2.4 Le problème de Dirichlet.

On considère à présent le cas où l'on s'arrête avant l'infini selon une procédure d'arrêt.

Proposition 1.2.4. *Étant donnée la chaîne de Markov homogène (X_n) de générateur A , B un ensemble inclus dans \mathcal{E} , ν_B le temps de sortie de B , la fonctionnelle :*

$$v(x) = \mathbb{E} \left[\sum_{n=0}^{\nu_B-1} \frac{1}{(1+\lambda)^{n+1}} c(X_n) + \frac{1}{(1+\lambda)^{\nu_B+1}} \Phi(X_{\nu_B}) \mid X_0 = x \right]$$

est solution du problème de Dirichlet

$$\begin{cases} (A - \lambda I)v + c = 0 & \text{sur } B, \\ v = \frac{\Phi}{1+\lambda} & \text{sur } C_B. \end{cases} \quad (1.19)$$

Preuve : Soit $Y(\omega)$ la variable aléatoire définie par

$$Y(\omega) = \mathbb{1}_{\{\nu_B > 0\}} \left\{ \sum_{n=0}^{\nu_B-1} \frac{1}{(1+\lambda)^{n+1}} c(X_n) \right\} + \frac{1}{(1+\lambda)^{\nu_B+1}} \Phi(X_{\nu_B}). \quad (1.20)$$

La fonctionnelle $v(x)$ vaut par définition $\mathbb{E}[Y \mid X_0 = x]$. En utilisant l'opérateur de translation θ_p (voir section 1.2.3) on obtient

$$\nu_B \circ \theta_p = \inf\{n \geq 0 : X_{n+p} \notin B\} = \inf\{n \geq p : X_n \notin B\} - p.$$

On obtient donc, sur $\{\nu_B \geq p\}$, la propriété suivante, $\nu_B = p + \nu_B \circ \theta_p$.

Regardons tout d'abord $Y(\omega)$ sur $\{\nu_B \geq 1\}$, on a alors

$$Y = \frac{1}{(1+\lambda)} c(X_0) + \mathbb{1}_{\{\nu_B > 1\}} \left\{ \sum_{n=1}^{\nu_B-1} \frac{1}{(1+\lambda)^{n+1}} c(X_n) \right\} + \frac{1}{(1+\lambda)^{\nu_B+1}} \Phi(X_{\nu_B})$$

et, en utilisant $\nu_B = 1 + \nu_B \circ \theta$ sur $\{\nu_B \geq 1\}$, on obtient :

$$Y = \frac{1}{(1+\lambda)} c(X_0) + \frac{1}{(1+\lambda)} Y \circ \theta \quad (1.21)$$

d'où finalement :

$$Y = \mathbb{I}_{\{\nu_B=0\}} \frac{1}{(1+\lambda)} \Phi(X_0) + \mathbb{I}_{\{\nu_B \geq 1\}} \left\{ \frac{1}{(1+\lambda)} c(X_0) + \frac{1}{(1+\lambda)} Y \circ \theta \right\}$$

l'équation (1.19) se déduit alors facilement de l'équation précédente. Le terme $\mathbb{E} [\mathbb{I}_{\{\nu_B \geq 1\}} Y \circ \theta | X_0 = x]$ se calcule en utilisant la propriété de Markov forte, pour la v.a bornée $Y(\omega)$, comme suit :

$$\begin{aligned} \mathbb{E} [\mathbb{I}_{\{\nu_B \geq 1\}} Y \circ \theta | X_0 = x] &= \mathbb{E} [\mathbb{I}_{\{\nu_B \circ \theta \geq 0\}} Y \circ \theta | X_0 = x] \\ &= \mathbb{E} [\mathbb{E} [\mathbb{I}_{\{\nu_B \circ \theta \geq 0\}} Y \circ \theta | X_1] | X_0 = x] \\ &= \mathbb{E} [\mathbb{E}_{X_1} [\mathbb{I}_{\{\nu_B \geq 0\}} Y] | X_0 = x] = \mathbb{E} [\mathbb{E}_{X_1} [Y] | X_0 = x] \\ &= Mv(x) \end{aligned}$$

Pour conclure il suffit de remarquer que si $x \in C_B$ alors $\nu_B = 0$ et si $x \in B$ on a $\nu_B \geq 1$. \square

Remarque : Le problème de Dirichlet (1.19) admet une unique solution. En effet v est connu sur C_B et la sous matrice $(A - \lambda)|_B$ est une matrice inversible car elle reste à diagonale fortement dominante (voir définition 1.3.2) de paramètre λ . \square

Exercice 1.2.4. Montrer que la fonctionnelle :

$$v(x) = \mathbb{E} \left[\sum_{i=0}^{\nu_B-1} \left(\prod_{k=0}^{i-1} \xi(X_k) \right) c^i(X_i) + \left(\prod_{k=0}^{\nu_B-1} \xi(X_k) \right) \Phi(X_{\nu_B}) | X_0 = x \right]$$

satisfait l'équation de Kolmogorov suivante :

$$\begin{cases} v(x) = c(x) + \xi(x)Mv(x) & \text{sur } B \\ v(x) = \Phi(x) & \text{pour } x \in C_B \end{cases}$$

1.3 Principe du Maximum discret

On note $|v|_\infty = \max_{x \in \mathcal{E}} |v_x|$ pour $v \in \mathbb{R}^E$ et $|A|_\infty = \max_{|v|_\infty \leq 1} |Av|_\infty$ pour $A \in \mathbb{R}^{E \times E}$.

Proposition 1.3.1. Toute matrice stochastique M vérifie $|M|_\infty = 1$ et donc tous les modules des valeurs propres de M sont inférieurs à 1.

Preuve :

$$|Mv|_\infty = \max_{x \in \mathcal{E}} \left| \sum_{y \in \mathcal{E}} M_{xy} v_y \right| \leq (\max_{y \in \mathcal{E}} |v_y|) (\max_{x \in \mathcal{E}} \sum_y M_{xy}) = \max_{y \in \mathcal{E}} |v_y|.$$

et donc $|M|_\infty \leq 1$. En prenant $v' = (1 \ 1 \ \dots \ 1)$ on obtient l'égalité. \square

On peut retrouver ce résultat en considérant une chaîne de Markov homogène de matrice de transition M . On a alors

$$\|(Mc)_x\| = \|\mathbb{E} [c(X_1) | X_0 = x]\| \leq |c|_\infty \mathbb{E} [1 | X_0 = x] \leq |c|_\infty.$$

Definition 1.3.1. On dit qu'un opérateur A satisfait au :

– Principe du Maximum Discret si et seulement si :

$$(Av)_x \leq 0, \forall x \in \operatorname{Argmax}_{x \in \mathcal{E}} v_x$$

– Principe du Maximum Discret Positif [resp. strictement positif] si et seulement si :

$$\max_{x \in \mathcal{E}} v_x \geq 0 \Rightarrow (Av)_x \leq 0, \forall x \in \operatorname{Argmax}_{x \in \mathcal{E}} v_x$$

$$[\text{resp. } \max_{x \in \mathcal{E}} v_x > 0 \Rightarrow (Av)_x < 0, \forall x \in \operatorname{Argmax}_{x \in \mathcal{E}} v_x].$$

Remarque : On notera que la définition donnée ici n'est pas conforme (il faut changer A en $-A$) à celle que l'on trouve dans les ouvrages d'analyse numérique (par exemple [28] p. 31) \square

Lemme 1.3.1. Si A satisfait le Principe du Maximum strictement positif, alors

$$((Av)_x \geq 0 \forall x \in \mathcal{E}) \Rightarrow v_x \leq 0 \forall x \in \mathcal{E}. \quad (1.22)$$

Proposition 1.3.2. Si M est une matrice stochastique, alors le générateur $A = M - I$ satisfait le Principe du Maximum Discret et $A - D$, D diagonale, $D_{ii} \geq 0$ [resp. > 0], satisfait le Principe du Maximum Discret [resp. strictement] positif.

Preuve : Soit $x \in \operatorname{Argmax}_{x \in \mathcal{E}} \{v_x\}$, on a :

$$(Av)_x = \sum_{y \in \mathcal{E}} M_{xy} v_y - v_x \leq \sum_{y \in \mathcal{E}} M_{xy} v_x - v_x \leq 0,$$

puisque $M_{xy} \geq 0$, $\sum_{y \in \mathcal{E}} M_{xy} = 1$ et $v_y \leq v_x \forall y \in \mathcal{E}$. Les autres cas sont laissés en exercice. \square

Proposition 1.3.3. Toute matrice vérifiant le Principe du Maximum strictement positif est inversible et les éléments de l'inverse sont négatifs.

Exercice 1.3.1. Démontrer la proposition précédente.

Remarque : Soit $\mu = 1/(1 + \lambda)$ avec $\lambda > 0$, on a alors la relation suivante $(I - \mu M) = (-1/(1 + \lambda))(A - \lambda I)$. $(A - \lambda I)$ vérifie le principe du maximum strictement positif et le lemme (1.3.1) se traduit par une propriété de positivité de l'opérateur $(I - \mu M)^{-1}$ ($(I - \mu M)$ est inversible d'après la proposition précédente 1.3.3 et ses éléments sont positifs) :

$$\text{si } u \geq w \text{ alors } (I - \mu M)^{-1}u \geq (I - \mu M)^{-1}w$$

En fait comme on l'a déjà noté

$$(I - \mu M)^{-1}u = \sum_{i=0}^{\infty} \mu^i M^i u$$

et de la positivité de M on déduit une autre propriété si $u \geq 0$ alors $(I - \mu M)^{-1}u \geq u$. \square

Definition 1.3.2. Soit A une matrice $E \times E$ telle que

$$A_{xy} \geq 0, \quad \forall x \neq y, x, y \in \mathcal{E}$$

$\exists \lambda > 0$ tel que $\sum_{y \in \mathcal{E}} A_{xy} \leq -\lambda < 0 \quad \forall x \in \mathcal{E}$. La matrice A est dite à diagonale fortement dominante.

Proposition 1.3.4. Soit A une matrice à diagonale fortement dominante de paramètre λ , alors A satisfait le Principe du maximum discret strictement positif. A est donc inversible (Proposition 1.3.3) et on a la majoration $\|A^{-1}\|_{\infty} \leq 1/\lambda$

Preuve : Soit $x^* \in \text{Argmax}_{x \in \mathcal{E}} \{v_x\}$. Supposons $\max_{x \in \mathcal{E}} v_x > 0$. On a tout d'abord :

$$(Av)_{x^*} = \sum_{y \in \mathcal{E}} A_{x^*y} v_y = A_{x^*x^*} v_{x^*} + \sum_{y \neq x^*} A_{x^*y} v_y \leq \sum_{y \in \mathcal{E}} A_{x^*y} v_{x^*} \leq -\lambda v_{x^*} < 0.$$

En notant par e le vecteur dont toutes les composantes valent 1, la propriété de diagonale dominante s'écrit aussi $Ae \leq -\lambda e$. En utilisant la Proposition 1.3.3 on a $(-A^{-1}) \geq 0$ et donc $(-A^{-1})Ae \leq -\lambda(-A^{-1})e$. Soit en fait $(-A^{-1})e \leq (1/\lambda)e$ ce qui nous donne le résultat $\|A^{-1}\|_{\infty} \leq 1/\lambda$ (sachant que pour une matrice positive B $\|B\|_{\infty} = \|Be\|_{\infty}$) \square

1.4 Étude analytique des matrices

Nous donnons ici quelques résultats de structure sur les matrices et en particulier sur les matrices stochastiques. Ces résultats constitueront les outils nécessaires pour comprendre le comportement asymptotique des chaînes de Markov et traiter les problèmes ergodiques. La présentation faite ici suit [13] ou le lecteur pourra trouver les preuves de certains théorèmes cités ici sans démonstration. On pourra aussi se reporter au livre de Kato [14] sur la perturbation des opérateurs.

1.4.1 Développement de la résolvante d'un opérateur autour d'une valeur propre

Soit un opérateur A ici une matrice $E \times E$.

Definition 1.4.1. On note R^{λ} la résolvante de l'opérateur A , définie par $R^{\lambda}(A) = (A - \lambda I)^{-1}$ où $\forall \lambda \in \mathbb{C} - \mathcal{S}_A$ avec \mathcal{S}_A le spectre de A — c.a.d. l'ensemble des $\lambda \in \mathbb{C}$ où $A - \lambda I$ est singulière.

Théorème 1.4.1. Dans un voisinage de la valeur propre λ_0 de A , on a le développement

$$R^{\lambda} = -(\lambda - \lambda_0)^{-1}P - \sum_{n=1}^{m-1} (\lambda - \lambda_0)^{-n-1} N^n + \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n S^{n+1}, \quad (1.23)$$

avec :

- P est le projecteur spectral sur l'espace propre V_{λ_0} associé à la valeur propre λ_0 , de dimension m ; ($P(A - \lambda_0 I) = (A - \lambda_0 I)P = N$).
- N est le nilpotent associé à la valeur propre λ_0 ; ($N^m = 0$).
- S vérifie $S(A - \lambda_0 I) = (A - \lambda_0 I)S = I - P$. C'est donc une pseudo inverse de A sur le supplémentaire de l'espace propre V_{λ_0} .

Preuve : On se reportera à [13, 14] □

Remarque : Les matrices P , N et S vérifient aussi les propriétés suivantes que nous utiliserons par la suite $PN = NP = N$, $PS = SP = 0$ et $NS = SN = 0$. □

1.4.2 Application aux matrices stochastiques

Nous particularisons le théorème précédent pour les matrices stochastiques :

Corollaire 1.4.1. *Les nilpotents associés aux valeurs propres de module 1 d'une matrice stochastique sont nuls.*

Preuve : Soient, λ_0 une valeur propre de module 1 de M , matrice stochastique, et $\lambda = \lambda_0 \alpha \in \mathbb{C}$, $\alpha > 1 \in \mathbb{R}$. Pour une fonction c arbitraire donnée, la fonctionnelle $v_x^\lambda = \mathbb{E} \left[\sum_{n=0}^{+\infty} \frac{\lambda - \lambda_0}{\lambda^{n+1}} c_{X_n} | X_0 = x \right]$ est bornée :

$$\begin{aligned} |v_x^\lambda| &= \left| \mathbb{E} \left[\sum_{n=0}^{+\infty} \frac{\lambda - \lambda_0}{\lambda^{n+1}} c_{X_n} | X_0 = x \right] \right| \leq \frac{|\lambda - \lambda_0|}{|\lambda|} \sup_x |c_x| \frac{1}{1 - |\frac{1}{\lambda}|}, \\ &= \sup_x |c_x| \frac{|\lambda - \lambda_0|}{|\lambda| - 1} = \sup_x |c_x|. \end{aligned}$$

D'autre part v^λ est solution de l'équation de Kolmogorov (1.16) :

$$(M - \lambda)v^\lambda + (\lambda - \lambda_0)c = 0,$$

et donc $|v^\lambda| = |(\lambda - \lambda_0)R^\lambda(M)c|$. En utilisant $\alpha = 1 + \epsilon$ avec $\epsilon > 0$ et en utilisant le théorème précédent on voit que v^λ ne peut rester bornée dans tout voisinage de λ_0 que si le nilpotent de M associé à la valeur propre λ_0 est nul. □

Remarque : Notons que l'on conclut de cela que dans la décomposition de Jordan d'une matrice M stochastique, le bloc correspondant à la valeur propre 1 (de multiplicité m) de la matrice M est un bloc diagonal. Les noyaux de $A = M - I$ ($\mathcal{N}(A)$) et de A^T ($\mathcal{N}(A^T)$) sont tous deux de dimension m . Il en est de même pour les autres valeurs propres de module 1. □

1.4.3 Une autre caractérisation de P

On a également un théorème ergodique qui peut être vu comme un corollaire du théorème 1.23.

Corollaire 1.4.2 (Théorème ergodique).

$$\lim_{k \rightarrow \infty} \frac{1}{k} [I + M + M^2 + \dots + M^{k-1}] = P,$$

où M désigne une matrice stochastique et P le projecteur spectral sur l'espace propre associé à la valeur propre 1 de M . Ce résultat s'énonce aussi en disant que P est la limite au sens de Césaro de $(M^n)_{n \in \mathbb{N}}$.

Preuve : Nous donnons ici une preuve succincte qui suit la preuve de [13]. Les projecteurs spectraux et les nilpotents permettent d'écrire M après changement de base sous la forme : $M = \sum_{\lambda_i \in \mathcal{S}_M} (\lambda_i P_i + N_i)$

On note que sous cette forme : $M^n = \sum_{\lambda_i \in \mathcal{S}_M} (\lambda_i P_i + N_i)^n$

On a $\lim_{k \rightarrow \infty} \frac{1}{k} (\lambda_i P_i + N_i)^k = 0$ dès que $|\lambda_i| < 1$.

D'autre part en utilisant le fait que les nilpotents associés aux valeurs propres de module 1 sont nuls on obtient :

$$\frac{1}{m} [I + M + \dots + M^{m-1}] = \sum_{i, |\lambda_i|=1} \frac{1}{m} \sum_{k=0}^{m-1} \lambda_i^k P_i + \epsilon .$$

Mais

$$\frac{1}{m} \sum_{k=0}^{m-1} \lambda_i^k = \begin{cases} \frac{1}{m} \frac{\lambda_i^m - 1}{\lambda_i - 1} & \text{si } \lambda_i \neq 1 , \\ 1 & \text{si } \lambda_i = 1 , \end{cases}$$

d'où le résultat. □

Remarque : Une chaîne est dite apériodique si le p.g.c.d des m vérifiant $M_{x,x}^m > 0$ vaut 1 pour tout x . Pour une chaîne apériodique le résultat précédent est vrai avec une limite usuelle $\lim_{m \rightarrow \infty} M^m = P$. □

1.5 Problèmes ergodiques

Nous pouvons maintenant étudier l'équation de Kolmogorov dans le cas ergodique c.a.d. lorsque le taux d'actualisation tend vers zéro ou plus classiquement lorsqu'on veut évaluer le coût moyen par unité de temps d'une fonctionnelle additive.

Corollaire 1.5.1. *Si l'on note*

$$w_x^\lambda = \mathbb{E}_x \left[\sum_{n=0}^{+\infty} \frac{\lambda}{(1+\lambda)^{n+1}} c(X_n) \right], \quad \lambda \in \mathbb{R}_+^*, \quad (1.24)$$

on a $\lim_{\lambda \rightarrow 0} w^\lambda = v = Pc$ où P désigne le projecteur spectral sur l'espace propre associé à la valeur propre 1 de M — matrice de transition de la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$.

Preuve : $w^\lambda = \lambda v^\lambda$ et d'après l'équation de Kolmogorov actualisé (1.16), v^λ est solution de l'équation :

$$(A - \lambda)v^\lambda + c = 0 ,$$

où $A = M - I$. La matrice A admet la valeur propre 0 et le nilpotent associé est nul grâce au corollaire (1.4.1). Comme $v^\lambda = -R^\lambda(A)c$ le développement (??) s'écrit ici :

$$v^\lambda = \frac{Pc}{\lambda} - Sc + O(\lambda)$$

et cela termine la preuve. □

Remarque : En utilisant la caractérisation de P du corollaire (1.4.2) on voit que :

$$\lim_{n \rightarrow \infty} \mathbb{E}_x \left[\frac{1}{n} \sum_{k=0}^{n-1} c(X_k) \right] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (M^k c)(x) = (Pc)_x$$

ce qui donne une autre caractérisation de v . \square

La proposition qui suit donne une caractérisation de $v = Pc$ comme solution d'un système linéaire. Cette caractérisation sera utilisée numériquement pour trouver v sans passer par le calcul du projecteur P .

Notons tout d'abord que l'on trouve de façon heuristique le système linéaire de la proposition suivante. On remplace dans l'équation (1.16) v^λ par son développement $e/\lambda + r + O(\lambda)$ pour obtenir le système :

$$Ae = 0 \quad \text{et} \quad Ar - e + c = 0$$

Proposition 1.5.1. *les solutions (e, r) du système :*

$$Ar - e + c = 0 \quad \text{et} \quad Ae = 0 \tag{1.25}$$

sont $e = Pc$ et $r = -Sc + z$ où z est un élément quelconque du noyau de A .

Preuve : Tout d'abord, soient $e = Pc$ et $r = -Sc + z$ on a alors $Ae = APc = 0$ et de l'égalité $AS = I - P$ on déduit $A(-Sc + z) = Pc - c = e - c$ soit $Ar - e + c = 0$. Pour $z \in \mathcal{N}(A)$, le couple $(Pc, -Sc + z)$ est solution de l'équation (1.25).

Réciproquement, soit (e, r) un couple solution de (1.25) on a alors en multipliant la première équation par P , en rajoutant la seconde $(P - A)e = Pc$ (puisque $PA = 0$). Mais il est facile de vérifier que $(P - A)(P - S) = I$. En conséquence $P - A$ est inversible et donc $e = (P - A)^{-1}Pc = (P - S)Pc = Pc$. Comme on a vu que $(Pc, -Sc)$ est une solution particulière de l'équation (1.25) on doit avoir $A(r + Sc) = 0$ et donc $r = -Sc + z$ avec $z \in \mathcal{N}(A)$. \square

1.6 Propriétés combinatoires

Étant donnée une chaîne de Markov de matrice de transition M on lui associe l'application :

$$\begin{aligned} \Gamma : \mathcal{E} &\rightarrow \mathcal{P}(\mathcal{E}), \\ x &\rightarrow \Gamma(x) = \{y \in \mathcal{E} \mid M_{xy} > 0\} \end{aligned}$$

et on associe à la chaîne de Markov un graphe orienté G dont les sommets sont les éléments de \mathcal{E} et dont les arêtes sont $\{(x, y) \mid x \in \mathcal{E} \text{ et } y \in \Gamma(x)\}$. Γ associe à un noeud du graphe l'ensemble des états accessibles par une transition de probabilité non nulle. La matrice d'incidence du graphe G est alors $M_{x,y}^G = \mathbb{I}_{\Gamma(x)}(y)$.

Un chemin de longueur n allant de x^0 à $x^n \in \mathcal{E}$ sera la donnée de $x^0 x^1 \dots x^n$ avec $x^k \in \Gamma(x^{k-1})$. Un circuit sera un chemin revenant à son point de départ. Les chemins de longueur zéro sont réduits aux noeuds.

On pourra montrer à titre d'exercice que l'existence d'un chemin de longueur n joignant x à y se traduit par $M_{x,y}^n > 0$.

On définit alors sur \mathcal{E} :

– la relation d'équivalence :

$$x \sim y \Leftrightarrow x \text{ et } y \text{ appartiennent à un même circuit ;}$$

– le préordre

$$x > y \Leftrightarrow \exists \text{ un chemin de longueur quelconque allant de } x \text{ à } y .$$

Par passage au quotient ce préordre définit une relation d'ordre sur les classes d'équivalences. On a alors les définitions :

- une classe d'équivalence est dite *finale* si elle est minimale pour la relation d'ordre entre les classes,
- une classe non finale est dite *transitoire*,
- un ensemble $A \in E$ est dit *clos* s'il n'existe pas d'arc ayant son origine dans A et son extrémité dans C_A , . Les classes finales sont des ensembles clos.
- une chaîne de Markov ayant une seule classe est dite *irréductible* (pour tout couple $(x, y) \in \mathcal{E}$, Il existe $n > 0$ tel que $M_{x,y}^n > 0$)

Classer les états de la chaîne consiste à trouver les classes d'équivalence transitoires et finales.

1.7 Recherche de composantes fortement connexes

Le problème de Classification des états d'une chaîne de Markov peut se résoudre numériquement en utilisant un algorithme de recherche de composantes fortement connexes d'un graphe. Les composantes fortement connexes du graphe G sont exactement les classes d'équivalence de la relation \sim . Un algorithme de recherche des composantes connexes d'un graphe G basé sur les parcours en profondeur d'abord du graphe G et de son graphe adjoint G^T est décrit dans [4][p-488]. Sa complexité est linéaire en $card(\mathcal{E}) + card(V)$ (ou V est l'ensemble des arcs du graphe).

Il reste ensuite à chercher parmi les composantes connexes celles qui constituent une classe finale et d'agréger les composantes connexes non finales en une seule classe que l'on nommera transitoire (Il n'y a pas vraiment lieu pour l'étude d'une chaîne de distinguer plusieurs classes transitoires).

La classification permet d'obtenir une permutations des états de la chaîne telle que la matrice de transition s'écrive :

$$M = \begin{pmatrix} M_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & M_m & 0 \\ C_1 & \dots & C_m & C_{m+1} \end{pmatrix} \quad (1.26)$$

Les supports des M_i étant les classes finales de la chaîne notés f_i . Le support de C_{m+1} étant la classe transitoire notée T . La proposition qui suit permet de caractériser C_{m+1} : $(I - C_{m+1})$ est une matrice inversible et

$$(I - C_{m+1})^{-1} = \sum_{i=0}^{\infty} C_{m+1}^i. \quad (1.27)$$

On notera que réciproquement si M s'écrit sous la forme précédente avec $I - C_{m+1}$ inversible alors 1 n'est pas valeur propre de C_{m+1} et le support de C_{m+1} ne peut contenir de classes finales.

Proposition 1.7.1. *Toutes les valeurs propres de C_{m+1} sont de module strictement inférieur à 1.*

Preuve : Puisque les états appartenant à T sont transitoires, il existe un chemin de longueur m_x allant de tout élément $x \in T$ à l'une des classes finales f_i , mais alors tout chemin de longueur n plus grande que $\sup_{x \in T} m_x$, va encore de T à l'une des classes finales f_i puisque le chemin précédent ne peut que se prolonger par un chemin dans f_i , il existe donc n tel que :

$$M^n = \begin{pmatrix} M_1^n & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & M_m^n & 0 \\ D_1 & \dots & D_m & C_{m+1}^n \end{pmatrix}$$

possède dans chaque ligne indexée dans T , un élément non nul en dehors de ceux appartenant au bloc diagonal C_{m+1}^n . Et donc la somme des coefficients en ligne de $(C_{m+1})^n$ est strictement plus petite que 1. D'où $|(C_{m+1})^n|_\infty < 1$. On conclut en utilisant le fait que le rayon spectral $\rho(C_{m+1})$ vérifie $\rho(C_{m+1}) \leq (|C_{m+1}^n|_\infty)^{1/n} < 1$. \square

Remarque : La proposition précédente permet de donner un sens à la fonctionnelle de la Proposition 1.2.3 quand $\lambda = 0$ et que le support de c est dans la classe transitoire de la chaîne de Markov. On peut soit remarquer que v_λ admet une limite quand λ tend vers zéro en utilisant l'équation (1.2.3) et la décomposition (1.26) soit remarquer que $\mathbb{E} \left[\sum_{n=0}^N c(X_n) | X_0 = x \right]$ admet une limite quand N tend vers l'infini : $(I - C_{m+1})^{-1} c_T$, où c_T est la restriction de c à la classe T . \square

1.7.1 Un exemple numérique

Nous donnons ici un exemple numérique de calcul de la forme canonique d'une chaîne de Markov à l'aide du logiciel Scilab qui permet des calculs numériques matriciels et des calculs sur les graphes. Dans les cas réels les dimensions des matrices de transition d'une chaîne pouvant être très grandes et il est important de pouvoir utiliser des matrices creuses (ce que l'on peut faire en Scilab) pour leur codage numérique.

-->M

← Une matrice de transition

M =

```
!  0.1  0.7  0.  0.1  0.  0.1 !
!  0.  0.3  0.7  0.  0.  0.  !
!  0.  0.4  0.6  0.  0.  0.  !
!  0.  0.  0.  0.7  0.  0.3 !
!  0.  0.8  0.  0.  0.1  0.1 !
!  0.  0.  0.  0.8  0.  0.2 !
```

```

-->Mb=M<>0 ;                               ← Matrice Booléene associée
-->g=mat_2_graph(sparse(bool2s(Mb)),1,'node-node'); ← Graphe  $G$  associé.
-->show_graph(g);                            ← visualisation du graphe
-->[nc,ncomp]=strong_connex(g);             ← Calcul des composantes connexes

-->nc                                         ← nombre de composantes connexes
nc =

4.
// recherche des classes finales
-->recur=0*ones(1,nc); indsRec=[]; indsT=[];
-->for i=1:nc
--> inds= find(ncomp==i);
--> B1= Mb(inds,:);                          ← Extraction des lignes de la classe  $i$ 
--> B1(:,inds)=[];                            ← si  $\sim B1$  ne contient que
                                                des %t la classe est une classe finale
--> if  $\sim B1$  then recur(i)=1;indsRec=[indsRec,inds];
--> else indsT=[indsT,inds];end
-->end
-->indsT                                       ← Indices des classes transitoires
indsT =

! 2. 4. !

-->ncomp1=ncomp;                               ← On agrège toutes les classes non finale
-->ncomp1(indsT)= maxi(ncomp)+1;

-->perm=[indsRec,indsT];                       ← Permutation des états

-->Mc=M(perm,perm)                             ← Forme canonique de  $M$ 
Mc =

! 0.2 0.8 0. 0. 0. 0. !
! 0.3 0.7 0. 0. 0. 0. !
! 0. 0. 0.3 0.7 0. 0. !
! 0. 0. 0.4 0.6 0. 0. !
! 0.1 0. 0.8 0. 0.1 0. !
! 0.1 0.1 0.7 0. 0. 0.1 !

```

1.8 Mesures de probabilité invariantes

Soit une chaîne de Markov de matrice de transition M , notons $A = M - I$ son générateur. Soit $p \geq 0$ un vecteur ligne tel que $p^T \in \mathcal{N}(A^T)$. Il vérifie $A^T p^T = 0$ donc $pA = 0$ ou encore $pM = p$. C'est une mesure invariante pour M . En utilisant la proposition 1.4 on voit que pour une loi initiale $p^0(x) \equiv p$ invariante par M (quitte à renormaliser p pour que ce soit une

mesure de probabilité), la loi marginale de X_n vaut $p^n(x) = p$.

Supposons que l'on ait m classes finales notées f_i . Pour chacune d'elle M_i est une matrice stochastique et admet donc au moins un vecteur propre à gauche pour M ($\text{card}(\mathcal{N}(A_i^T)) = \text{card}(\mathcal{N}(A_i)) \geq 1$). On admettra qu'en fait il peut être choisi positif (en fait strictement positif) et définir une probabilité invariante q_i , la mesure $r_i = q_i \mathbb{I}_{f_i}$ est alors une mesure invariante pour M . On construit ainsi m vecteurs indépendants (r_1^T, \dots, r_m^T) appartenant à $\mathcal{N}(A^T)$ et on a donc $m \leq \text{card}(\mathcal{N}(A))$.

Inversement, on peut montrer que $\text{card}(\mathcal{N}(A)) \leq m$ ou (m est le nombre de classes finales) en construisant une base de $\mathcal{N}(A^T)$ dont les supports sont des ensembles clos et qui sont des mesures de probabilité. On trouvera une démonstration de ce résultat basée sur le principe du maximum dans [13].

1.9 Calcul de la matrice P sur la forme canonique

Dans la décomposition canonique vu précédemment les matrices M_i sont des matrices de chaîne irréductibles ne contenant qu'une classe finale. D'après le paragraphe précédent, le noyau de A_i est un sous espace de dimension 1 et P_i est un projecteur sur cet espace propre. On vérifie qu'il s'écrit $P_i = e_i r_i$ avec $e_i = (1, \dots, 1)^T$ et r_i est la mesure de probabilité invariante associée à M_i (e_i et r_i^T sont respectivement des vecteurs propres à droite et à gauche associés à la valeur propre 1).

On a vu que $P = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} M^i$ et en appliquant cela à la sous matrice C_{m+1} et en utilisant la proposition 1.7.1 on voit que $P \equiv 0$ sur le support de la classe transitoire T.

On a donc :

$$P = \begin{pmatrix} P_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & P_m & 0 \\ C_1^* & \dots & C_m^* & 0 \end{pmatrix} \quad (1.28)$$

et, en utilisant le fait que $MP = P$ (1.23), on obtient facilement : $C_i^* = (I - C_{m+1})^{-1} C_i P_i$ pour $i = 1, m$.

Remarque : Quand il y a une seule classe récurrente et une transiente : P est un projecteur sur un espace propre de dimension 1 engendré par e et a donc toutes ses lignes identiques. Dans ce cas particulier on a donc $C_1^* = P_1$. \square

Nous reprenons ici l'exemple de la session précédente et nous utilisons directement ce qui précède pour obtenir un calcul numérique de P . Pour chaque classe finale on cherche une solution r_i positive de $(M_i^T - I)r_i^T = 0$. On sait que l'ensemble des solutions est un espace de dimension 1 et que les solutions sont de la forme λr où r peut être choisi positif. En rajoutant la contrainte $(1, \dots, 1)r_i^T = 1$ il y a unicité de la solution et la solution vérifiera bien $r_i^T \geq 0$. On cherche donc numériquement une solution au système linéaire :

$$\begin{pmatrix} M_i^T - I \\ 1 \dots 1 \end{pmatrix} r_i^T = \begin{pmatrix} 0 \\ 1 \dots 1 \end{pmatrix} r_i^T =$$

-->M1=Mc(1:2,1:2);

← calcul de P_1

```

-->M1=[M1'-eye(2,2);ones(1,2)];
-->mu1=(M1)\ [zeros(2,1);1];          ← résolution d'un système linéaire
-->P1=ones(2,1)*mu1'

-->M2=Mc(3:4,3:4);                    ← calcul de P2
-->M2=[M2'-eye(2,2);ones(1,2)];
-->mu2=(M2)\ [zeros(2,1);1];P2=ones(2,1)*mu2'
                                     ← calcul des lignes sur la classe transitoire
-->C1star= (eye(2,2)- Mc(5:6,5:6))\ (Mc(5:6,1:2)*P1);          ← C1*
-->C2star= (eye(2,2)- Mc(5:6,5:6))\ (Mc(5:6,3:4)*P2);          ← C2*
-->P=[P1,zeros(2,4);zeros(P1),P2,zeros(2,2);C1star,C2star,zeros(2,2)]
P =

!   0.2727273   0.7272727   0.           0.           0.           0.   !
!   0.2727273   0.7272727   0.           0.           0.           0.   !
!   0.           0.           0.3636364   0.6363636   0.           0.   !
!   0.           0.           0.3636364   0.6363636   0.           0.   !
!   0.0303030   0.0808081   0.3232323   0.5656566   0.           0.   !
!   0.0606061   0.1616162   0.2828283   0.4949495   0.           0.   !

```

Le calcul numérique de la matrice P permet de résoudre numériquement des problèmes ergodiques.

Nous montrons ici que la caractérisation par l'équation (1.25) permet de calculer directement $v = Pc$ sans passer par l'intermédiaire de la matrice P .

Notons tout d'abord que le calcul des composantes de $v \in \mathcal{N}(A)$ pour $x \in T$ se déduit du calcul de v sur les classes finales f_i (cela résulte de l'inversibilité de $C_{m+1} - I$ proposition 1.7.1 dans forme canonique (1.26)).

Sur chaque classe finale f_i , $v \in \mathcal{N}(A)$ se traduit par le fait que $v = v_i e_{f_i}$ où v_i est un scalaire. On doit résoudre $v_i e_{f_i} + A w_{f_i} = c_{f_i}$ et on sait que $w_{f_i} = (Sc)_{|_{f_i}} + \alpha_i e_{f_i}$. w_{f_i} est donc connu à une constante près et il suffit de fixer une composante arbitraire de w_{f_i} pour se ramener à un système linéaire qui aura une unique solution. Reprenons notre exemple numérique :

```

-->c=[1:6]'                            ← le vecteur c
-->A1=Mc(1:2,1:2)-eye(2,2)              ← Classe 1
-->B1=A1;B1(:,1)=[1;1];                ← on fixe la première
    composante de w à zéro et l'équation à résoudre
    devient B1 ( v1, w1 )^T = c1
-->x=B1\ c(1:2)                          ← résolution du système linéaire
-->v1=x(1)*ones(2,1); w1=[0;x(2)];

-->A2=Mc(3:4,3:4)-eye(2,2)              ← Classe 2
-->A2(:,1)=[1;1];                       ← A2 ( v2, w2 )^T = c2
-->x=A2\ c(3:4)                          ← résolution du système linéaire
-->v2=x(1)*ones(2,1); w2=[0;x(2)];

                                     calcul de v sur la classe transitoire
-->vt= -(Mc(5:6,5:6)-eye(2,2))\ (Mc(5:6,1:4)*[v1;v2]);

```

```

-->wt= -(Mc(5:6,5:6)-eye(2,2))\ (Mc(5:6,1:4)*[w1;w2]+vt -c(5:6));
-->v=[v1;v2;vt]; w = [w1;w2;wt]
-->norm( (Mc- eye(6,6))*v) ← vérification Av = 0
ans =

1.110E-16
-->norm (v+ (Mc- eye(6,6))*w -c) ← vérification v + Aw = c
ans =

0.

```

L'exercice qui suit permet de montrer que l'on peut construire une chaîne de Markov apériodique et irréductible ayant une unique mesure invariante donnée π . Ce résultat peut être utilisé pour des simulations de tirages de loi donnée π [9] et à donné lieu à de nombreux développements (Mcmc, Monte Carlo Markov Chains).

Exercice 1.9.1. Soit $\pi(x)$ une loi de probabilité donnée sur \mathcal{E} telle que $\pi(x) > 0$. On se donne une Matrice stochastique Q avec $Q_{x,y} > 0$ et on définit une matrice de transition :

$$M_{x,y} = Q_{x,y} \min \left(1, \frac{\pi(y)Q_{y,x}}{\pi(x)Q_{x,y}} \right)$$

Montrer que la propriété suivante est vérifiée :

$$\pi(x)M_{x,y} = \pi(y)M_{y,x}$$

Montrer que si M vérifie la propriété précédente alors $\pi = \pi M$. Montrer que la chaîne de matrice de transition M est apériodique (voir remarque page (16)) et irréductible et en déduire que $\lim_{m \rightarrow \infty} M_{x,y}^m = \pi(y)$.

1.10 Interprétation probabiliste d'une base de $\mathcal{N}(A)$

Soit ici $\nu_T = \inf\{n \geq 0 \mid X_n \notin T\}$ où T est l'ensemble des états transitoires et considérons le processus tué $Y_n = X_n \mathbb{I}_{\{n \leq \nu_T\}} + \Delta \mathbb{I}_{\{n > \nu_T\}}$ (On notera que le processus tué ne reste que pour au plus une seule valeur de n dans une classe finale du processus X_n). On regarde ici les probabilités d'atteindre une classe finale donnée en partant d'un état transitoire. Soit donc $F = \bigcup_{i=1}^m f_i$ et $\Gamma \subset F$. On vérifie facilement que :

$$\begin{aligned} v_\Gamma(x) &= \mathbb{E}_x [\mathbb{I}_\Gamma(X_{\nu_T}) \mathbb{I}_{\{\nu_T < \infty\}}] = v_\Gamma^0(x) \quad \text{avec} \\ v_\Gamma^\lambda(x) &\equiv \mathbb{E}_x \left[\sum_{n=0}^{\infty} \frac{1}{(1+\lambda)^{n+1}} \mathbb{I}_\Gamma(Y_n) \right] \end{aligned} \quad (1.29)$$

$v_\Gamma^\lambda(x)$ pour tout $\lambda > 0$ est solution de l'équation de Kolmogorov $(\bar{A} - \lambda I)v + \mathbb{I}_\Gamma(x) = 0$ sur l'espace d'état $\mathcal{E} \cup \{\Delta\}$. La matrice de transition de la chaîne Y_n s'écrit en utilisant les mêmes notations que dans (1.26) :

$$\bar{M} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ e_F & 0 & 0 & 0 & 0 \\ 0 & C_1 & \dots & C_m & C_{m+1} \end{pmatrix}$$

La première ligne et la première colonne correspondant a l'état Δ et e_F désignant un vecteur colonne dont toutes les composantes valent 1 et ce vecteur parcourt les lignes des états finaux. En utilisant la forme particulière de \overline{M} on voit que :

$$\begin{aligned} v_\Gamma^\lambda(\Delta) &= 0 \\ v_\Gamma^\lambda(x) &= \mathbb{I}_\Gamma(x)/(1+\lambda) \quad \text{pour } x \in F \\ v_\Gamma^\lambda|_T &= -(C_{m+1} - (1+\lambda)I)^{-1}(C_1, \dots, C_m)v_\Gamma^\lambda|_F \end{aligned}$$

La matrice $(C_{m+1} - I)$ est inversible et on en déduit que v_Γ^λ admet une limite quand $\lambda \mapsto 0$. On vérifie que cette limite w_Γ (en fait sa restriction à \mathcal{E}) vérifie :

$$\begin{cases} (Aw_\Gamma)(x) = 0 & \text{pour } x \in \mathcal{C}_T, \\ w_\Gamma(x) = 1 & \text{pour } x \in \Gamma \\ w_\Gamma(x) = 0 & \text{pour } x \in F \setminus \Gamma \end{cases} \quad (1.30)$$

Si Γ est une union de f_i , $\Gamma \equiv \bigcup_{i \in I} f_i$, on déduit de l'équation précédente que w_Γ vérifie aussi $Aw_\Gamma = 0$ on a donc $w_\Gamma \in \mathcal{N}(A)$. Il suffit alors d'appliquer le théorème de convergence dominée aux fonctions $g_\Gamma^\lambda = \sum_{n=0}^{\infty} \frac{1}{(1+\lambda)^n} \mathbb{I}_\Gamma(Y_n)$ (les séries sont en fait réduites à un point et $|g_\Gamma^\lambda| \leq 1$) pour conclure que $w_\Gamma = v_\Gamma^0(x)$.

On peut construire ainsi m vecteur $(v_j(x))_{j=1,m}$ indépendants, dans $\mathcal{N}(A)$ donnés par $v_j \equiv v_{f_j}$ et qui ont une interprétation probabiliste donnée par (1.29) . Comme m le nombre de classes finales est aussi la dimension de $\mathcal{N}(A)$ (voir paragraphe 1.8), on construit ainsi une base de $\mathcal{N}(A)$.

Si l'on choisit $\Gamma \equiv F = \bigcup_{i=1}^m f_i$, il est facile de voir que la solution v_F obtenue vérifie $v_F(x) = 1$ pour tout $x \in \mathcal{E}$. On a donc :

$$\mathbb{E}_x [\mathbb{I}_F(X_{\nu_T})\mathbb{I}_{\{\nu_T < \infty\}}] = 1$$

mais par définition de ν_T on a aussi $\mathbb{I}_T(X_{\nu_T})\mathbb{I}_{\{\nu_T < \infty\}} = 0$ d'où :

$$\mathbb{E}_x [\mathbb{I}_{\{\nu_T < \infty\}}] = 1$$

On a donc aussi montré que la Probabilité de quitter l'ensemble des états transitoires vaut 1.

Remarque : De façon plus générale on peut montrer que la fonctionnelle :

$$v(x) = \mathbb{E}_x [c(X_{\nu_B})\mathbb{I}_{\{\nu_B < \infty\}}] \quad \text{avec } B \subset T$$

est solution de l'équation :

$$\begin{aligned} v(x) &= c(x) \quad \text{pour } x \in C_B \\ v(x) &= (I - M_{BB})^{-1}M_{BT}c|_{C_B} I_B = \left(\sum_{k=0}^{\infty} M_{BB}^k \right) M_{BT}c|_{C_B} I_B \end{aligned}$$

en notant par M_{XY} les matrices $M_{XY} = \{(M_{xy}), x \in X \text{ et } y \in Y\}$. La preuve qui est simple est laissée en exercice au lecteur. \square

Exercice 1.10.1. Montrer que v_Γ est la solution minimale positive parmi les solutions du problème (1.30).

Exercice 1.10.2. *Un joueur dans un casino peut, tant que sa fortune est non nulle, parier $1F$. Il peut alors récupérer $2F$ avec une probabilité p ou perdre sa mise avec une probabilité $1 - p \neq 0$. Le joueur s'arrête quand il est ruiné ou quand sa fortune atteint la valeur E . Interpréter le résultat de cette section sur le problème du joueur. On suppose ici $p = q$, que se passe-t-il si E tends vers l'infini (i.e tant que le joueur a de l'argent il joue et la banque est supposée avoir une provision d'argent infinie).*

Remarque : Nous reprenons maintenant l'équation (1.19) en supposant ici $\mathbb{E}[\nu_B] < \infty$. On doit alors avoir $B \subset T$ et donc la matrice $M_{BB} - I$ est inversible (Proposition 1.7.1). L'équation (1.19) pour $\lambda = 0$ a alors une solution unique. On conclut que sous l'hypothèse $\mathbb{E}[\nu_B] < \infty$ la Proposition 1.2.4 est encore valable pour $\lambda = 0$ et $v = \lim_{\lambda \downarrow 0} v_\lambda$. \square

1.11 États transients et états récurrents

Soit $\tau_\Gamma^{p+1} = \inf\{n; n > \tau_\Gamma^p, X_n \in \Gamma\}$ avec $\tau_\Gamma^0 = 0$. On montre facilement que sur $\{\tau_\Gamma^p < \infty\}$, $\tau_\Gamma^{p+1} = \tau_\Gamma^p + \tau_\Gamma \circ \theta_{\tau_\Gamma^p}$ où θ est l'opérateur de translation et $\tau_\Gamma = \tau_\Gamma^1$. Pour $\Gamma \equiv \{x\}$ on notera $\tau_x^p \equiv \tau_{\{x\}}^p$.

Nous montrons dans ce paragraphe que l'on peut utiliser les résultats précédents sur les équations de Kolmogorov pour caractériser les éléments des classes finales par $\mathbb{P}_x \{\tau_x < \infty\} = 1$ (on parle aussi d'états récurrents) et les états des classes transitoires par $\mathbb{P}_x \{\tau_x < \infty\} < 1$ (on parle aussi dans ce cas d'états transients).

Nous commençons par un exercice :

Exercice 1.11.1. *Montrer la propriété suivante*

$$\mathbb{P}_x \{\tau_x^p < \infty\} = (\mathbb{P}_x \{\tau_x < \infty\})^p$$

et en déduire la formule :

$$\mathbb{E}_x \left[\sum_{n=0}^{+\infty} \mathbb{I}_x(X_n) \right] = \sum_{n=0}^{\infty} (\mathbb{P}_x \{\tau_x < \infty\})^n$$

Théorème 1.11.1. *Soit $x \in f_i$, on a alors $\mathbb{P}_x \{\tau_x < \infty\} = 1$. Soit $x \in T$, on a alors $\mathbb{P}_x \{\tau_x < \infty\} < 1$.*

Preuve : Considérons tout d'abord l'équation (1.24) avec $c(x) \equiv \mathbb{I}_\Gamma(x)$. On a vu qu'alors $\lim_{\lambda \rightarrow 0} v^\lambda = v = P\mathbb{I}_\Gamma$. Considérons ici $x \in f_i$ et $\Gamma \cap f_i \neq \emptyset$. En utilisant l'équation (1.28) et la forme particulière de P_i , on déduit que $P\mathbb{I}_\Gamma(x) \neq 0$. D'où :

$$\lim_{\lambda \rightarrow 0} \mathbb{E}_x \left[\sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} \mathbb{I}_\Gamma(X_n) \right] = \infty \quad \text{pour } x \in f_i \quad \Gamma \cap f_i \neq \emptyset$$

On a alors aussi $\mathbb{E}_x \left[\sum_{n=0}^{+\infty} \mathbb{I}_\Gamma(X_n) \right] = \infty$ (sinon on pourrait appliquer le théorème de Lebesgue à la suite $f_k = \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda_k)^{n+1}} \mathbb{I}_\Gamma(X_n)$ pour λ_k une suite décroissant vers 0 et arriver à une contradiction). En utilisant $\Gamma \equiv \{x\}$ et l'exercice 1.11.1 on conclut que $\mathbb{P}_x \{\tau_x < \infty\} = 1$.

Considérons maintenant le problème actualisé en horizon infini :

$$w_x^\lambda = \mathbb{E}_x \left[\sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} \mathbb{I}_\Gamma(X_n) \right],$$

On a vu au paragraphe 1.2.3 que w^λ est solution de l'équation (1.16) :

$$(M - (1+\lambda)I)w^\lambda + \mathbb{I}_\Gamma = 0.$$

On suppose ici $\Gamma \subset T$, la solution est alors nécessairement à support dans T et en appelant $w|_T$ la restriction de w à T on doit avoir :

$$(\lambda I - (C_{m+1} - I))w|_T^\lambda = \mathbb{I}_\Gamma$$

$C_{m+1} - I$ est une matrice inversible et on en déduit que :

$$\lim_{\lambda \rightarrow 0} w|_T^\lambda = (I - C_{m+1})^{-1} \mathbb{I}_\Gamma.$$

En utilisant une suite λ_k qui décroît vers 0 et le théorème de Bepo-Levi [?] On conclut que $\mathbb{E}_x \left[\sum_{n=0}^{+\infty} \mathbb{I}_\Gamma(X_n) \right] < \infty$ pour $\Gamma \subset T$, et à nouveau en utilisant l'exercice 1.11.1 on conclut que $\mathbb{P}_x \{ \tau_x < \infty \} < 1$. \square

Remarque : La fonction $w = \lim_{\lambda \rightarrow 0} w^\lambda$ que l'on a construite $w(x) \equiv (I - C_{m+1})^{-1} \mathbb{I}_\Gamma(x)$ est une fonction positive solution de l'équation $Av + \mathbb{I}_\Gamma = 0$. Cette équation n'admet pas une unique solution puisque $v(x) = e$ est dans le noyau de A . Soit z une autre solution positive de $Av + \mathbb{I}_\Gamma = 0$ on a alors $(A - \lambda I)(w^\lambda - z) = \lambda z \geq 0$ et du principe du maximum strictement positif on déduit $w^\lambda \leq z$. w est en fait la solution minimale positive de $Av + \mathbb{I}_\Gamma = 0$. \square

1.12 Corrigé des exercices

Exercice 1.1.1 Considérons tout d'abord le cas $\sigma(n+1) = \sigma(n) + 1$. On considère la partition des entiers $[0, \sigma(n)]$ en $I = (\sigma(i))_{i=1, \dots, n}$ et $J = [1, \sigma(n)] \setminus I$. On note $B = \{X_{\sigma(n)} = x_{\sigma(n)}, \dots, X_{\sigma(1)} = x_{\sigma(1)}\}$ et $C_y = \{X_i = y^i\}_{i \in J}$ et $A = \{X_{\sigma(n+1)} = x_{\sigma(n+1)}\}$. On suppose $\mathbb{P}\{B\} \neq 0$, par la formule de décomposition on obtient :

$$\mathbb{P}\{A|B\} = \sum_{y_j \in \mathcal{E}, j \in J} \mathbb{P}\{A|B \cap C_y\} \mathbb{P}\{B \cap C_y\} \mathbb{P}\{B\}^{-1}$$

Par la propriété de Markov $\mathbb{P}\{A|B \cap C_y\} = \mathbb{P}\{A|X_{\sigma(n)} = x_{\sigma(n)}\}$ et donc :

$$\begin{aligned} \mathbb{P}\{A|B\} &= \sum_{y_j \in \mathcal{E}, j \in J} \mathbb{P}\{A|X_{\sigma(n)} = x_{\sigma(n)}\} \mathbb{P}\{B \cap C_y\} \mathbb{P}\{B\}^{-1} \\ &= \mathbb{P}\{A|X_{\sigma(n)} = x_{\sigma(n)}\} \end{aligned}$$

On montre ensuite par récurrence sur p la formule pour $\sigma(n+1) = \sigma(n) + p$. On pose $A_p(x) = \{X_{\sigma(n)+p} = x\}$ et on suppose la formule vérifiée pour $A_p(\cdot)$.

$$\mathbb{P}\{A_{p+1}(x)|B\} = \sum_{y \in \mathcal{E}} \mathbb{P}\{A_{p+1}(x)|A_p(y) \cap B\} \mathbb{P}\{A_p(y)|B\}$$

La récurrence à l'étape p assure :

$$\mathbb{P}\{A_p(y)|B\} = \mathbb{P}\{A_p(y)|X_{\sigma(n)} = x_{\sigma(n)}\}$$

et la première partie assure

$$\mathbb{P}\{A_{p+1}(x)|A_p(y) \cap B\} = \mathbb{P}\{A_{p+1}(x)|A_p(y)\} = \mathbb{P}\{A_{p+1}|A_p(y) \cap X_{\sigma(n)} = x_{\sigma(n)}\}.$$

On a donc

$$\begin{aligned} \mathbb{P}\{A_{p+1}(x)|B\} &= \sum_{y \in \mathcal{E}} \mathbb{P}\{A_{p+1}(x)|A_p(y), X_{\sigma(n)} = x_{\sigma(n)}\} \mathbb{P}\{A_p(y)|X_{\sigma(n)} = x_{\sigma(n)}\} \\ &= \mathbb{P}\{A_{p+1}(x)|X_{\sigma(n)} = x_{\sigma(n)}\} \end{aligned}$$

Exercice 1.1.2

$$\mathbb{E}[f(X_{n+k+p})|X_n] = g(X_n) \quad \text{avec} \quad g(x) \equiv \sum_{z \in \mathcal{E}} f(z) \mathbb{P}\{X_{n+k+p} = z | X_n = x\}$$

On a donc par la formule de décomposition :

$$g(x) = \sum_{z \in \mathcal{E}, y \in \mathcal{E}} f(z) \mathbb{P}\{X_{n+k+p} = z | X_{n+k} = y, X_n = x\} \mathbb{P}\{X_{n+k} = y | X_n = x\}.$$

En utilisant l'exercice précédent (1.1.1)

$$\mathbb{P}\{X_{n+k+p} = z | X_{n+k} = y, X_n = x\} = \mathbb{P}\{X_{n+k+p} = z | X_{n+k} = y\}$$

et donc en posant $h(y) \equiv \mathbb{E}[f(X_{n+k+p})|X_{n+k} = y]$ on a :

$$\begin{aligned} g(x) &= \sum_{y \in \mathcal{E}} h(y) \mathbb{P}\{X_{n+k} = y | X_n = x\}. \\ &= \mathbb{E}[h(X_{n+k})|X_n = x] = \mathbb{E}[\mathbb{E}[f(X_{n+k+p})|X_{n+k}] | X_n = x] \end{aligned}$$

Exercice 1.1.3 Soient $A = \{X_{n+p+k} = \alpha\}$, $B_i = \{X_{n+p} = \beta_i\}$ et $C = \{X_n = \gamma\}$. En utilisant la formule dite des probabilités composées :

$$\mathbb{P}\{A \cap B_i \cap C\} = \mathbb{P}\{A|B_i \cap C\} \mathbb{P}\{B_i|C\} \mathbb{P}\{C\}$$

et la formule de décomposition

$$\mathbb{P}\{A \cap C\} = \sum_{i \in \mathbb{E}} \mathbb{P}\{A \cap C \cap B_i\}$$

On obtient :

$$\mathbb{P}\{A|C\} = \sum_{i \in \mathbb{E}} \mathbb{P}\{A|B_i \cap C\} \mathbb{P}\{B_i|C\}$$

La propriété de Markov assure $\mathbb{P}\{A|B_i \cap C\} = \mathbb{P}\{A|B_i\}$ et la formule précédente devient :

$$P_{\gamma, \alpha}^{(n, n+k+p)} = \sum_{i \in \mathbb{E}} P_{\gamma, i}^{(n, n+k)} P_{i, \alpha}^{(n+k, n+k+p)}$$

Exercice 1.2.1 Soit $g_n((X_i)_{i \geq n})$ définie par

$$g_n \equiv \sum_{i=n}^{N-1} \left(\prod_{k=n}^{i-1} \xi^k(X_k) \right) c^i(X_i) + \left(\prod_{k=n}^{N-1} \xi^k(X_k) \right) f(N, X_N),$$

on a alors $g_n(\cdot) = c^n(X_n) + \xi^n(X_n)g_{n+1}(\cdot)$. On conclut en raisonnant comme dans la preuve de la proposition 1.9.

Exercice 1.2.2 Soit $Y_n = (X_n, L_n)$ on peut toujours supposer que la chaîne X_n est donnée par $X_{n+1} = f(X_n, U_n)$ on a alors :

$$Y_{n+1} = F(Y_n, U_n) \quad \text{avec} \quad F((x, l), u) = \begin{pmatrix} f(x, u) \\ l + f(x, u) \end{pmatrix}$$

Y_n est donc bien une chaîne de Markov et le problème à résoudre s'écrit :

$$v_n(x, l) = \mathbb{E} [\phi((Y_N)_2) | Y_n = (x, l)]$$

L'équation de récurrence vérifiée par v_n s'écrit alors en notant par P la matrice de transition de la chaîne X_n :

$$\begin{aligned} v_n(x, l) &= \sum_{x' \in \mathcal{E}} P_{x, x'} v_{n+1}(x', l + x') \\ v_N(x, l) &= \phi(l) \end{aligned}$$

Le raisonnement est le même pour $L_n = \max_{i=0}^n X_i$ où à nouveau $Y_n = (X_n, L_n)$ est une chaîne de Markov compte tenu de la formule :

$$Y_{n+1} = F(Y_n, U_n) \quad \text{avec} \quad F((x, l), u) = \begin{pmatrix} f(x, u) \\ \max(l, f(x, u)) \end{pmatrix}$$

de même pour $L_n = \prod_{i=0}^n$ par :

$$Y_{n+1} = F(Y_n, U_n) \quad \text{avec} \quad F((x, l), u) = \begin{pmatrix} f(x, u) \\ l * f(x, u) \end{pmatrix}$$

Exercice 1.2.3 Soit $B = \mathcal{E} \cap (\infty, K]$ et $\nu_B^n \equiv \inf\{k \geq n : X_k \notin B\}$, $v_n(x)$ s'écrit aussi :

$$v_n(x) = \mathbb{E} \left[\frac{1}{(1+r)^{N-n}} \phi(X_N) \mathbb{1}_{\nu_B^n > N} | X_n = x \right]$$

L'équation vérifiée par $v_n(x)$ est donc :

$$\begin{cases} v_n(x) &= \frac{1}{1+r} M^{(n)} v_{n+1}(x), \forall n = 0, \dots, N-1, & \text{pour } x \leq K, \\ v_N(x) &= \phi(x) & \text{pour } x \leq K, \\ v_n(x) &= 0, \forall n = 0, \dots, N & \text{pour } x > K. \end{cases} \quad (1.31)$$

Exercice 1.3.1 Soit L vérifiant le principe de Maximum strictement positif. Pour la première partie on montre facilement par contradiction que le noyau de L est réduit à 0.

Soit $L^{-1}e_j$ la j -ième colonne de la matrice L . Si $\max v > 0$ ($v \equiv L^{-1}e_j$) alors on devrait avoir $(e_j)_\alpha = (LL^{-1}e_j)_\alpha < 0$, pour $\alpha \in \text{Argmax } v$ ce qui n'est pas possible. Les éléments de la j -ième colonne de L^{-1} sont donc tous négatifs.

Exercice 1.9.1

1. La relation $\pi(x)M_{x,y} = \pi(y)M_{y,x}$ se déduit de la définition de M en remarquant que :
 $\pi(x)M_{x,y} = \min(\pi(y)Q_{y,x}, \pi(x)Q_{x,y})$.
2. En utilisant la propriété $\pi(x)M_{x,y} = \pi(y)M_{y,x}$ on obtient $\sum_{y \in \mathcal{E}} \pi(y)M_{y,x} = \sum_{y \in \mathcal{E}} \pi(x)M_{x,y} = \pi(x)$. π est une mesure invariante pour M .
3. $M_{x,x} = Q_{x,x} > 0$ et $M_{x,y} > 0$ la chaîne est donc apériodique et irréductible. On a donc $\lim_{m \rightarrow \infty} M_{x,y}^m = P_{x,y}$. On a vu que dans le cas d'une chaîne irréductible P avait toutes ses lignes identiques et égales à l'unique mesure invariante qui est π d'après la deuxième question.

Exercice 1.10.1 Soit z une solution positive de l'équation (1.30). On peut prolonger z à $\mathcal{E} \cup \{\Delta\}$ par $z(\Delta) = 0$. et on constate alors que z vérifie $\bar{M}z = 0$. On a donc :

$$(\bar{A} - \lambda I)(v_\Gamma^\lambda - z) = \lambda z$$

On a donc $(\bar{A} - \lambda I)(v_\Gamma^\lambda - z) \geq 0$ et du principe du maximum strictement positif on déduit $v_\Gamma^\lambda \leq z$. Il suffit alors de faire tendre λ vers zéro pour conclure.

Exercice 1.11.1 Soit $T = \tau_x^p$, la propriété de Markov forte (1.15) en utilisant le fait que $X_T = x$ sur $\{T < \infty\}$ nous donne :

$$\mathbb{E}[\mathbb{I}_{\tau_x \circ \theta_T < \infty} \mathbb{I}_{T < \infty} | \mathcal{F}_T] = \mathbb{I}_{T < \infty} \mathbb{E}_x[\mathbb{I}_{\tau_x < \infty}]. \quad (1.32)$$

D'où :

$$\begin{aligned} \mathbb{P}_x \{ \tau_x^{p+1} < \infty \} &= \mathbb{E}_x \left[\mathbb{I}_{\{ \tau_x^p < \infty \}} \mathbb{I}_{\{ \tau_x \circ \theta_{\tau_x^p} < \infty \}} \right] = \mathbb{E}_x \left[\mathbb{E} \left[\mathbb{I}_{\{ \tau_x^p < \infty \}} \mathbb{I}_{\{ \tau_x \circ \theta_{\tau_x^p} < \infty \}} | \mathcal{F}_T \right] \right] \\ &= \mathbb{P}_x \{ \tau_x^p < \infty \} \mathbb{P}_x \{ \tau_x < \infty \} \end{aligned}$$

ce qui répond à la première question. Il suffit ensuite de remarquer que par définition de τ_x^p on a :

$$\mathbb{E}_x \left[\sum_{n=0}^{+\infty} \mathbb{I}_x(X_n) \right] = \sum_{n=0}^{\infty} \mathbb{P}_x \{ \tau_x^n < \infty \} \quad (1.33)$$

$$(1.34)$$

qui combiné avec la première question termine l'exercice.

Chapitre 2

Contrôle de chaînes de Markov

On s'intéresse à présent à l'optimisation d'un système dont l'état est décrit par une chaîne de Markov. On considère le cas où la matrice de transition de la chaîne de Markov dépend d'un paramètre de contrôle, et l'on cherche à optimiser, par rapport au contrôle, un certain critère. Le lecteur pourra se référer en particulier à [33] et [2].

On se place dans le cas de l'observation complète, c'est à dire qu'on observe l'état du système à tout instant.

On dit que $(X^n)_{n \in \mathbb{N}}$ est une chaîne (de Markov) contrôlée sur l'espace d'état $\mathcal{E} = \{1, 2, \dots, E\}$, si les probabilités de transition sont fonctions de variables de contrôles qui prennent leurs valeurs dans un espace \mathcal{C} . On note u l'élément générique de \mathcal{C} . On note M_{xy}^{nu} la probabilité d'aller de l'état x à l'état y si l'on prend la décision $u \in \mathcal{C}$ à l'instant n (parfois la décision à l'instant n quand on se trouve dans l'état x sera choisie dans $\mathcal{C}_x^n \subset \mathcal{C}$).

Pour tout n , le contrôle est une variable aléatoire que l'on note U^n . Soit $\mathbb{U} = (U^0, U^1, \dots)$ la suite des variables aléatoires à valeurs dans \mathcal{C} qui sont les contrôles exercés aux instants $0, 1, \dots$, on appellera \mathbb{U} une stratégie.

Une stratégie fixée \mathbb{U} , permet de définir une probabilité sur les trajectoires de la chaîne contrôlée ($\mathbb{P}^{\mathbb{U}}$) et on dira que \mathbb{U} est admissible si la propriété de Markov continue à être vérifiée sous l'action de \mathbb{U} .

De façon générale le contrôle U^n se définit par la donnée d'une loi de probabilité $q_{U^n(H^n)}(u)$ sur l'espace \mathcal{C} qui pourra dépendre de la trajectoire jusqu'à l'instant n et des contrôles passés $H^n = (H^{n-1}, U^{n-1}, X^n)$ (avec $H^0 = X^0$). On notera \mathcal{D}^{HR} l'ensemble des contrôles et \mathcal{U}^{HR} l'ensemble des stratégies précédentes, \mathcal{U}^{HD} désignera le cas où $q_{U^n(H^n)}(u)$ est concentrée en un point de \mathcal{C} ce qui veut dire que le contrôle ($U^n \in \mathcal{D}^{HD}$) à l'instant n est une fonction de l'ensemble des trajectoires possibles sur $[0, n]$ du processus contrôlé à valeur dans \mathcal{C} .

\mathcal{U}^{MR} désigne les stratégies Markoviennes relaxées, le contrôle ($U^n \in \mathcal{D}^{MR}$) à l'instant n ne dépend que de l'état courant du processus. Il est donné par une fonction sur \mathcal{E} à valeur dans les lois de probabilité sur \mathcal{C} ($x \mapsto q_{U^n(x)}(u)$). Enfin \mathcal{U}^{MD} désigne les stratégies Markoviennes déterministes, $q_{U^n(x)}(u)$ est concentrée en un point de \mathcal{C} pour chaque x . Le contrôle (\mathcal{D}^{MR}) à l'instant n est une fonction de \mathcal{E} à valeur dans \mathcal{C} et on le notera simplement $U_n(x)$. On parle aussi dans ce cas d'ensemble de stratégies feedback.

Dans le cas de l'observation complète pour les problèmes présentés dans ce chapitre, nous montrerons que l'on ne perd rien à optimiser dans la classe des stratégies feedback (aussi appelées boucles fermées sur l'état) c'est à dire dans l'ensemble \mathcal{U}^{MD} , Dans ce cas on utilise la même notation U pour référer aussi bien à la fonction qu'à la suite des contrôles.

Pour une stratégie \mathbb{U} donnée et une loi initiale p_0 , l'évolution de la chaîne sous la loi induite s'exprime par :

$$\begin{aligned} \mathbb{P}^{\mathbb{U}} \{U_t = u | H_t = h_t\} &= q_{U_t(h_t)}(u) \quad , \quad \mathbb{P}^{\mathbb{U}} \{X_0 = s\} = p_0(s) \\ \text{et} \quad \mathbb{P}^{\mathbb{U}} \{X^{t+1} = s | H_t = (h_{t-1}, u_{t-1}, x_t), U_t = u\} &= M_{x_t, s}^{(n), u}. \end{aligned} \quad (2.1)$$

On supposera que pour u fixé $M^{(n), u}$ est bien une matrice stochastique. Comme pour les chaînes de Markov, on peut construire une représentation canonique du processus contrôlé [5].

Le problème de commande optimale consiste à optimiser un critère s'exprimant sous la forme d'une fonctionnelle de la trajectoire de la chaîne, sur une période de temps finie ou infinie, c'est à dire résoudre des problèmes du type :

$$\min_{\mathbb{U} \in \mathcal{U}(\cdot)} \mathbb{E} \left[\sum_{n=0}^{N-1} c^{U^n}(X^n) | X^0 = x \right] ,$$

où N , fini ou infini, est appelé l'horizon du problème, et la fonction c est appelée le coût instantané. L'horizon peut être un temps d'arrêt et devenir lui même un contrôle à optimiser.

Dans le cas des stratégies Markoviennes on utilisera les notations suivantes :

$$\begin{aligned} c^U(x) &= \sum_{u \in \mathcal{C}_x} c^u(x) q_{U(x)}(u), \quad \text{pour } U \in \mathcal{D}^{MR} \\ c^U(x) &= c^{U(x)}(x), \quad \text{pour } U \in \mathcal{D}^{MD} \\ M_{x,s}^{(n), U} &= \sum_{u \in \mathcal{C}_x} M_{x,s}^{(n), u} q_{U(x)}(u) \quad \text{pour } U \in \mathcal{D}^{MR} \\ M_{x,s}^{(n), U} &= M_{x,s}^{(n), U(x)} \quad \text{pour } U \in \mathcal{D}^{MD} \end{aligned} \quad (2.2)$$

La chaîne contrôlée pour une stratégie Markovienne relaxée ou non est une chaîne de Markov de matrice de transition $M_{x,s}^{(n), U}$ et l'évaluation pour cette stratégie fixée de $\mathbb{E} \left[\sum_{n=0}^{N-1} c^{U^n}(X^n) \right]$ se traite comme au chapitre précédent avec la fonction coût instantané $c^{U^n}(x)$ définie précédemment.

Pour les problèmes stationnaires nous serons amené à rechercher des stratégies elle mêmes stationnaires. On remplacera la notation \mathcal{U} par \mathcal{S} pour désigner ces stratégies.

2.1 Problème de temps d'arrêt optimal

Le problème de temps d'arrêt optimal est un problème de contrôle où à tout instant, les deux seuls contrôles possibles sont soit arrêter le processus soit le laisser continuer. Soit (X^n) une chaîne de Markov de matrice de transition $M^{(n)}$ sur un espace d'état fini \mathcal{E} . Soit $\psi^n(x)$ une fonction définie pour $0 \leq n \leq N$ et pour $x \in \mathcal{E}$. On cherche à évaluer

$$u_n(x) = \sup_{\tau \text{ } \mathcal{F}_n \text{ t.a., } n \leq \tau \leq N} \mathbb{E} \left[\sum_{k=n}^{\tau-1} c^{(k)}(X^k) + \psi^\tau(X^\tau) | X^n = x \right].$$

La variable de commande est ici un temps d'arrêt $n \leq \tau \leq N$.

Théorème 2.1.1. *La fonction u_n , pour $n \in [0, N]$, est l'unique solution de l'équation :*

$$\begin{cases} u_n(x) &= \max(M^{(n)}u_{n+1} + c^{(n)}, \psi^n), \quad n = 0, \dots, N-1 \\ u_N &= \psi^N. \end{cases} \quad (2.3)$$

De plus, le temps d'arrêt

$$\tau_n = \inf\{N \geq k \geq n, u_k(X^k) = \psi^k(X^k)\}$$

est tel que

$$u_n(X^n) = \mathbb{E} \left[\sum_{k=n}^{\tau_n-1} c^{(k)}(X^k) + \psi^{\tau_n}(X^{\tau_n}) | X^n \right].$$

On dit que τ_n est un temps d'arrêt optimal.

Ce résultat peut se prouver par l'approche contrôle stochastique par le principe de programmation dynamique ou bien par une méthode probabiliste (par les martingales). Nous utilisons une approche programmation dynamique pour établir l'équation d'optimalité et nous utilisons une preuve par les martingales pour caractériser la stratégie optimale.

Preuve : Nous montrons tout d'abord que u_n est solution de l'équation d'optimalité (l'unicité de la solution étant un point évident). Soit τ un temps d'arrêt $n \leq \tau \leq N$ on a alors :

$$\begin{aligned} & \mathbb{E} \left[\sum_{k=n}^{\tau-1} c^{(k)}(X^k) + \psi^\tau(X^\tau) | X^n \right] \\ &= \mathbb{E} [\mathbb{I}_{\{\tau=n\}} \psi^\tau(X^\tau) | X^n] + \mathbb{E} \left[\mathbb{I}_{\{\tau>n\}} \left(\sum_{k=n}^{\tau-1} c^{(k)}(X^k) + \psi^\tau(X^\tau) \right) | X^n \right] \\ &= \psi^n(X^n) \mathbb{I}_{\{\tau=n\}} + \mathbb{I}_{\{\tau>n\}} \mathbb{E} \left[\left(\sum_{k=n}^{\tau'-1} c^{(k)}(X^k) + \psi^{\tau'}(X^{\tau'}) \right) | X^n \right] \\ & \quad (\text{en posant } \tau' \equiv \max(n+1, \tau)) \\ &\leq \psi^n(X^n) \mathbb{I}_{\{\tau=n\}} + \mathbb{I}_{\{\tau>n\}} (c^{(n)}(X^n) + \mathbb{E} [u_{n+1}(X^{n+1}) | X^n]) \\ &\leq \max(\psi^n(X^n), c^{(n)}(X^n) + M^{(n)}u_{n+1}(X^n)) \end{aligned}$$

On a donc $u_n(x) \leq \max(\psi^n(x), c^{(n)}(x) + M^{(n)}u_{n+1}(x))$. D'autre part, le temps d'arrêt $\tau \equiv n$ est dans \mathcal{F}_n et donc $\psi^n(x) \leq u_n(x)$. Soit maintenant ϵ fixé et τ_ϵ un temps d'arrêt \mathcal{F}_{n+1} mesurable tel que :

$$u_{n+1}(x) \leq \epsilon + \mathbb{E} \left[\left(\sum_{k=n+1}^{\tau_\epsilon-1} c^{(k)}(X^k) + \psi^{\tau_\epsilon}(X^{\tau_\epsilon}) \right) | X^{n+1} = x \right]$$

, on a alors :

$$c^{(n)}(x) + \mathbb{E} [u_{n+1}(X^{n+1}) | X^n = x] \leq \epsilon + c^{(n)}(x) + \mathbb{E} \left[\left(\sum_{k=n+1}^{\tau_\epsilon-1} c^{(k)}(X^k) + \psi^{\tau_\epsilon}(X^{\tau_\epsilon}) \right) | X^n = x \right] \leq u_n(x)$$

ce qui termine la preuve de la première partie.

Montrons maintenant que $\tau_n \equiv \inf\{k \geq n, u_k(X^k) = \psi^k(X^k)\}$ définit une stratégie optimale. Soit le processus discret $(Y_p)_{N \geq p \geq n}$ définie par :

$$Y_p = u_p(X^p)\mathbb{I}_{\{p < \tau_n\}} + \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq p\}} + \sum_{j=n}^{p-1} c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}}.$$

Sachant que $\tau_n \leq N$ on a tout d'abord :

$$\begin{aligned} Y_N &= \psi^{\tau_n}(X^{\tau_n}) + \sum_{j=n}^{N-1} c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}} \\ &= \psi^{\tau_n}(X^{\tau_n}) + \sum_{j=n}^{\tau_n-1} c^{(j)}(X^j). \end{aligned}$$

D'autre part $Y_n = u_n(X^n)\mathbb{I}_{\{n < \tau_n\}} + \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq n\}}$, ce qui, en utilisant le fait que $\tau_n \geq n$ et que $u_n(X^n)\mathbb{I}_{\{\tau_n = n\}} = \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n = n\}}$ nous donne $Y_n = u_n(X^n)$. On a ensuite :

$$\begin{aligned} Y_p &= (M^{(p)}u_{p+1}(X^p) + c^{(p)}(X^p))\mathbb{I}_{\{p < \tau_n\}} + \sum_{j=n}^{p-1} c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}} + \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq p\}} \\ &= \mathbb{E}[u_{p+1}(X^{p+1})|\mathcal{F}_p]\mathbb{I}_{\{p < \tau_n\}} + \sum_{j=n}^p c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}} + \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq p\}} \\ &= \mathbb{E}[u_{p+1}(X^{p+1})\mathbb{I}_{\{p < \tau_n\}}|\mathcal{F}_p] + \mathbb{E}\left[\psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq p\}} + \sum_{j=n}^p c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}}|\mathcal{F}_p\right] \end{aligned}$$

En utilisant maintenant $u_{p+1}(X^{p+1})\mathbb{I}_{\{p+1 = \tau_n\}} = \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{p+1 = \tau_n\}}$, on obtient :

$$\begin{aligned} Y_p &= \mathbb{E}\left[u_{p+1}(X^{p+1})\mathbb{I}_{\{p+1 < \tau_n\}} + \psi^{\tau_n}(X^{\tau_n})\mathbb{I}_{\{\tau_n \leq p+1\}} + \sum_{j=n}^p c^{(j)}(X^j)\mathbb{I}_{\{j < \tau_n\}}|\mathcal{F}_p\right] \\ &= \mathbb{E}[Y_{p+1}|\mathcal{F}_p] \end{aligned}$$

Y_p est donc une martingale discrète et vérifie en conséquence $Y_n = \mathbb{E}[Y_N|\mathcal{F}_n]$. En remplaçant Y_n et Y_N par leurs valeurs, on conclut que :

$$u_n(X_n) = \mathbb{E}\left[\psi^{\tau_n}(X^{\tau_n}) + \sum_{j=n}^{\tau_n-1} c^{(j)}(X^j)|\mathcal{F}_n\right].$$

□

Exercice 2.1.1. On considère ici une version actualisée du problème précédent (avec $c^{(j)} = 0$) :

$$v_n(x) = \sup_{\tau \in \mathcal{F}_n - t.a., n \leq \tau \leq N} \mathbb{E}\left[\frac{1}{(1+r)^{\tau-n}} f^\tau(X_\tau)|X_n = x\right].$$

Montrer que la fonction v est l'unique solution de l'équation

$$\begin{cases} v_n(x) = \max\left(\frac{1}{1+r} Mv_{n+1}, f^n\right) & n = 0, \dots, N-1 \\ u_N = f^N. \end{cases}$$

Et que le temps d'arrêt $\tau_n = \inf\{k \geq n, v_k(X_k) = f^k(X_k)\}$ est un temps d'arrêt optimal. Calculer le prix d'un put américain dans le modèle de Cox Ross qui est donné par la formule suivante :

$$v_0(x) = \sup_{\tau \text{ t.a., } 0 \leq \tau \leq N} \mathbb{E} \left[\frac{1}{(1+r)^\tau} (K - X_\tau)_+ | X_0 = x \right].$$

2.1.1 Application numérique

```

-->function y=f(x); y=max(0,K-x) ; endfunction          ← fonction f
-->pa=0.5; pb=1-pa; X1= 1; a = -0.05 ; b = +0.05; K=1; r=0.01; N=50
  Les points accessibles en i étapes à partir de X1 stockés dans X(i,1:i)
function y=f(x); y=max(0,K-x) ; endfunction
  Les points accessibles en i étapes à partir de X1 stockés dans X(i,1:i)

X = zeros(N,N) ;
X(:,1) = X1*(1+a).^ (0:N-1)';          ← première colonne X(1,i)= X_1(1+a)^(i-1)
for i=2:N                                ← calcul récursif de chaque lignes de X
  X(i,2:i) = X(i-1,1:i-1)*(1+b);
end
  calcul de  $v_n(x)$  aux points donnés par X:  $V(n,i) \equiv v_n(X(n,i))$ 
V = zeros(N,N) ; V(N,:) = f(X(N,:));    ← initialisation

for i=N-1:-1:1
  V(i,1:i) = V(i+1,1:i)*pa/(1+r);
  V(i,1:i) = V(i,1:i)+ (pb/(1+r))* V(i+1,2:i+1);
  V(i,1:i) = max( V(i,1:i),f(X(i,1:i)));
end

// dessin dans le plan
xstop=[];ystop=[]                        ← Points ou il faut s'arreter
xcont=[];ycont=[]                        ← Points ou il faut continuer

for i=1:N
  Xi=X(i,1:i); [m,im]=maxi(f(Xi),V(i,1:i));
  istop=find(im==1);
  xstop=[xstop;i*ones(istop)']; ystop=[ystop;Xi(istop)'];
  icont=find(im==2);
  xcont=[xcont;i*ones(icont)']; ycont=[ycont;Xi(icont)'];
end

plot2d(xcont,ycont,0);                    ← graphique~2.1
plot2d(xstop,ystop,-1);
legends(['continuer';'arreter'],[0,-1],2)
xtitle('commande optimale','temps','X');

```

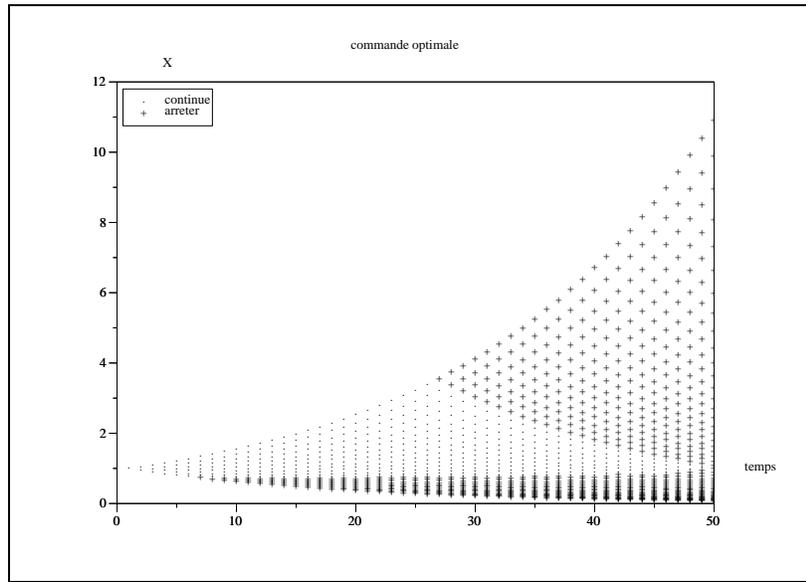


FIG. 2.1 – arret optimal

2.1.2 Un exemple :Le problème de la secrétaire

On se pose le problème suivant : On doit choisir une secrétaire parmi N candidates. Pour ce faire, on reçoit les candidates une à une et à chaque fois on peut décider de retenir la candidate et de s'arrêter ou bien de continuer. Si l'on continue on ne peut pas revenir en arrière, les candidates non retenues ne peuvent plus l'être ultérieurement. Chaque candidate a bien sur une «valeur» que le décideur est capable d'évaluer exactement au moment de l'entretien. Le but du décideur est de maximiser la probabilité d'avoir choisi la meilleure des N candidates au moment ou il décide de s'arrêter.

On peut formaliser le problème de la façon suivante. On dispose de N objets de valeurs v_k ($v_1 < v_2 < \dots < v_N$) inconnues que l'on tire au hasard (de façon équiprobable) les uns après les autres. Quand on tire le k -ième objet on peut soit le choisir et s'arrêter soit continuer. Les objets que l'on va tirer successivement auront pour valeur $v_{\sigma(k)}$ où σ est une permutation de $[1, N]$ et toutes les suites $(v_{\sigma(k)})_{k=1, N}$ sont équiprobables.

Introduisons les notations suivantes :

$$S_k = \begin{cases} 1 & \text{si } W_k = k \text{ avec } W_k = \text{Argmax}_{j \in [1, k]} v_{\sigma(j)} \\ 0 & \text{sinon} \end{cases}$$

La probabilité que sur N tirages la k -ième valeurs obtenues soit la valeur optimale est donnée par :

$$\begin{aligned} \mathbb{P}\{W_N = k\} &= \mathbb{P}\{W_N \in [1, k]\} \mathbb{P}\{W_k = k\} \\ &= (k/N) \mathbb{P}\{S_k = 1\} \\ &= \mathbb{E}[g_k(S_k)] \text{ avec } g_k(1) \equiv k/N \text{ et } g_k(0) \equiv 0. \end{aligned}$$

La quantité que le décideur cherche à maximiser ($\mathbb{P}\{W_N = k\}$) s'exprime donc sous la forme de l'espérance d'une fonction du processus S_k . L'espace d'état de S_k est $\{0, 1\}$ et S_k est

une chaîne de Markov dont les matrices de transitions $M^{(k)}$ vérifient : $M_{i,1}^{(k)} = 1/(k+1)$ et $M_{i,0}^{(k)} = k/(k+1)$ (On peut remarquer qu'elles ne dépendent pas de l'état du processus à l'instant k , les S_k sont en fait indépendants).

Le problème du décideur est donc un problème d'arrêt optimal. Il cherche à résoudre $u_1(1)$ (noter que $S_1 \equiv 1$), $u_n(x)$ étant donné par la formule générale :

$$u_n(x) \equiv \sup_{\tau \text{ } \mathcal{F}_n \text{ t.a., } n \leq \tau \leq N} \mathbb{E}[g^\tau(S^\tau) | S_n = x].$$

En utilisant le Théorème (2.3), on peut affirmer que $u_n(x)$ est solution de l'équation suivante :

$$u_n(x) = \max \left(\frac{1}{n+1} u_{n+1}(1) + \frac{n}{n+1} u_{n+1}(0), \frac{n}{N} \mathbb{1}_{x \neq 0} \right) \quad ; \quad u_N(x) = g_N(x) = x \quad (2.4)$$

La solution $u_n(x)$ étant positive on peut simplifier l'équation précédente sous la forme :

$$u_n(0) = \frac{1}{n+1} u_{n+1}(1) + \frac{n}{n+1} u_{n+1}(0), \quad u_n(1) = \max \left(u_n(0), \frac{n}{N} \right). \quad (2.5)$$

Lemme 2.1.1. *Pour $N > 1$, Il doit exister un $\alpha \in [1, N)$ tel que $\max(u_\alpha(0), \frac{\alpha}{N}) = u_\alpha(0)$.*

Preuve : Si cette propriété n'est pas vérifiée, la deuxième équation (2.5) impose $u_n(1) = n/N > u_n(0)$ pour tout n dans $[1, N]$. On obtient alors avec la première équation (2.5) et $u_N(0) = 0$ la formule explicite suivante pour $u_n(0)$:

$$u_n(0) = \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i}$$

Pour $N > 1$ on déduit de la formule précédente $u_1(0) \geq 1/N$ qui contredit l'inégalité $u_n(0) < n/N$ pour $n \in [1, N)$. \square

Lemme 2.1.2. *Pour $n \in [1, N)$, si $u_n(1) = u_n(0) = \beta \geq n/N$ alors $u_p(1) = u_p(0) = \beta > p/N$ pour tout $p \in [1, n-1]$*

Preuve : La preuve se fait par récurrence. Pour $n = 1$ la propriété est bien évidemment vérifiée. Supposons qu'elle soit vraie pour n . Si $u_{n+1}(1) = u_{n+1}(0) = \beta \geq (n+1)/N$ l'équation (2.5) nous donne $u_n(0) = \beta$ et $u_n(1) = \max(\beta, \frac{n}{N}) = \beta$ et $\beta > n/N$. On conclut avec l'hypothèse de récurrence \square

On appelle α le plus grand α qui vérifie le Lemme 2.1.1 et des deux Lemmes précédents on déduit la forme de $u_n(1)$:

$$\begin{aligned} u_n(1) &= u_n(0) = u_\alpha(1) > n/N \quad \text{pour } n \in [1, \alpha[\\ &u_\alpha(1) = u_\alpha(0) \geq \alpha/N \\ u_n(1) &= n/N > u_n(0) = \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i} \quad \text{pour } n \in]\alpha, N] \end{aligned}$$

Et la stratégie optimale s'en déduit. On a tout d'abord $u_n(0) > 0$ sauf éventuellement en $n = N$ (pour $n < \alpha$ cela vient de l'équation précédente et pour $n \geq \alpha$ cela vient de la formule explicite pour $u_n(0)$) on ne s'arrête donc jamais en 0. La forme de $u_n(1)$ indique qu'il faut s'arrêter après le temps déterministe α dès que l'on atteint l'état 1 (c'est à dire dès qu'une candidate est meilleure que les précédentes).

α est caractérisé par :

$$\alpha = \max \left\{ n \geq 1 : \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i} \geq \frac{n}{N} \right\}$$

et pour N grand on obtient facilement $\lim_{N \rightarrow \infty} \alpha_N/N = e^{-1}$. Après avoir vu environ le tiers des N candidates, il faut choisir la première qui à une valeur supérieures à toutes celles déjà vu. La probabilité d'avoir choisit la meilleure est alors $u_1(1) = u_\alpha(0)$ et cette quantité est de l'ordre de α/N (on sait que $\alpha/N \leq u_\alpha(0)$ et on déduit de la formule explicite de $u_n()$ pour $n \geq \alpha$ que $u_\alpha(0) \leq (\alpha + 1)/N$). Pour N grand, la probabilité d'avoir choisit la meilleure candidate est de l'ordre de $1/e$.

2.2 Programmation dynamique en horizon fini

On s'intéresse à l'optimisation d'un système régi par une chaîne de Markov contrôlée, sur une période de gestion finie à N étapes. On suppose que la matrice de transition et le coût dépendent du temps et on les note respectivement $M^{(n)u}$ et c^{nu} , $n \in \mathbb{N}$.

On définit la fonction valeur :

$$v^n(x) = \min_{\mathbb{U} \in \mathcal{U}^{MD}} \mathbb{E} \left[\sum_{j=n}^{N-1} c^{jU^j}(X^j) + \phi(X^N) | X^n = x \right]. \quad (2.6)$$

Nous allons montrer qu'en fait ce problème admet la même solution que le problème plus général où on autorise $\mathbb{U} \in \mathcal{U}^{HR}$ et que cette solution s'obtient par la Programmation Dynamique.

2.2.1 Calcul de la fonction coût pour une stratégie donnée dans \mathcal{U}^{HR}

Lemme 2.2.1. *Soit f une fonction sur $\mathcal{H}^{n+1} \equiv (\mathcal{E} \times \mathcal{C})^n \times \mathcal{E}$, pour une stratégie donnée $\mathbb{U} \in \mathcal{U}^{HR}$ et en utilisant les équations (2.1) et la notation $h^n = (h^{n-1}, u^{n-1}, x^n)$, on obtient :*

$$\mathbb{E}^{\mathbb{U}} [f(H^{n+1}) | H^n = h^n] = \sum_{u \in \mathcal{C}, y \in \mathcal{E}} q_{U^n(h^n)}(u) M_{x^n, y}^{(n), u} f(h^n, u, y) \quad (2.7)$$

Preuve :

$$\begin{aligned} \mathbb{E}^{\mathbb{U}} [f(H^{n+1}) | H^n = h^n] &= \sum_{u \in \mathcal{C}, y \in \mathcal{E}} f(h^n, u, y) \mathbb{P}^{\mathbb{U}} \{X^{n+1} = y, U^n = u | H^n = h^n\} \\ &= \sum_{u \in \mathcal{C}, y \in \mathcal{E}} f(h^n, u, y) \mathbb{P}^{\mathbb{U}} \{U^n = u | H^n = h^n\} \\ &\quad \times \mathbb{P}^{\mathbb{U}} \{X^{n+1} = y | H^n = h^n, U^n = u\} \\ &= \sum_{u \in \mathcal{C}, y \in \mathcal{E}} q_{U^n(h^n)}(u) M_{x^n, y}^{(n), u} f(h^n, u, y) \end{aligned}$$

□

Théorème 2.2.1. *Pour une stratégie donnée $\mathbb{U} \in \mathcal{U}^{HR}$ la fonctionnelle du processus $(X_n)_{n \in \mathbb{N}}$ donnée par l'expression suivante pour $0 \leq n \leq N$:*

$$v_n^{\mathbb{U}}(h^n) = \mathbb{E}_{h^n}^{\mathbb{U}} \left[\sum_{i=n}^{N-1} c^{iU_i}(X_i) + \phi(X^N) \right] \quad \text{et} \quad v_N^{\mathbb{U}}(h^N) = \phi(x^N) \quad (2.8)$$

se calcule par récurrence arrière de la façon suivante :

$$v_n^{\mathbb{U}}(h^n) = \sum_{u \in \mathcal{C}} q_{U^n(h^n)}(u) \left\{ c^{n,u}(x^n) + \sum_{y \in \mathcal{E}} M_{x^n,y}^{(n)u} v_{n+1}^{\mathbb{U}}(h^n, u, y) \right\} \quad (2.9)$$

avec la notation $h^n = (h^{n-1}, u^{n-1}, x^n)$.

La preuve qui découle de l'équation (2.7) est laissée en exercice au lecteur.

Dans le cas particulier où la stratégie est Markovienne relaxée, l'équation précédente se simplifie, $(X^n)_{n \in \mathbb{N}}$ est une chaîne de Markov de transition (2.2) $M_{x,y}^{(n),U} \equiv \sum_{u \in \mathcal{C}} q_{U^n(x)}(u) M_{x,y}^{(n)u}$ et pour la fonction $c^{n,U}(x) \equiv \sum_{u \in \mathcal{C}} q_{U^n(x)}(u) c^{nu}(x)$ on retrouve l'équation (1.9) du chapitre précédent.

2.2.2 Équation de la programmation dynamique pour $\mathbb{U} \in \mathcal{U}^{HR}$

Nous nous plaçons ici dans le cadre général où les stratégies peuvent appartenir à \mathcal{U}^{HR} et nous allons montrer que v_n^* définie par :

$$v_n^*(h^n) = \min_{\mathbb{U} \in \mathcal{U}^{HR}} \mathbb{E}_{h^n}^{\mathbb{U}} \left[\sum_{i=n}^{N-1} c^{iU_i}(X^i) + \phi(X^N) \right] \quad \text{et} \quad v_N^*(h^N) = \phi(x^N) \quad (2.10)$$

est l'unique solution de l'équation de Programmation Dynamique suivante :

$$\begin{cases} u_n(h^n) = \inf_{u \in \mathcal{C}} \left\{ c^{(n)u}(x^n) + \sum_{y \in \mathcal{E}} M_{x^n,y}^{(n)u} u_{n+1}(h^n, u, y) \right\} \\ u_N(h^N) = \phi(x^N) \end{cases} \quad (2.11)$$

Nous supposons ici que le minimum est bien défini dans l'équation précédente. Nous donnerons dans le théorème qui termine cette section des hypothèses qui assurent l'existence du minimum.

On peut tout d'abord remarquer que dans l'équation précédente u_n ne dépend en fait que de x^n (cela se montre par récurrence et vient du fait que le coût instantané $c^{(n)u}$ ne dépend que de X^n et le coût final ϕ que de X^N).

Lemme 2.2.2. *Soit $(u_n(h_n))_{n \in [0,N]}$ la solution de l'équation (2.11), u_n vérifie :*

$$u_n(h^n) \leq v_n^*(h^n) \equiv \min_{\mathbb{U} \in \mathcal{U}^{HR}} v_n^{\mathbb{U}}(h^n) \quad \text{pour} \quad n \in [0, N] \quad (2.12)$$

Preuve : Soit $\mathbb{U} \in \mathcal{U}^{HR}$ et $k < N$, on a alors en utilisant l'équation (2.7) :

$$\begin{aligned}
\mathbb{E}^{\mathbb{U}} \left[c^{kU^k}(X^k) + u_{k+1}(H^{k+1}) | H^k = h^k \right] \\
&= \sum_{u \in \mathcal{C}} q_{U^k(h^k)}(u) \left\{ c^{k,u}(x^k) + \sum_{y \in \mathcal{E}} M_{x^k,y}^{(k)u} u_{k+1}(h^k, u, y) \right\} \\
&\geq \min_{u \in \mathcal{C}} \left\{ c^{k,u}(x^k) + \sum_{y \in \mathcal{E}} M_{x^k,y}^{(k)u} u_{k+1}(h^k, u, y) \right\} \\
&\geq u_k(h^k)
\end{aligned} \tag{2.13}$$

L'inégalité (2.13) se déduit de la propriété suivante $\sum_{u \in \mathcal{C}} q(u)f(u) \geq \min_{u \in \mathcal{C}} f(u)$. On déduit du résultat précédent, en reprenant une espérance conditionnelle par rapport à H_n que :

$$\mathbb{E}_{h_n}^{\mathbb{U}} \left[c^{kU^k}(X^k) \right] + \mathbb{E}_{h_n}^{\mathbb{U}} \left[u_{k+1}(H^{k+1}) \right] \geq \mathbb{E}_{h_n}^{\mathbb{U}} \left[u_k(H^k) \right] \quad \text{pour } n \leq k \leq N-1 \tag{2.14}$$

et en sommant toutes les équations précédentes pour $k = n, \dots, N-1$ on obtient :

$$\mathbb{E}_{h_n}^{\mathbb{U}} \left[\sum_{j=n}^{N-1} c^{jU^j}(X^j) + \phi(X^N) \right] \geq \mathbb{E}_{h_n}^{\mathbb{U}} \left[u_n(H_n) \right] = u_n(h_n).$$

D'où le résultat. □

Nous montrons maintenant que l'on peut toujours construire une stratégie ϵ -optimale (même quand le minimum n'est pas atteint dans l'équation (2.11)) et que l'on a $u_n(x^n) = u_n(h^n) = v_n^*(h^n)$

Lemme 2.2.3. $(u_n(h_n))_{n \in [0, N]}$ solution de l'équation (2.11) vérifie

$$u_n(h^n) = \min_{\mathbb{U} \in \mathcal{U}^{HR}} v_n^{\mathbb{U}}(h_n)$$

pour $n \in [0, N]$. Pour $\epsilon > 0$ fixé on peut construire une stratégie Markovienne déterministe $\mathbb{U}_\epsilon \in \mathcal{U}^{MD}$ telle que $u^n(x) \leq v_n^{\mathbb{U}_\epsilon}(x) \leq u^n(x) + \epsilon$.

Preuve : Soit ϵ fixé, en utilisant la définition de $u_n(x)$ (en utilisant le fait que u_n ne dépend en fait que de l'état à l'instant n) on peut construire une stratégie \mathbb{U}_ϵ telle que :

$$u_n(x) + \epsilon/N \geq c^{(n)U_\epsilon^n}(x) + \sum_{y \in \mathcal{E}} M_{x,y}^{(n)U_\epsilon^n} u_{n+1}(y) \tag{2.15}$$

Il est alors facile de montrer par récurrence que $v_n^{\mathbb{U}_\epsilon}(x) \leq u_n(x) + (N-n)\epsilon/N$, ce qui permet de conclure. On a en effet $v_n^*(h_n) \leq v_n^{\mathbb{U}_\epsilon}(x) \leq u_n(x) + \epsilon$, ce qui combiné avec le lemme précédent assure $u_n(x) = v_n^*(x)$. La stratégie \mathbb{U}_ϵ construite dans la démonstration est clairement dans \mathcal{U}^{MD} et elle permet d'atteindre le minimum de la fonction coût à ϵ près. □

Si maintenant le minimum est atteint dans l'équation (2.11) cela permet de construire une stratégie Markovienne déterministe $U^*(x)$:

$$U^*(x) \in \underset{u \in \mathcal{C}}{\text{Argmin}} \left\{ c^{(n)u}(x) + \sum_{y \in \mathcal{E}} M_{x,y}^{(n)u} u_{n+1}(y) \right\}$$

qui en utilisant la même démonstration que plus haut vérifie $v_n^{U^*} = v_n^*$.

Nous récapitulons les résultats précédents dans le théorème qui suit.

Théorème 2.2.2. *Supposons \mathcal{E} fini, et l'une des trois conditions suivantes :*

1. \mathcal{C} est de dimension finie.
2. \mathcal{C} est compact et $u \rightarrow (M^{(n)u}, c^{(n)u})$ est continue pour tout $n \in \mathbb{N}$.
3. \mathcal{C} est compact et, pour (n, x, y) fixés, $u \rightarrow (M_{x,y}^{(n)u})$ est semi continue inférieurement et $c^{(n)u}(x)$ est semi continue inférieurement (comme \mathcal{E} est fini on déduit de l'hypothèse qui précède que c est minorée)

Alors la fonction valeur v^n (2.6) est solution de l'équation de la Programmation Dynamique (ou équation de Bellman) :

$$\begin{cases} v^n(x) &= \min_{u \in \mathcal{C}} \{ (M^{(n)u} v^{n+1})(x) + c^{(n)u}(x) \} \quad \forall x \in \mathcal{E}, n = 0, \dots, N-1; \\ v^N(x) &= \phi(x), \quad \forall x \in \mathcal{E}. \end{cases} \quad (2.16)$$

et une commande optimale feedback est donnée par l'équation suivante :

$$U^*(x) \in \underset{u \in \mathcal{C}}{\text{Argmin}} \left\{ c^{(n)u}(x) + \sum_{y \in \mathcal{E}} M_{x,y}^{(n)u} u_{n+1}(y) \right\}$$

Ce théorème permet de calculer par récurrence arrière v^n . On part de $v^N = \phi$, et on calcule récursivement v^{N-1}, \dots, v^0 .

Il faut montrer que les hypothèses du théorème assurent l'existence d'un minimum à chaque étape de l'algorithme récursif et la preuve du théorème découle alors des paragraphes précédents ou nous avons montré que minimiser sur \mathcal{U}^{HR} et \mathcal{U}^{MD} donnait en fait le même résultat.

Preuve : Nous limitons la preuve pour les hypothèses 3. c est minorée ainsi que ϕ . Montrons par récurrence que $u_n(x)$ donné par (2.11) est minorée et que le minimum est atteint à chaque étape de l'algorithme récursif. Pour $k = N$, $u_N(x) = \phi(x)$ et l'hypothèse est vérifiée. Supposons $u_{k+1}(x) \geq K$, alors $\sum_{y \in \mathcal{E}} M_{x,y}^{(k)u} (u_{k+1}(x) - K)$ est s.c.i en u [24, p-603]. Comme $\sum_{y \in \mathcal{E}} M_{x,y}^{(k)u} K = K$, on conclut que $c^{(k)u}(x) + \sum_{y \in \mathcal{E}} M_{x,y}^{(k)u} u_{k+1}(x)$ est aussi s.c.i et le minimum est atteint ([24, p-603]). $u_k(x)$ est donc aussi minorée. \square

On peut sans difficulté généraliser ce qui précède si des contraintes sur le contrôle imposent $u^n \in \mathcal{C}_x^n$.

2.2.3 Liens avec les problèmes de temps d'arrêt

Nous montrons ici pour conclure que le problème du temps d'arrêt optimal (d'horizon N) peut se plonger dans un problème de programmation dynamique en horizon fini. Soit (X^n) une chaîne de Markov de matrice de transition $M^{(n)}$ sur un espace d'état fini \mathcal{E} . On augmente l'espace d'état d'un point Δ ($\bar{\mathcal{E}} = \{\mathcal{E}, \Delta\}$). Et on considère une chaîne contrôlée sur l'espace d'état augmenté avec deux contrôles possibles $\mathcal{C} = \{C, S\}$ (C pour continuer et S pour stopper). On se donne la matrice de transition de la chaîne contrôlée $\bar{M}^{(n)u}$ par :

$$\overline{M}^{(n)C} = \begin{pmatrix} M^{(n)} & \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ 0 & 1 \end{pmatrix} \quad \overline{M}^{(n)S} = \begin{pmatrix} 0 & \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ 0 & 1 \end{pmatrix}$$

En d'autres termes, tant que l'on utilise la commande C , la chaîne se comporte sur \mathcal{E} comme la chaîne X^n et quand on choisit la commande S la chaîne saute en Δ et y reste. On cherche à montrer que la fonctionnelle :

$$v_n(x) = \sup_{\tau \text{ } \mathcal{F}_n \text{ t.a.}, n \leq \tau \leq N} \mathbb{E} \left[\sum_{k=n}^{\tau-1} c^{(k)}(X^k) + \psi^\tau(X^\tau) | X^n = x \right]$$

coincide sur \mathcal{E} avec la solution du problème de commande en horizon fini :

$$v^n(x) = \max_{\mathbb{U} \in \mathcal{U}^{MD}} \mathbb{E} \left[\sum_{j=n}^N \bar{c}^{(j)U^j}(X^j) | X^n = x \right].$$

avec les contraintes suivantes $U^j \in \mathcal{C}_x^j \equiv \{C, S\}$ pour $j \neq N$ et $U^N \in \mathcal{C}_x^N \equiv \{S\}$. D'autre part, on choisit comme fonction coût c :

$$\bar{c}^{(j)p}(x) \equiv \begin{cases} \psi^j(x) & \text{pour } p = S \text{ et } x \in \mathcal{E} \\ c^{(j)}(x) & \text{pour } p = C \text{ et } x \in \mathcal{E} \\ 0 & \text{pour } x = \Delta \end{cases}$$

L'équation de la programmation dynamique s'écrit :

$$\begin{cases} v^n(x) = \max_{u \in \mathcal{C}_x^n} \{ (\overline{M}^{(n)u} v^{n+1})(x) + \bar{c}^{(n)u}(x) \} \quad \forall x \in \overline{\mathcal{E}}, n \in [0, N-1], \\ v^N(x) = \bar{c}^{(N)S}(x) = \psi^N(x) \quad \forall x \in \mathcal{E} \quad \text{et} \quad v^N(\Delta) = 0. \end{cases}$$

Tout d'abord pour $x = \Delta$ l'équation précédente s'écrit $v^n(\Delta) = \max_{u \in \mathcal{C}_x^n} v^{n+1}(\Delta) = v^{n+1}(\Delta)$. Compte tenu de la condition finale on a donc $v^n(\Delta) = 0$ pour tout n . Pour $x \in \mathcal{E}$, on obtient en considérant les deux commandes possibles C et S l'équation d'optimalité :

$$\begin{aligned} v^n(x) &= \max \left(M^{(n)} v^{n+1}(x) + c^{(n)}(x), v^{n+1}(\Delta) + \psi^n(x) \right) \\ &= \max \left(M^{(n)} v^{n+1}(x) + c^{(n)}(x), \psi^n(x) \right), \end{aligned}$$

On retrouve bien l'équation (2.3) et la formule pour le temps d'arrêt optimal se déduit aussi de la même façon.

Si on se donne une stratégie \mathbb{U} , on peut lui associer un temps d'arrêt $\tau_n = \inf\{k \geq n : U^k(x) = S\}$, (qui est $\leq N$) et en raisonnant comme au dessus on obtient une nouvelle démonstration de la Proposition 1.2.2 : la fonctionnelle

$$v_n(x) = \mathbb{E} \left[\sum_{k=n}^{\tau_n-1} c^{(j)}(X^j) + \psi^{\tau_n}(X^{\tau_n}) | X^n = x \right]$$

est solution de l'équation de Kolmogorov rétrograde :

$$v_n(x) = \mathcal{L}_{\bar{U}}^n[c, \psi]v_{n+1}$$

où :

$$\mathcal{L}_{\bar{U}}^j[c, \Psi]w \equiv \begin{cases} M^{(j)}w(x) + c^{(j)}(x) & \text{si } U^j(x) = C \\ \Psi^j(x) & \text{si } U^j(x) = S \end{cases}$$

Nous terminons ce paragraphe en montrant que l'on peut comparer les solutions de deux inéquations variationnelles, ce résultat nous sera utile au Chapitre 4 lors de l'étude de la stabilité des méthodes de discrétisation par différences finies.

Proposition 2.2.1. *Soient v et w les solutions des équations*

$$v_n(x) = \max \left(M^{(n)}v_{n+1}(x) + c_v^{(n)}(x), \Psi^n(x) \right) \quad (2.17)$$

$$w_n(x) = \max \left(M^{(n)}w_{n+1}(x) + c_v^{(n)}(x) + e^{(n)}x, \Psi^n(x) \right) \quad (2.18)$$

$$v_N(x) = w_N(x) = \Psi^N(x) \quad (2.19)$$

on a alors la majoration suivante :

$$\sup_{j \in [0, N]} \|v_j - w_j\|_{\infty} \leq N \sup_{j \in [0, N-1]} \|e^{(j)}\|_{\infty}$$

Preuve : Soit \bar{U}_v la stratégie optimale qui réalise le maximum dans l'équation (2.17) on a alors par définition de \bar{U}_v , $v_n = \mathcal{L}_{\bar{U}_v}^n[c_v, \Psi]v_{n+1}$ mais, pour la même stratégie, on a aussi $w_n(x) \geq \mathcal{L}_{\bar{U}_v}^n[c_v + e, \Psi]w_{n+1}$. D'où :

$$\begin{aligned} v_n - w_n &\leq \mathcal{L}_{\bar{U}_v}^n[c_v, \Psi]v_{n+1} - \mathcal{L}_{\bar{U}_v}^n[c_v + e, \Psi]w_{n+1} \\ &\leq \mathcal{L}_{\bar{U}_v}^n[-e, 0](v_{n+1} - w_{n+1}) \\ &= \mathcal{L}_{\bar{U}_v}^n[-e - \Phi, 0](v_{n+1} - w_{n+1}) \end{aligned} \quad (2.20)$$

où Φ est une fonction positive. En utilisant maintenant les résultats du paragraphe précédent, $v_n - w_n$ s'interprète comme la fonctionnelle :

$$v_n(x) - w_n(x) = \mathbb{E} \left[\sum_{k=n}^{\tau_n-1} (-e^{(j)}(X^j) - \Phi^{(j)}(X^j)) | X^n = x \right]$$

avec $\tau_n = \inf\{k \geq n : \bar{U}^k(x) = S\}$. τ_n étant majoré par N et la fonction Φ étant positive, on obtient la majoration $v_n(x) - w_n(x) \leq 0$. Inverser les rôles de v et w permet d'obtenir l'équation précédente où e est remplacé par $-e$, on obtient alors la majoration

$$w_n(x) - v_n(x) \leq N \sup_{j \in [0, N-1]} \|e^{(j)}\|_{\infty}.$$

On a également $v_N(x) - w_N(x) = \Psi^N(x) - \Psi^N(x) = 0$, ce qui termine la preuve. \square

2.3 Programmation dynamique en horizon infini

On considère à présent le problème de la commande optimale d'une chaîne de Markov en horizon infini, pour un coût actualisé ($\lambda > 0$). L'actualisation du coût permet de gérer l'importance relative des événements à court et long terme tout en conduisant à la résolution d'une équation de la programmation dynamique stationnaire.

On définit la fonction valeur :

$$v^*(x) = \inf_{\mathbb{U} \in \mathcal{U}^{MR}} v^{\mathbb{U}}(x) \quad \text{avec} \quad v^{\mathbb{U}}(x) = \mathbb{E}_{x,0} \left[\sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} c^{U^n}(X^n) \right] \quad (2.21)$$

Remarque : On peut remarquer que les modèles actualisés en horizon infini peuvent aussi provenir de problème en horizon fini pour lesquels l'horizon est aléatoire et indépendant des actions du contrôleur. Par exemple, considérons un critère qui tient compte de la valeur moyenne de l'horizon inconnu :

$$w(x) = \mathbb{E}_{x,0} \left[\mathbb{E}_{\nu} \left[\sum_{n=0}^{\nu} c^{U^n}(X^n) \right] \right].$$

Si ν suit une loi géométrique sur \mathbb{N} de paramètre $p = \frac{\lambda}{1+\lambda}$ ($\mathbb{P}\{\nu = n\} = p(1-p)^n$), ce critère se réécrit par un calcul facile laissé en exercice, sous la forme :

$$w(x) = \mathbb{E}_{x,0} \left[\sum_{n=0}^{\infty} \frac{1}{(1+\lambda)^n} c^{U^n}(X^n) \right].$$

□

2.3.1 Évaluation pour une stratégie stationnaire de la fonction coût

On considère ici des stratégies stationnaires Markoviennes et on confond donc dans la notation \mathbb{U} et U . Pour une stratégie fixée U , le calcul de $v^U(x)$ se ramène (voir les équations 2.2), pour la chaîne contrôlée, au calcul de fonctionnelles vues à la section (1.2.3). $v^U(x)$ est l'unique solution de l'équation fonctionnelle (1.16).

On peut remarquer que l'équation (1.16) s'écrit aussi en notation vectorielle :

$$v^U = \frac{1}{1+\lambda} c^U + \frac{1}{1+\lambda} M^U v^U \quad (2.22)$$

v^U apparaît donc comme l'unique point fixe de l'application $\mathcal{L}_U : V \mapsto V$ telle que $\mathcal{L}_U(v) = \frac{1}{1+\lambda} c^U + \frac{1}{1+\lambda} M^U v$. Ici V désigne les fonctions bornées sur \mathcal{E} muni de la norme L^∞ et \mathcal{L}_U est bien à valeur dans V pour $c^u(x)$ bornée. Nous supposons dans la suite de ce chapitre $|c^u(x)| \leq C < \infty$ pour tout $x \in \mathcal{E}$ et tout $u \in \mathcal{C}_x$.

2.3.2 Équation d'optimalité

On cherche dans cette section à montrer que la fonction valeur (2.21) est l'unique point fixe de l'application : $\mathcal{L} : V \mapsto V$

$$\mathcal{L}v = \inf_{U \in \mathcal{D}^{MD}} \mathcal{L}_U v$$

ou bien (quand le minimum est atteint) de l'application $L : V \mapsto V$

$$Lv = \min_{U \in \mathcal{D}^{MD}} \mathcal{L}_U v$$

Exercice 2.3.1. *Montrer que $\inf_{U \in \mathcal{D}^{MD}} \mathcal{L}_U = \inf_{U \in \mathcal{D}^{MR}} \mathcal{L}_U$. En d'autres termes on ne gagne rien à utiliser des stratégies Markoviennes relaxées.*

Le théorème qui suit permet de minorer ou majorer la fonction valeur si l'on dispose de sur ou sous-solutions de l'équation d'optimalité $v = \mathcal{L}v$. Il montre aussi que si l'équation d'optimalité admet une solution, celle-ci ne peut être que la fonction valeur v^* (2.21). nous commençons par un lemme.

Lemme 2.3.1. *Pour $(v_1, v_2) \in V^2$ on a la propriété suivante (principe du maximum) :*

$$\mathcal{L}v_1 - \mathcal{L}v_2 \leq v_1 - v_2 \implies v_1 \geq v_2$$

La même propriété est vérifiée pour \mathcal{L}_W et L .

Preuve : On commence par l'opérateur \mathcal{L}_W . De $\mathcal{L}_W v_1 - \mathcal{L}_W v_2 \leq v_1 - v_2$ on déduit $(I - \frac{1}{1+\lambda} M^W)(v_1 - v_2) \geq 0$. Soit aussi, en utilisant le générateur A^W de la chaîne contrôlée, $(A^W - \lambda I)(v_1 - v_2) \leq 0$. En appliquant le corollaire 1.3.1 on obtient $v_1 \geq v_2$.

Supposons maintenant $\mathcal{L}v_1 - \mathcal{L}v_2 \leq v_1 - v_2$. Soit $\epsilon > 0$ fixé, par définition de \mathcal{L} on peut trouver U_ϵ tel que $\mathcal{L}_{U_\epsilon} v_1 \leq \mathcal{L}v_1 + \epsilon \frac{\lambda}{1+\lambda} e$. Pour ce même U_ϵ on a aussi $\mathcal{L}(v_2 - \epsilon e) \leq \mathcal{L}_{U_\epsilon}(v_2 - \epsilon e)$. On a donc :

$$\begin{aligned} \mathcal{L}_{U_\epsilon} v_1 - \mathcal{L}_{U_\epsilon}(v_2 - \epsilon e) &\leq \mathcal{L}v_1 - \mathcal{L}(v_2 - \epsilon e) + \epsilon \frac{\lambda}{1+\lambda} e = \mathcal{L}v_1 - \mathcal{L}(v_2) + \epsilon e \\ &\leq v_1 - v_2 + \epsilon e = v_1 - (v_2 - \epsilon e) \end{aligned} \quad (2.23)$$

Le résultat montré pour l'opérateur \mathcal{L}_W nous assure alors $v_1 \geq v_2 + \epsilon e$ pour tout $\epsilon > 0$ et donc $v_1 \geq v_2$. \square

Le théorème suivant est la clef de l'existence d'une solution des problèmes de points fixe $v = \mathcal{L}v$ (resp. $v = \mathcal{L}_W v$).

Théorème 2.3.1. *L et \mathcal{L} sont des contractions strictes sur V .*

Preuve : Nous montrons ici que \mathcal{L} est une contraction sur V . Soient u, v dans V et $x \in \mathcal{E}$, supposons que $\mathcal{L}v(x) \leq \mathcal{L}u(x)$. Par définition de $\mathcal{L}v$, pour $\epsilon > 0$ fixé il existe W_ϵ tel que $\mathcal{L}_{W_\epsilon} v(x) \leq \mathcal{L}v(x) + \epsilon$. On a alors aussi de façon évidente $\mathcal{L}u(x) \leq \mathcal{L}_{W_\epsilon} u(x)$ d'où $0 \leq \mathcal{L}u(x) - \mathcal{L}v(x) \leq \mathcal{L}_{W_\epsilon} u(x) - \mathcal{L}_{W_\epsilon} v(x) + \epsilon$. En utilisant le fait que M^{W_ϵ} est une matrice stochastique et en explicitant \mathcal{L}_{W_ϵ} on obtient $0 \leq \mathcal{L}u(x) - \mathcal{L}v(x) \leq \frac{1}{1+\lambda} \|u - v\|_\infty$. Le cas $\mathcal{L}u(x) \leq \mathcal{L}v(x)$ se déduit en intervertissant les rôles de u et v et conclut la preuve, le coefficient de contraction étant $\alpha = \frac{1}{1+\lambda} < 1$. \square

Théorème 2.3.2. *Soit $v \in V$ vérifiant $v \geq \mathcal{L}v$ (resp. $v \leq \mathcal{L}v$) alors $v \geq \hat{v}$ (resp. $v \leq \hat{v}$) où \hat{v} est un point fixe de \mathcal{L} . Soit $v \in V$ vérifiant $v = \mathcal{L}v$ alors $v = \hat{v}$. Les mêmes résultats sont obtenus avec l'opérateur L et avec l'opérateur \mathcal{L}_W (en remplaçant dans ce dernier cas \hat{v} par v^W le point fixe de \mathcal{L}_W).*

Preuve : Nous donnons ici la preuve pour l'opérateur \mathcal{L} qui est immédiate en utilisant le lemme qui précède. Supposons $v \geq \mathcal{L}v$ alors compte tenu de $\mathcal{L}\hat{v} = \hat{v}$ on a $\mathcal{L}v - \mathcal{L}\hat{v} \leq v - \hat{v}$ et donc $v \geq \hat{v}$. Le résultat symétrique se démontre de la même façon. \square

Corollaire 2.3.1. Soient $v \in V$, $W \in \mathcal{D}^{MD}$ et $\epsilon \geq 0$ vérifiant $\mathcal{L}_W v \leq v + \epsilon$ alors $v^W \leq v + \frac{1+\lambda}{\lambda}\epsilon$ (on rappelle que v^W est l'unique solution de $v^W = \mathcal{L}_W v^W$).

Preuve : Soit $v_1 = v + \epsilon(1 + \lambda)/\lambda\epsilon$, on a :

$$\mathcal{L}_W(v_1) \leq \mathcal{L}_W(v) + \frac{\epsilon}{\lambda}e \leq v_1$$

et il suffit d'appliquer le théorème précédent. \square

Exercice 2.3.2. Montrer que les résultats du théorème 2.5 sont aussi vrais si on remplace \mathcal{L} par l'un ou l'autre des deux opérateurs suivants $\mathcal{I}_m = \inf_U \mathcal{L}_U^m v$ ou $\mathcal{J}_m = \mathcal{L}^m$ (pour tout m dans \mathbb{N}^+). \mathcal{L}^m désignant l'opérateur \mathcal{L} itéré m fois.

Théorème 2.3.3. Soit \hat{v} l'unique point fixe de l'application \mathcal{L} et v^* la fonction valeur définie en (2.21), $v^* = \hat{v}$.

Preuve : En utilisant la définition (2.21 et les résultats du chapitre précédent v^* vérifie $v^* = \inf_{W \in \mathcal{D}^{MD}} v^W$ où v^W est l'unique point fixe de \mathcal{L}_W . Soit donc W fixé, de $\text{cal}L v^W \leq \text{cal}L_W$ et de $\text{cal}L_W v^W = v^W$ on déduit $\text{cal}L v^W \leq v^W$ ce qui en utilisant le théorème 2.5 assure $v^W \geq \hat{v}$. La propriété étant vraie pour tout W , on obtient $v^* \geq \hat{v}$.

Réciproquement, soit ϵ fixé et U_ϵ vérifiant :

$$\mathcal{L}_{U_\epsilon} \hat{v} \leq \mathcal{L}\hat{v} + \epsilon$$

. \hat{v} étant le point fixe de \mathcal{L} on obtient $\mathcal{L}_{U_\epsilon} \hat{v} \leq \hat{v} + \epsilon$ et en utilisant le Corollaire 2.3.1 on obtient :

$$v^* \leq v^{U_\epsilon} \leq \hat{v} + \frac{1+\lambda}{\lambda}\epsilon e$$

et il sort $v^* \leq \hat{v}$. \square

2.3.3 Politique optimale

Théorème 2.3.4. Nous supposons ici que pour tout $v \in V$ le minimum est atteint dans la définition de $\mathcal{L}v$ (voir par exemple les hypothèses du théorème 2.2.2) alors il existe une politique optimale stationnaire Markovienne déterministe U^* . Cette politique s'explicite sous la forme :

$$U^*(x) \in \underset{u \in \mathcal{C}_x}{\text{Argmin}} \left\{ c^u(x) + \sum_{y \in \mathcal{E}} M_{x,y}^u v^*(y) \right\} \quad (2.24)$$

Preuve : Comme v^* est dans V , l'hypothèse faite sur \mathcal{L} assure l'existence de $U^* \in \text{Argmin}_U \mathcal{L}_U v^*$ et l'équation d'optimalité peut s'écrire $v^* = \mathcal{L}_{U^*} v^*$. Nous avons vu par ailleurs (2.22) que la fonction valeur v^{U^*} associée à la stratégie stationnaire construite avec U^* est l'unique solution de l'équation de point fixe $v = \mathcal{L}_{U^*} v$. On doit donc avoir $v^* = v^{U^*}$ et ainsi U^* définit bien une stratégie Markovienne déterministe optimale. \square

On peut aussi remarquer que cette stratégie optimale est la même pour tous les points de départ de la chaîne.

Corollaire 2.3.2. *S'il existe une politique optimale dans \mathcal{U}^{HR} alors il existe une politique optimale stationnaire markovienne déterministe.*

Preuve : Soit $\mathbb{U}^* \in \mathcal{U}^{HR}$ une politique optimale (i.e $v^*(x) = v^{\mathbb{U}^*}(x)$), elle se décompose en $\mathbb{U}^* = (U^1, W)$ ou U^1 , le contrôle à la première étape, est dans \mathcal{D}^{MD} . on obtient alors en utilisant la définition de $v^{\mathbb{U}^*}$:

$$v^{U^1} = \frac{1}{1+\lambda} c^{U^1} + \frac{1}{1+\lambda} M^{U^1} v^W \geq \frac{1}{1+\lambda} c^{U^1} + \frac{1}{1+\lambda} M^{U^1} v^{U^1} \geq \mathcal{L} v^{U^1} = v^{U^1}$$

et donc $v^{U^1} = \mathcal{L}_{U^1} v^{U^1}$ et on conclut qu'une politique stationnaire Markovienne déterministe basée sur U^1 est optimale. \square

Remarque : On peut toujours obtenir une politique ϵ optimale. Pour $\epsilon > 0$ fixé, il existe W_ϵ vérifiant $\mathcal{L}_{W_\epsilon} v^* \leq \mathcal{L} v^* + \frac{\lambda}{1+\lambda} \epsilon e$, comme $\mathcal{L} v^* = v^*$, on conclut en utilisant le corollaire (2.3.1) que $v^{W_\epsilon} \leq v^* + \epsilon$ et on obtient bien ainsi une politique ϵ -optimale. \square

2.3.4 Dépendance de la fonction valeur vis à vis de la fonction coût

Nous regardons ici une propriété de continuité de la fonction qui à c associe la fonction valeur v . Nous changeons ici légèrement les notations en faisant apparaître dans la définition de \mathcal{L} la dépendance en c . Nous noterons donc $\mathcal{L}_U[c](v) \equiv \frac{1}{1+\lambda} c^U + \frac{1}{1+\lambda} M^U v^U$ et $\mathcal{L}[c](v) = \inf_{U \in \mathcal{D}^{MD}} \mathcal{L}_U[c]v$. D'autre part dans ce paragraphe M^U désigne une matrice contractante (pas forcément Markovienne). λ peut dépendre de x dans la définition de \mathcal{L} mais on suppose que $0 < \inf_x \lambda(x) = \underline{\lambda}$.

Proposition 2.3.1. *Soient v_1 et v_2 les points fixes de $\mathcal{L}[c_1]$ et $\mathcal{L}[c_2]$ on a alors :*

$$\|v_1 - v_2\|_\infty \leq \frac{1}{\underline{\lambda}} \sup_{u \in U} \|c_1^u - c_2^u\|_\infty$$

Preuve : Soit $\epsilon > 0$ fixé, On peut trouver une stratégie U_1^ϵ telle que $\mathcal{L}_{U_1^\epsilon}[c_1]v_1 \leq \mathcal{L}[c_1](v_1) + \epsilon$. On a alors aussi par définition de $\mathcal{L}[c_2]$, $\mathcal{L}[c_2](v_2) \leq \mathcal{L}_{U_1^\epsilon}[c_2]v_2$. En utilisant les définition de v_1 et v_2 on a donc :

$$\begin{aligned} v_2 - v_1 &\leq \mathcal{L}_{U_1^\epsilon}[c_2]v_2 - \mathcal{L}_{U_1^\epsilon}[c_1]v_1 + \epsilon \\ &\leq \mathcal{L}_{U_1^\epsilon}[c_2 - c_1](v_2 - v_1) + \epsilon \\ &\leq \frac{1}{1+\underline{\lambda}} (\|v_2 - v_1\|_\infty + \sup_{u \in U} \|c_1^u - c_2^u\|_\infty) + \epsilon \end{aligned}$$

l'inégalité est vérifiée pour tout ϵ . En intervertissant les rôles de v_2 et v_1 on obtient l'inégalité inverse et finalement la majoration cherchée. \square

2.3.5 Itération sur les valeurs

Nous avons vu que v^* est l'unique solution dans V de l'équation de point fixe $v = \mathcal{L}v$ (resp. $v = Lv$), l'existence de v^* étant basée sur le caractère contractant de l'opérateur \mathcal{L} (resp. L). v^* est obtenu comme limite des suites $v_{n+1} = \mathcal{L}v_n$ avec $v_0 \in V$. Cette algorithmme itératif est appelée ici méthode d'itérations sur les valeurs :

$$v^{n+1}(x) = \frac{1}{1 + \lambda} \min_{u \in \mathcal{C}} \{ (M^u v^n)(x) + c^u(x) \} \quad (2.25)$$

(de valeur initiale $v^0 = 0$ par exemple).

Cet algorithmme est peu efficace lorsque λ est proche de 0. On peut alors utiliser l'algorithmme d'Howard (paragraphe 2.3.6)).

Pour être complet il faut rajouter un critère d'arrêt dans l'algorithmme d'itérations sur les valeurs et indiquer comment on construit une stratégie optimale. Nous faisons cela sous forme d'exercice.

Exercice 2.3.3. Soit $v_n = \mathcal{L}^n v_0$ la suite générée par l'algorithmme d'itération sur les valeurs. On arrête les itérations à l'étape $N + 1$ avec :

$$N \equiv \inf \left\{ n, \|v_{n+1} - v_n\|_\infty \leq \frac{\lambda}{2} \epsilon \right\}$$

et on construit la loi de commande U_ϵ :

$$U_\epsilon(x) \in \underset{U}{\text{Argmin}} \mathcal{L}_U v_{N+1} = \underset{u \in \mathcal{C}}{\text{Argmin}} \frac{1}{1 + \lambda} \{ c^u(x) + (M^u v^{N+1})(x) \}$$

On cherche à montrer que U_ϵ définit une stratégie ϵ -optimale.

1. Montrer que $\|v_{N+1} - v^*\|_\infty \leq \frac{1}{\lambda} \|v_{N+1} - v_N\|_\infty$
2. Montrer que $\|v^{U_\epsilon} - v^*\|_\infty \leq \frac{2}{\lambda} \|v_{N+1} - v_N\|_\infty$ (On pourra utiliser $\|v^{U_\epsilon} - v^*\|_\infty \leq \|v^{U_\epsilon} - v_{N+1}\|_\infty + \|v_{N+1} - v^*\|_\infty$ et majorer séparément chacun des termes).
3. Montrer que U_ϵ est une politique ϵ -optimale

Exercice 2.3.4. Supposons que l'on ait la propriété suivante $v_{p+1} \leq v_p$ où v_p est la suite générée par l'algorithmme d'itérations sur les valeurs, alors $v_n \downarrow v^*$ pour $n \geq p$.

La vitesse de convergence de l'algorithmme d'itération sur les valeurs est lié au coefficient de contraction, On peut chercher à améliorer ce coefficient en procédant comme suit.

Pour une loi de commande fixée U , par exemple si \mathcal{C}_x est réduit à un point, l'algorithmme d'itération sur les valeurs consiste comme on l'a vu en une méthode itérative $v_{n+1} = \mathcal{L}_U v_n$ pour résoudre le système linéaire :

$$\Gamma^U v = \frac{1}{1 + \lambda} c^U \quad \text{avec} \quad \Gamma^U = \left(I - \frac{1}{1 + \lambda} M^U \right) \quad (2.26)$$

Supposons que l'on décompose Γ^U en $\Gamma^U = Q_U - R_U$ avec Q_U inversible et $Q_U^{-1} \geq 0$ et $R_U \geq 0$ (ici une matrice est ≥ 0 si tous ses coefficients sont positifs ou nuls), on dira alors que (Q_U, R_U) définissent une décomposition admissible. L'équation (2.26) peut s'écrire $v = (Q_U)^{-1} \left\{ \frac{1}{1+\lambda} c^U + R_U v \right\}$ ce qui conduit à une nouvelle méthode itérative possible : $v_{n+1} = (Q_U)^{-1} \left\{ \frac{1}{1+\lambda} c^U + R_U v_n \right\}$. On peut donc espérer généraliser l'algorithme d'itération sur les valeurs en considérant :

$$v_{n+1} = \mathcal{H}_{R,Q}(v_n) \equiv \min_{u \in \mathcal{C}} \left\{ Q_u^{-1} R_u v_n + Q_u^{-1} \frac{1}{1+\lambda} c^u \right\} \quad (2.27)$$

On peut évidemment remarquer que le choix $Q_u = I$ et $R_u = \frac{1}{1+\lambda} M^u$ est un choix possible de décomposition admissible et qu'on retrouve pour ce choix (2.25).

Dan l'exercice qui suit on propose une condition suffisante sur la décomposition admissible qui assure la convergence de l'algorithme proposé en (2.27).

Exercice 2.3.5. *Pour une décomposition admissible (Q, R) vérifiant*

$$\sup_u \|Q_u^{-1} R_u\|_\infty < 1$$

la suite v_n obtenue avec l'équation (2.27) converge vers v^* .

1. Montrer que l'opérateur $\mathcal{H}_{R,Q}$ est contractant sur V et donc que la suite v_n converge vers une limite que l'on notera w^* .
2. Montrer que la matrice $Q_u^{-1} R_u - I$ est a diagonale fortement dominante (1.3.4) et vérifie donc le Principe du maximum discret strictement positif. Utiliser ce résultat pour montrer que $w^* = v^*$.

Exercice 2.3.6. *Soit $Q_u = I - S_u/(1+\lambda)$ avec S_u une matrice sous stochastique et α_u un vecteur donné. v et w étant donnés par les formules :*

$$\begin{aligned} w &= \min_{u \in \mathcal{C}} [Q_u^{-1} \alpha_u] \\ v &= \min_{u \in \mathcal{C}} \left[\frac{S_u}{1+\lambda} v + \alpha_u \right]. \end{aligned}$$

Montrer que $v = w$. En conclure que la suite w_n donnée par :

$$w_{n+1} = \min_{u \in \mathcal{C}} \left\{ \frac{S_u}{1+\lambda} w_{n+1} + R_u w_n + \frac{1}{1+\lambda} c^U \right\} \quad (2.28)$$

coincide avec la suite v_n donnée par l'équation (2.27).

Deux décomposition admissibles classiques sur lesquelles nous reviendrons au Chapitre 3 faire un pageref sont la décomposition de Gauss-Seidel et la décomposition de Jacobi.

- Considérons la décomposition de $M^u = T_B^u + T_H^u$ où T_B^u est la partie triangulaire inférieure de M^u ($(T_B^u)_{i,j} = 0$ si $j \geq i$) et T_H^u la partie triangulaire supérieure de M^u ($(T_H^u)_{i,j} = 0$ si $j < i$) alors il est facile de vérifier que la décomposition (Q_u, R_u) avec $Q_u = I - \frac{1}{1+\lambda} T_B^u$ et $R_u = T_H^u \frac{1}{1+\lambda}$ est admissible. $-Q_u$ est une matrice a diagonale fortement dominante qui est inversible et qui vérifie $Q_u^{-1} \geq 0$ en utilisant les propositions 1.3.4 et 1.3.3.

Cette décomposition est celle de Gauss-Seidel et il est important de noter que dans son implémentation l'inversion de la matrice Q_u est inutile. En effet on peut utiliser les résultats de l'exécice 2.3.6 pour écrire les équations sous la forme

$$v_{n+1}(x) = \frac{1}{1+\lambda} \min_{u \in \mathcal{C}} \{T_B^u v_{n+1}(x) + T_H^u v_n + c^u\} \quad (2.29)$$

et on se convaincra en que la résolution séquentielle des équations précédentes permet d'obtenir la solution. Quand on traite les équations de façon séquentielles, la structure triangulaire inférieure de la matrice T_B^u fait que le calcul de $v_{n+1}(x)$ ne nécessite que la connaissance des $v_{n+1}(z)$ avec $z < x$ et que ces valeurs ont été calculées aux itérations précédentes.

- La décomposition de Jacobi est la décomposition admissible suivante : Q_u est la matrice diagonale construite avec la diagonale principale de la matrice $I - \frac{1}{1+\lambda}M^u$ et R_u vaut alors bien sûr $Q_u - I + \frac{1}{1+\lambda}M^u$. A nouveau $-Q^u$ est à diagonale fortement dominante et la décomposition est bien admissible.

Le résultat que nous énonçons ici assure la convergence des variantes Gauss-Seidel et Jacobi de l'algorithme d'itérations sur les valeurs

Théorème 2.3.5. *Soient (Q_1, R_1) et (Q_2, R_2) deux décompositions admissibles de $I - \frac{1}{1+\lambda}M^U$ telles que $R_2 \leq R_1 \leq \frac{1}{1+\lambda}M^U$ alors*

$$\|Q_2^{-1}R_2\|_\infty \leq \|Q_1^{-1}R_1\|_\infty \leq \frac{1}{1+\lambda}.$$

Preuve : Montrons tout d'abord que si $(Q_i, R_i)_{i=1,2}$ sont deux décomposition admissibles telle que $R_2 \leq R_1$ alors $Q_2^{-1} \geq Q_1^{-1}$. Soit (Q_i, R_i) une décomposition admissible, on a alors $Q_i^{-1}R_i \geq 0$ et donc $\|Q_i^{-1}R_i\|_\infty = \|Q_i^{-1}R_i e\|_\infty$. De $I - Q_i = M/(1+\lambda) - R_i$ on déduit $0 \leq I - Q_i \leq M/(1+\lambda)$ et donc $\|I - Q_i\|_\infty \leq 1/(1+\lambda) < 1$. Q_i^{-1} s'écrit donc aussi :

$$Q_i^{-1} = \sum_{k=0}^{\infty} (I - Q_i)^k \quad (2.30)$$

En utilisant la relation $I - Q_i = M/(1+\lambda) - R_i$ on voit aussi que $R_2 \leq R_1$ implique $(I - Q_2) \geq (I - Q_1)$, ce qui combiné avec l'expression (2.30) donne $Q_2^{-1} \geq Q_1^{-1}$.

Compte tenu de $(I - M/(1+\lambda))e = \lambda/(1+\lambda)e \geq 0$, on a alors :

$$\begin{aligned} & Q_2^{-1}(I - M/(1+\lambda))e \geq Q_1^{-1}(I - M/(1+\lambda))e \\ \Rightarrow & Q_2^{-1}(Q_2 - R_2)e \geq Q_1^{-1}(Q_1 - R_1)e \\ \Rightarrow & Q_2^{-1}R_2e \leq Q_1^{-1}R_1e \Rightarrow \|Q_2^{-1}R_2\|_\infty \leq \|Q_1^{-1}R_1\|_\infty \end{aligned}$$

Soit maintenant $(Q_0, R_0) \equiv (I, M/(1+\lambda))$ qui définit une décomposition admissible on conclut :

$$\|Q_2^{-1}R_2\|_\infty \leq \|Q_1^{-1}R_1\|_\infty \leq \|Q_0^{-1}R_0\|_\infty = \frac{1}{1+\lambda} < 1$$

□

Ce théorème permet de comparer les trois algorithmes entre eux. On a en effet $R_{GS} \leq R_J \leq \frac{1}{1+\lambda}M$. Pour des résultats plus complets et plus fin de comparaisons des différents algorithmes on pourra se référer à [24].

Sur les mêmes idée que la décomposition de Gauss-Seidel ou celle de Jacobi on peut construire des décompositions de domaine (le domaine étant ici l'espace d'état de la chaîne de Markov) qui peuvent permètre la résolution de problèmes de grande dimensions en les décomposant en problèmes de plus petite taille avec un mécanisme de coordination pour assurer la convergence vers la fonction de coût globale.

Reprenons ici la décomposition de Jacobi mais en considérant des blocs diagonaux. On découpe M_u en 4 blocs en considérant une partition de l'espace d'état en deux sous espace $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$:

$$M^u = \begin{pmatrix} M_{1,1}^u & M_{1,2}^u \\ M_{2,1}^u & M_{2,2}^u \end{pmatrix}$$

et on considère la décomposition (Q_u, R_u) donnée par :

$$Q^u = I - \frac{1}{1+\lambda} \begin{pmatrix} M_{1,1}^u & 0 \\ 0 & M_{2,2}^u \end{pmatrix} \quad \text{et} \quad R^u = \frac{1}{1+\lambda} \begin{pmatrix} 0 & M_{1,2}^u \\ M_{2,1}^u & 0 \end{pmatrix} \quad (2.31)$$

Pour les mêmes raisons qu'auparavant la décomposition (Q_u, R_u) est admissible et $Q_u^{-1}R_u$ est contractant (Théorème 2.3.5). L'algorithme de Jacobi apparaissant comme un cas particulier ou on partitionne l'espace d'état de taille n en n sous espaces réduits a 1 élément.

On notera aussi que les hypothèses de l'exercice 2.3.6 sont vérifiées et on utilise donc la forme de l'algorithme donné par l'équation 2.28. Les itérations sur la fonction valeur s'écrivent alors en notant par v_n^i (resp. c_i^u) la restriction de v_n (resp. c^u) a \mathcal{E}_i :

$$\begin{aligned} v_{n+1}^1 &= \frac{1}{1+\lambda} \min_{u \in \mathcal{C}} \{ M_{1,1}^u v_{n+1}^1 + M_{1,2}^u v_n^2 + c_1^u \} \\ v_{n+1}^2 &= \frac{1}{1+\lambda} \min_{u \in \mathcal{C}} \{ M_{2,2}^u v_{n+1}^2 + M_{2,1}^u v_n^1 + c_2^u \} \end{aligned}$$

Les matrices $M_{1,1}^u$ et $M_{2,2}^u$ sont sous-Markoviennes, et donc quitte à rajouter des points cimetières à \mathcal{E}_1 et \mathcal{E}_2 le calcul de v_{n+1}^1 (resp. v_{n+1}^2) revient à résoudre un problème de commande optimale sur \mathcal{E}_1 (resp. sur \mathcal{E}_2) dont la fonction coût tient compte de v_n . Les sous problèmes d'optimisation sont ici résolus en parallèle (à nouveau par une méthode itérative) la coordination étant faite par mise à jour du coût instantané. On pourrait procéder de même avec une décomposition par bloc utilisant Gauss-seidel. Les mêmes idées et des développements sous des hypothèses moins restrictives sont développées dans [16].

2.3.6 Itération sur les politiques ou algorithme de Howard

On suppose ici à nouveau que le minimum est atteint pour tout $v \in V$ dans l'expression $\mathcal{L}v = \inf_{U \in \mathcal{D}^{MD}} \mathcal{L}_U v$.

La méthode d'itérations sur les politiques consiste à alterner deux étapes :

1. une étape de calcul de la politique optimale u en fonction de la fonction valeur v déjà calculée, ceci en minimisant le second membre de l'équation de Bellman.
2. une étape de résolution exacte, par inversion de matrice, de l'équation obtenue en remplaçant dans l'équation de Bellman, le contrôle optimal réel par le contrôle précédemment calculé.

L'algorithme consiste donc, à calculer deux suites $(v_n)_{n \in \mathbb{N}}$, $(u_n)_{n \in \mathbb{N}}$ (en partant d'une politique initiale u_0) définies par :

– *Étape 2n – 1.* A une stratégie u^n on associe v^n solution de :

$$(A^{u^n} - \lambda I)v^n + c^{u^n} = 0 .$$

Comme nous l'avons déjà vu cette solution existe et est unique ($v^n = \mathcal{L}_{u^n} v^n$).

– *Étape 2n.* À un coût v^n on associe une nouvelle stratégie u^{n+1} telle que

$$x \in \mathcal{E} \rightarrow u^{n+1}(x) \in \underset{u \in \mathcal{C}}{\text{Argmin}}(x) \{ ((A^u - \lambda I)v^n)(x) + c^u(x) \} .$$

u_{n+1} existe car on a supposé que le minimum est atteint pour tout $v \in V$ dans l'expression $\mathcal{L}v = \inf_{U \in \mathcal{D}^{MD}} \mathcal{L}U$.

Théorème 2.3.6. *Soit $(v_n)_{n \in \mathbb{N}}$ la suite obtenue par l'algorithme d'itération sur les politiques :*

1. *la suite v_n est décroissante et minorée par v^* .*
2. *la suite v_n est majorée par $w_n = \mathcal{L}^n v_0$. Où la suite w_n n'est autre que la suite obtenue par l'algorithme d'itération sur les valeurs que l'on initialise avec v_0 .*
3. *La suite v_n converge vers v^* de façon monotone et en norme L_∞ .*

Preuve : Par définition v_n est l'unique solution de l'équation $v = \mathcal{L}_{u_n} v$ et par définition de u_{n+1} :

$$\mathcal{L}_{u_{n+1}} v_n = \mathcal{L}v_n = \min_U \mathcal{L}U v_n \leq \mathcal{L}_{u_n} v_n = v_n$$

On a tout d'abord $\mathcal{L}v_n \leq v_n$ dont on déduit que $v_n \geq v^*$ (théorème 2.5). D'autre part, nous avons aussi $\mathcal{L}_{u_{n+1}} v_n \leq v_n$ ce qui conduit (corollaire 2.3.1) à $v_{n+1} = v^{u_{n+1}} \leq v_n$. M^U est une matrice à coefficients positifs et par conséquent l'opérateur \mathcal{L}_U est un opérateur monotone et on a donc $\mathcal{L}_{u_{n+1}} v_{n+1} \leq \mathcal{L}_{u_{n+1}} v_n$. D'où :

$$v_{n+1} = \mathcal{L}_{u_{n+1}} v_{n+1} \leq \mathcal{L}_{u_{n+1}} v_n = \mathcal{L}v_n$$

Montrons maintenant le point 2 par récurrence. La formule précédente nous donne $v_1 \leq \mathcal{L}v_0 = w_0$. Supposons $v_n \leq w_n$, on a alors :

$$v_{n+1} \leq \mathcal{L}v_n \leq \mathcal{L}w_n = w_{n+1}$$

Le point 3 est alors évident. □

Remarque : La relation $\mathcal{L}v_n \leq v_n$ donne pour $n = 0$, $w_1 = \mathcal{L}v_0 \leq v_0 = w_0$. On a alors, pour tout n , $w_{n+1} \leq w_n$ en utilisant le caractère monotone de l'opérateur \mathcal{L} (ce point est laissé en exercice). La suite générée par l'algorithme d'itération sur les valeurs est décroissante si on initialise l'algorithme avec v_0 . □

Exercice 2.3.7. *Soit $W \in V$ tel que $0 \leq \min_U [(A^U - \lambda I)W + c^U]$ montrez qu'alors $W \leq v^* \leq v^n$ pour tout n . Trouver un minorant de la suite v_n et de v^* dans le cas où la fonction $c^u(x) \geq \alpha$.*

2.3.7 Variantes sur Howard

L'algorithme de Howard s'écrit encore : Pour u_n fixé on calcule v_n vérifiant $\mathcal{L}_{u_n} v_n = v_n$ et ensuite $u_{n+1} \in \text{Argmin}_U \mathcal{L}_U v_n$. La solution v_n de $\mathcal{L}_{u_n} v_n = v_n$ est aussi la limite quand m tend vers l'infini de $\mathcal{L}_{u_n}^m w$ pour tout $w \in V$. On peut donc chercher des variantes de l'algorithme de Howard où on va remplacer le calcul exact de $\mathcal{L}_{u_n} v_n = v_n$ par une solution obtenue avec un nombre fini d'itérations de l'opérateur $\mathcal{L}_{u_n} : v_n = \mathcal{L}_{u_n}^m v_{n-1}$ (on notera qu'il faut cette fois un v_0 pour initialiser l'algorithme). Les deux cas extrêmes sont $m = 1$ qui redonne l'algorithme d'itération sur les valeurs et $m = \infty$ qui donne l'algorithme de Howard. On peut aussi pour des problèmes de grande dimension utiliser des méthodes de décompositions de domaines comme nous avons fait pour les itérations sur les valeurs [16].

On rappelle ici les notations de l'exercice 2.3.2

$$\mathcal{I}_m = \inf_U \mathcal{L}_U^m v \quad \text{et} \quad \mathcal{J}_m = \mathcal{L}^m.$$

On considère $\mathcal{W}_m v = \mathcal{L}_U^m v$ avec $U \in \mathcal{D}(v)$ ($\mathcal{D}(v) \equiv \{U \mid \mathcal{L}_U v = \mathcal{L}v\}$). Avec les notations précédentes, la variante de l'algorithme de Howard s'écrit $v_{n+1} = \mathcal{W}_m v_n$ et nous étudions ici sa convergence en montrant qu'on peut l'encadrer par deux suites convergentes.

Proposition 2.3.2. *On se donne v_0 tel que $\mathcal{L}v_0 \leq v_0$ et on considère les trois suites $x_{n+1} = \mathcal{I}_m x_n$, $y_{n+1} = \mathcal{W}_m y_n$ et $z_{n+1} = \mathcal{L}z_n$ avec $x_0 = y_0 = z_0 = v_0$. On a alors $x_n \leq y_n \leq z_n$, la suite y_n est décroissante et $\mathcal{L}y_n \leq y_n$. La suite y_n converge vers v^* .*

Preuve : Supposons $x_n \leq y_n$ et soit $W \in \mathcal{D}(y_n)$ alors $\mathcal{I}_m x_n = \min_U \mathcal{L}_U^m x_n \leq \mathcal{L}_W^m x_n \leq \mathcal{L}_W^m y_n = y_{n+1}$.

Supposons $y_n \leq z_n$ et $\mathcal{L}y_n \leq y_n$ (pour $n = 0$ cela résulte des hypothèses). Soit $W \in \mathcal{D}(y_n)$ on a alors par définition de $W : \mathcal{L}_W y_n = \mathcal{L}y_n$. De l'hypothèse $\mathcal{L}y_n \leq y_n$ on déduit par récurrence que $\mathcal{L}_W^m y_n \leq \mathcal{L}y_n \leq y_n$. On a tout d'abord $y_{n+1} \leq y_n$ et aussi $y_{n+1} \leq \mathcal{L}y_n \leq \mathcal{L}z_n = z_{n+1}$. D'autre part $\mathcal{L}y_{n+1} = \mathcal{L}\mathcal{L}_W^m y_n \leq \mathcal{L}_W^{m+1} y_n = \mathcal{L}_W^m \mathcal{L}y_n \leq \mathcal{L}_W^m y_n = y_{n+1}$. D'où l'hypothèse de récurrence pour $n + 1$ et la décroissance de la suite y_n .

De $\mathcal{L}y_n \leq y_n$ on déduit que $y_n \geq v^*$. La suite z_n n'est autre que la suite donnée par un algorithme d'itération sur les valeurs. Elle converge donc vers v^* et même en décroissant compte tenu du résultat de l'exercice 2.3.4. La suite y_n converge donc aussi vers v^* \square

Proposition 2.3.3. *L'opérateur \mathcal{I}_m ($m \geq 1$) admet le même point fixe que l'opérateur \mathcal{L} . La suite x_n converge vers v^* en décroissant.*

Preuve : Pour $m = 1$ le résultat est évident $\mathcal{I}_1 = \mathcal{L}$. On montre facilement que l'opérateur \mathcal{I}_m est contractant et on note w^* son point fixe dans V . Il est facile et laissé au lecteur de démontrer par récurrence que $\mathcal{I}_m v \leq \mathcal{I}_m w^*$. On a donc $w^* = \mathcal{I}_m w^* \geq \mathcal{J}_m w^*$. En utilisant alors les résultats de l'exercice 2.3.2 et le fait que le point fixe de \mathcal{J}_m est v^* , on a $w^* \geq v^*$. Inversement, soit $\epsilon > 0$ fixé il existe alors $W_\epsilon \in V$ vérifiant $\mathcal{L}_{W_\epsilon}^m v^* \leq \mathcal{J}_m v^* + \epsilon$ (à nouveau la démonstration se fait par récurrence). On a donc, $\mathcal{L}_{W_\epsilon}^m v^* \leq v^* + \epsilon$, soit encore $\mathcal{I}^m v^* \leq v^*$, en utilisant à nouveau les résultats de l'exercice 2.3.2 on obtient $w^* \leq v^*$.

La suite x_n est obtenue en appliquant un algorithme d'itération sur les valeurs à l'opérateur \mathcal{L}_U^m . Elle converge donc vers v^* . Soit u_0 tel que $\mathcal{L}v_0 = \mathcal{L}_{u_0} v_0$, compte tenu de l'hypothèse $\mathcal{L}v_0 \leq v_0$ on a :

$$x_1 \leq \mathcal{L}_{u_0}^m v_0 \leq v_0 = x_0.$$

La suite x_n converge donc vers v^* en décroissant compte tenu à nouveau du résultat de l'exercice 2.3.4
□

Exercice 2.3.8. Nous considérons ici pour terminer la variante de l'algorithme de Howard qui utilise une décomposition $(Q_u, R_u) = (I - \frac{1}{1+\lambda}M_Q^u, \frac{1}{1+\lambda}M_R^u)$ de M^u construite sur une décomposition de $M^u = M_R^u + M_Q^u$ avec $M_R^u \geq 0$ et $M_Q^u \geq 0$. Montrez que cette décomposition est une décomposition admissible. On se donne v_0 et u_1 tels que $v_1 \leq v_0$ et les itérations sur les politiques :

$$\begin{aligned} v_n &= \frac{1}{1+\lambda} \left(M_Q^{u_n} v_n + M_R^{u_n} v_{n-1} + c^{u_n} \right) \equiv \mathcal{T}_{u_n}(v_n, v_{n-1}) \\ u_{n+1} &\in \operatorname{Argmin} \mathcal{L}_U v_n \end{aligned}$$

On montre que la suite $v_n \downarrow v^*$ en plusieurs étapes.

1. Montrer que si $v_n \leq v_{n-1}$ alors $v_{n+1} \leq v_n$.
2. Montrer que $v^* \leq v_0$.
3. Montrer que si $v^* \leq v_n$ alors $v^* \leq v_{n+1}$.

2.4 Inéquations variationnelles

Comme dans le paragraphe 2.2.3, on peut montrer que les problèmes de temps d'arrêt optimal en horizon infini sont un cas particulier de programmation dynamique en horizon infini. Les démonstrations sont très voisines et laissées en exercice au lecteur. Par contre, nous allons voir ici comment se réalisent les Algorithmes d'itérations sur les valeurs, de Howard et leurs variantes pour les problèmes de temps d'arrêt optimal en horizon infini.

$$v(x) = \sup_{\tau} \mathbb{E} \left[\sum_{j=0}^{\tau-1} \frac{1}{(1+\lambda)^j} c^{(j)}(X^j) + \psi^{\tau}(X^{\tau}) \mid X^0 = x \right]$$

est solution de l'équation :

$$v(x) = \frac{1}{1+\lambda} \max (Mv(x) + c(x), \psi(x)) \quad (2.32)$$

qui s'écrit aussi avec les notations du paragraphe 2.2.3

$$v(x) = \frac{1}{1+\lambda} \max_{u \in \{C, S\}} \mathcal{L}_u[c, \psi]v(x)$$

où :

$$\mathcal{L}_u[c, \psi]w \equiv \begin{cases} Mw(x) + c(x) & \text{si } u = C \\ \psi(x) & \text{si } u = S \end{cases}$$

Si l'on se donne une loi de commande $U : \mathcal{E} \mapsto \{C, S\}$. L'équation :

$$v(x) = \mathcal{L}_{U(x)}[c, \psi]v(x)$$

n'est autre que la solution du problème de Dirichlet vu a la proposition 1.2.4 avec $B = \{x : U(x) = C\}B$.

L'Algorithme d'itération sur les valeurs peut s'appliquer à l'équation 2.32. Comme on sait que la solution est $\geq \psi$ on peut choisir $v_0 = \psi$ pour initialiser l'algorithme.

La décomposition admissible de Gauss-Seidel qui consiste à résoudre séquentiellement l'équation (2.29) s'écrit ici :

$$v_{n+1}(x) = \frac{1}{1+\lambda} \max(T_B v_{n+1}(x) + T_H v_n + c^u, \psi(x)) \quad (2.33)$$

où T_B et T_H sont respectivement les parties triangulaires inférieures (stritement) et supérieures de M .

La décomposition de Jacobi s'écrit quant-à elle sous la forme :

$$v_{n+1} = \max\left(Q^{-1} R v_n + Q^{-1} \frac{1}{1+\lambda} c, \psi(x)\right) \quad (2.34)$$

où Q est la matrice diagonale construite avec la diagonale principale de la matrice $I - \frac{1}{1+\lambda} M$ et R vaut alors bien sûr $Q - I + \frac{1}{1+\lambda} M$.

L'algorithme d'itération sur les politiques ou algorithme de Howard.

L'algorithme consiste donc, à calculer deux suites $(v_n)_{n \in \mathbb{N}}$, $(u_n)_{n \in \mathbb{N}}$ (en partant d'une politique initiale u_0) définies par :

– *Étape 2n - 1.* A une stratégie u_n on associe v_n solution de :

$$v_n(x) = \begin{cases} \frac{1}{1+\lambda}(M v_n(x) + c(x)) & \text{si } u_n(x) = C \\ \frac{1}{1+\lambda}\psi(x) & \text{si } u_n(x) = S \end{cases}$$

qui n'est autre qu'un problème de Dirichlet (Proposition 1.2.4).

– *Étape 2n.* À un coût v_n on associe une nouvelle stratégie $u_{n+1}(x)$ telle que $u_{n+1}(x) = C$ si $M v_n(x) + c(x) > \psi$ et $u_{n+1} = S$ dans le cas contraire.

2.5 Corrigé des exercices

Exercice 2.1.1

$$\begin{cases} v^n(x) &= \max\left(\frac{p}{1+r} v^{n+1}(x(1+a)) + \frac{1-p}{1+r} v^{n+1}(x(1+b)), f(x)\right) \quad n = 0, \dots, N-1 \\ u_n(x) &= (K-x)_+ \end{cases}$$

Exercice 2.3.2 On montre facilement que :

$$\mathcal{L}_{U_m} \cdots \mathcal{L}_{U_1} v = \sum_{k=1}^m \frac{1}{(1+\lambda)^k} M^{U_m} \cdots M^{U_{m-k+2}} c^{U_{m-k+1}} + \frac{1}{(1+\lambda)^m} M^{U_m} \cdots M^{U_1} v$$

le produit de matrices stochastiques étant stochastiques $\mathcal{L}_{U_m} \cdots \mathcal{L}_{U_1}$ s'écrit aussi :

$$\mathcal{L}_{U_m} \cdots \mathcal{L}_{U_1} v = C(U_m, \dots, U_1) + \frac{1}{(1+\lambda)^m} \overline{M}^{U_m, \dots, U_1} v$$

où $\overline{M}^{(\cdot)}$ est une matrice stochastique. Le même raisonnement qu'avec \mathcal{L} permet alors de conclure pour l'opérateur \mathcal{J}_m . Pour l'opérateur \mathcal{I}_m on peut tout d'abord montrer par récurrence que pour $v \in V$ et $\epsilon > 0$ donné Il existe (U_m, \dots, U_1) vérifiant :

$$\mathcal{L}_{U_m} \cdots \mathcal{L}_{U_1} v \leq \mathcal{I}_m v + \epsilon$$

et ce point étant vérifié la démonstration suit celle que l'on a vu pour \mathcal{L}

Exercice 2.3.3

1. Montrons l'inégalité $\|v_{N+1} - v^*\|_\infty \leq \frac{1}{\lambda} \|v_{N+1} - v_N\|_\infty$

$$\begin{aligned} \|v_{N+1} - v^*\|_\infty &= \|Lv_N - Lv^*\|_\infty \leq \frac{1}{1+\lambda} \|v_N - v^*\|_\infty \\ &\leq \frac{1}{1+\lambda} \{ \|v_N - v_{N+1}\|_\infty + \|v_{N+1} - v^*\|_\infty \} \end{aligned}$$

2. Montrons tout d'abord $\|v_{N+1} - v^{U_\epsilon}\|_\infty \leq \frac{1}{\lambda} \|v_{N+1} - v_N\|_\infty$. On utilise le caractère contractant de L et \mathcal{L}_{U_ϵ} ainsi que le fait que par définition de U_ϵ on a $Lv_{N+1} = \mathcal{L}_{U_\epsilon} v_{N+1}$ et aussi $v^{U_\epsilon} = \mathcal{L}_{U_\epsilon} v^{U_\epsilon}$.

$$\begin{aligned} \|v_{N+1} - v^{U_\epsilon}\|_\infty &\leq \|Lv_{N+1} - v^{U_\epsilon}\|_\infty + \|Lv_{N+1} - v_{N+1}\|_\infty \\ &\leq \|\mathcal{L}_{U_\epsilon} v_{N+1} - \mathcal{L}_{U_\epsilon} v^{U_\epsilon}\|_\infty + \|Lv_{N+1} - v_{N+1}\|_\infty \\ &\leq \frac{1}{1+\lambda} \{ \|v_{N+1} - v^{U_\epsilon}\|_\infty + \|v_{N+1} - v_N\|_\infty \} \end{aligned}$$

d'où le résultat.

3. De la majoration $\|v^{U_\epsilon} - v^*\|_\infty \leq \|v^{U_\epsilon} - v_{N+1}\|_\infty + \|v_{N+1} - v^*\|_\infty$ et de ce qui précède on déduit $\|v^{U_\epsilon} - v^*\|_\infty \leq \frac{2}{\lambda} \|v_{N+1} - v_N\|_\infty$ ce qui puisque $\|v_{N+1} - v_N\|_\infty \leq \frac{\lambda}{2} \epsilon$ termine la preuve.

Exercice 2.3.4 Soit $v_n \leq v_{n-1}$, pour U fixé l'opérateur \mathcal{L}_U est croissant, on a donc $\mathcal{L}_U v_{n-1} \leq \mathcal{L}_U v_n$. d'où $\mathcal{L}_U v_{n-1} \leq \min_U \mathcal{L}_U v_{n-1} = v_n$ et donc aussi $v_{n+1} = \min_U \mathcal{L}_U v_{n-1} \leq v_n$.

Exercice 2.3.5 Montrons la deuxième question. $1/(1+\lambda)M^U - I = R_U - Q_U$ est à diagonale fortement dominante de paramètre $-\lambda/(1+\lambda)$. Il est tout d'abord évident que les éléments hors diagonaux de $Q_U^{-1}R_U - I$ sont positifs. Soit e le vecteur dont toutes les composantes valent 1, comme $Q_U^{-1} \geq 0$ on a :

$$(Q_U^{-1}R_U - I)e \leq -\frac{\lambda}{(1+\lambda)} Q_U^{-1}e \leq -\frac{\lambda}{(1+\lambda)} \min_i ((Q_U^{-1}e)_i) e$$

On conclut alors sur la première partie en remarquant que Q_U étant inversible on doit avoir $\min_i ((Q_U^{-1}e)_i) \neq 0$. Par définition de w^* , pour tout $\epsilon > 0$ il existe U_ϵ vérifiant :

$$Q_{U_\epsilon}^{-1} R_{U_\epsilon} w^* + Q_{U_\epsilon}^{-1} \frac{1}{1+\lambda} c^{U_\epsilon} \leq \mathcal{H}_{R,Q} w^* + \epsilon e = w^* + \epsilon e$$

Soit par ailleurs $v^{U_\epsilon} = \mathcal{L}_{U_\epsilon} v^{U_\epsilon}$ on a $(Q_{U_\epsilon} - R_{U_\epsilon})v^{U_\epsilon} = \frac{1}{1+\lambda}c^{U_\epsilon}$ d'où :

$$(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)(w^* - v^{U_\epsilon}) \leq \epsilon e$$

D'autre part $(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)$ est inversible et on a donc

$$(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)(w^* - v^{U_\epsilon} - (Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)^{-1}\epsilon e) \leq 0$$

d'où $w^* \geq v^{U_\epsilon} + \epsilon(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)^{-1}e$ et donc $w^* \geq v^* + \epsilon(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)^{-1}e$. Il suffit alors de faire tendre ϵ vers zéro pour conclure que $w^* \geq v^*$. Pour ce faire on remarque que l'hypothèse $\xi = \sup_u \|Q_u^{-1}R_u\|_\infty < 1$ nous assure $\|(Q_{U_\epsilon}^{-1}R_{U_\epsilon} - I)^{-1}\|_\infty < 1/(1 - \xi)$ est bornée indépendamment de ϵ . Montrons maintenant la réciproque. Soit ϵ fixé, on peut trouver v_{U_ϵ} tel que $v_{U_\epsilon} \leq v^* + \epsilon$. v_{U_ϵ} est le point fixe de l'opérateur \mathcal{L}_{U_ϵ} on obtient donc :

$$\mathcal{H}_{RQ}(v_{U_\epsilon}) \leq v_{U_\epsilon}$$

On peut se persuader que le théorème () s'applique avec l'opérateur \mathcal{H}_{RQ} ($Q_{U_\epsilon}^{-1}R_{U_\epsilon}$ est une matrice sous stochastique) et on obtient donc $w^* \leq v_{U_\epsilon}$ ce qui permet de conclure.

Exercice 2.3.7 On remarque que $\mathcal{L}_U v - v = \frac{1}{1+\lambda} \{c^U + (A^U - \lambda I)v\}$ et donc l'hypothèse sur W s'écrit $\mathcal{L}_W \geq W$. On en conclut que $W \leq v^* \leq v_n$ (théorème 2.5). Dans le cas $\alpha = 0$ une solution évidente est $W \equiv 0$. Le cas α non nul se déduit alors facilement et on obtient $\alpha e \leq v^* \leq v_n$.

Exercice 2.3.6 En utilisant la remarque qui suit la proposition 1.3.3 la matrice Q_u est inversible et son inverse à tous ses éléments positifs $Q_u^{-1} \geq 0$. Soit u fixé on a alors, par définition de v , $v \leq \frac{S_u}{1+\lambda}v + \alpha_u$ et Q_u^{-1} étant croissante $v \leq Q_u^{-1}\alpha_u$. On a donc $v \leq w$. Soit maintenant $\epsilon > 0$ fixé et u_ϵ tel que

$$\frac{1}{1+\lambda}S_{u_\epsilon}v + \alpha_{u_\epsilon} \leq v + \epsilon e.$$

En utilisant à nouveau le caractère croissant de $Q_{u_\epsilon}^{-1}$ on obtient

$$Q_{u_\epsilon}^{-1}\alpha_{u_\epsilon} \leq v + K\epsilon e$$

ou K est une constante indépendante de u_ϵ car on sait (proposition 1.3.4) que $\|Q_{u_\epsilon}^{-1}\|_\infty \leq \frac{1+\lambda}{\lambda}$. on a donc $w \leq v + K\epsilon e$ et on conclut finalement que $w = v$. La fin de l'exercice est maintenant évidente.

Exercice 2.3.8

1. Par définition de u_{n+1} , $\mathcal{L}_{u_{n+1}}v_n \leq \mathcal{L}_{u_n}v_n$ et si $v_n \leq v_{n-1}$ on a $\mathcal{L}_{u_n}v_n \leq \mathcal{T}_{u_n}(v_n, v_{n-1}) = v_n$ d'où $\mathcal{L}_{u_{n+1}}v_n \leq v_n$. D'autre part :

$$\begin{aligned} v_{n+1} &= \frac{1}{1+\lambda} \left(M_Q^{u_{n+1}}v_{n+1} + M_R^{u_{n+1}}v_n + c^{u_{n+1}} \right) \\ &= \frac{1}{1+\lambda} \left(M_Q^{u_{n+1}}(v_{n+1} - v_n) \right) + \mathcal{L}_{u_{n+1}}v_n \\ &\leq \frac{1}{1+\lambda} M_Q^{u_{n+1}}(v_{n+1} - v_n) + v_n \end{aligned}$$

On a donc $Q_{u_{n+1}}(v_{n+1} - v_n) \leq 0$ et en utilisant la positivité de $Q_{u_{n+1}}^{-1}$, $v_{n+1} \leq v_n$.

2. En utilisant la question précédente :

$$\begin{aligned} v_1 &= \frac{1}{1+\lambda} \left(M_Q^{u_1}(v_1 - v_0) \right) + \mathcal{L}_{u_1} v_0 \\ &\leq \frac{1}{1+\lambda} \left(M_Q^{u_1}(v_1 - v_0) \right) + \mathcal{L} v_0 \end{aligned}$$

On a alors $Q_{u_1}^{-1}(\mathcal{L}v_0 - v_0) = v_1 - v_0 \leq 0$ et donc $\mathcal{L}v_0 \leq v_0$ d'où l'on déduit (théorème 2.5) $v_0 \geq v^*$.

3. Supposons $v_n \geq v^*$ on a alors :

$$\begin{aligned} v^* &= \mathcal{L}v^* \leq \mathcal{L}_{u_{n+1}} v^* \\ &\leq \frac{1}{1+\lambda} M_Q^{u_{n+1}}(v^* - v_{n+1}) + \mathcal{T}_{u_{n+1}}(v_{n+1}, v_n) \end{aligned}$$

ce qui permet de montrer $Q_{u_{n+1}}(v^* - v_{n+1}) \leq 0$ et donc $v^* \leq v_{n+1}$.

Chapitre 3

Processus de Diffusion

3.1 Introduction

Ce chapitre est consacré à quelques notions de base du calcul stochastique. Aucune démonstration ne sera donnée; le lecteur pourra se référer aux ouvrages cités dans la bibliographie pour une étude approfondie de ce domaine.

Les systèmes apparaissant dans les applications sont souvent soumis à des perturbations que l'on peut considérer comme aléatoires. Un modèle stochastique dans lequel le système évolue selon un processus stochastique est alors approprié. On note $X(t) = (X_1(t), \dots, X_n(t))$ le vecteur décrivant l'état du système à l'instant t . On va considérer des processus X_t qui sont solutions d'une forme perturbée de l'équation différentielle :

$$\frac{dX}{dt} = b(t, X(t)) \quad (3.1)$$

où le vecteur $X(t) = (X_1(t), \dots, X_n(t))$ décrit l'état du système à l'instant t . L'équation perturbée peut s'écrire formellement :

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \frac{dW_t}{dt} \quad (3.2)$$

où $\frac{dW_t}{dt}$ est la dérivée formelle par rapport au temps d'un mouvement brownien W . En fait, la dérivée $\frac{dW_t}{dt}$ n'existe pas. (Dans la littérature d'ingénieur, la dérivée d'un Wiener est souvent appelée *bruit blanc*). L'équation (3.2) doit être réécrite avec la notation différentielle $dX = bdt + \sigma dW$, puis intégrée d'une limite inférieure fixe s à une limite supérieure variable t . L'intégrale $\int_s^t \sigma dW$ est définie comme l'intégrale stochastique au sens de Ito.

3.2 Processus stochastiques

Soit un *espace de Probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$. Une fonction $X(t, \omega) : \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}^n$ telle que $\forall t, X(t, \omega)$ est une variable (ou vecteur) aléatoire, est appelée un *processus stochastique*. La *trajectoire* est la famille, lorsque ω varie, de fonctions $X(., \omega) : t \rightarrow X(t, \omega)$.

Une *filtration* \mathcal{F}^t est une famille croissante de sous σ -algèbres de \mathcal{F} , (i.e $\mathcal{F}^s \subset \mathcal{F}^t$ si $s \leq t$).

On dit que $X(t)$ est un *processus adapté* à \mathcal{F}^t si $X(t)$ est \mathcal{F}^t -mesurable pour tout t .

Une variable aléatoire positive $T(\omega)$ est un *temps d'arrêt* par rapport à \mathcal{F}^t si $\forall t, \{\omega, T(\omega) \leq t\} \subset \mathcal{F}^t$. Il faut interpréter \mathcal{F}^t comme mesurant l'information disponible jusqu'à l'instant t , c'est à dire que tous les événements appartenant à \mathcal{F}^t sont mesurables au sens où on sait s'ils sont ou non réalisés à l'instant t .

Une famille adaptée $(M_t)_{t \geq 0}$ de variables aléatoires intégrables (c'est à dire vérifiant $\mathbb{E}[|M_t|] < +\infty \forall t$) est une *martingale* (resp. *sur-martingale*, *sous-martingale*) si

$$\forall s \leq t, \mathbb{E}[M_t | \mathcal{F}_s] = M_s \quad (\text{resp. } \leq M_s, \geq M_s)$$

Processus de Wiener. Soit \mathcal{F}^t une filtration continue à droite (i.e. $\mathcal{F}^t = \bigcap_{s>t} \mathcal{F}^s$) et complète (i.e. contenant les sous-ensembles de Ω inclus dans un ensemble de \mathcal{F} de probabilité nulle). On appelle *processus de Wiener* (ou *mouvement Brownien*) un processus stochastique $W(t)_{t \geq 0}$, adapté à \mathcal{F}^t , à accroissements indépendants et stationnaires dont les trajectoires sont continues. Ce qui signifie qu'il vérifie :

- $W(t)$ est continu,
- $\mathbb{E}[W(t) - W(s) | \mathcal{F}^s] = 0, \forall t \geq s$
- $\mathbb{E}[(W(t) - W(s))(W(t) - W(s))^T | \mathcal{F}^s] = I(t - s), \forall t \geq s$

ou I représente la matrice identité. (Les accroissements sont stationnaires dans le sens où la distribution de $W(t) - W(s)$ ne dépend que de la différence $t - s$). Un processus de Wiener est non différentiable.

Un processus de Wiener est dit *standard* si

$$W(0) = 0, \mathbb{E}[W(t)] = 0, \mathbb{E}[W(t)^2] = t.$$

Si $W(t)_{t \geq 0}$ est un processus de Wiener standard, alors $W(t)$ est une variable aléatoire gaussienne de loi $\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx$.

Soit $W(t)$ un processus de Wiener. A partir de $W(t)$, on peut construire un nouveau processus appelé *processus de Wiener avec drift* (ou dérive en français) :

$$X(t) = at + W(t).$$

Le paramètre a est le drift.

Le processus $s(t) = e^{X(t)}$ est appelé un processus de Wiener (ou mouvement Brownien) *géométrique*.

Proposition 3.2.1. *Si $(X_t)_{t \geq 0}$ est un processus de Wiener standard, alors X_t est une martingale.*

3.3 Equations différentielles stochastiques

Soit un espace de probabilité complet $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une filtration $(\mathcal{F}_t)_{t \geq 0}$. On considère le cas des équations différentielles stochastiques à valeurs vectorielles, c'est à dire au cas où le processus évolue dans \mathbb{R}^n . On se donne

- $W_t = (W_t^1, \dots, W_t^p)$ un \mathcal{F}_t -processus de Wiener p -dimensionnel (i.e. les (W_t^i) sont des \mathcal{F}_t -processus de Wiener standards indépendants),
- $b : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n, b(s, x) = (b^1(s, x), \dots, b^n(s, x)), \sigma : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times p}$ (l'ensemble des matrices $n \times p, \sigma(s, x) = (\sigma_{i,j}(s, x))_{1 \leq i \leq n, 1 \leq j \leq p}$),
- Z une variable aléatoire \mathcal{F}_0 -mesurable à valeur dans \mathbb{R}^n .

Une équation différentielle stochastique est une équation de la forme :

$$X_t = Z + \int_0^t b(t, X_t)dt + \int_0^t \sigma(t, X_t)dW_t \quad (3.3)$$

c'est à dire que l'on cherche un processus X_t à valeurs dans \mathbb{R}^n adapté à \mathcal{F}_t tel que

$$\forall 1 \leq i \leq n, \quad X_t^i = Z^i + \int_0^t b^i(s, X_s)ds + \sum_{j=1}^p \int_0^t \sigma_{i,j}(s, X_s)dW_s^j. \quad (3.4)$$

Nous ne revenons pas ici sur la construction de l'intégrale stochastique par rapport à un mouvement Brownien (voir les ouvrages cités en référence). Rappelons simplement les conditions d'existence de l'intégrale stochastique et sa propriété fondamentale : soit $(W_t)_{t \geq 0}$ un \mathcal{F}_t -Wiener et $(H_t)_{0 \leq t \leq T}$ un processus \mathcal{F}_t -adapté. On peut définir l'intégrale stochastique $(\int_0^t H_s dW_s)_{0 \leq t \leq T}$ dès que $\int_0^T H_s^2 ds < +\infty$ p.s. . Le processus $(\int_0^t H_s dW_s)_{0 \leq t \leq T}$ est une martingale si $\mathbb{E} \left[\int_0^T H_s^2 ds \right] < +\infty$.

Une solution de (3.3) est appelé processus de *diffusion*. On note formellement :

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = Z \quad (3.5)$$

Le théorème suivant donne des conditions suffisantes sur b et σ pour avoir l'existence et l'unicité d'une solution de (3.3).

Théorème 3.3.1. *Si b et σ sont des fonctions continues et s'il existe une constante K telle que*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|, \quad (3.6)$$

$$|b(t, x)| + |\sigma(t, x)| \leq K(1 + |x|), \quad (3.7)$$

$$\mathbb{E} [|Z|^2] < +\infty \quad (3.8)$$

alors $\forall T > 0$, il existe une solution unique de (3.3) dans $[0, T]$. De plus cette solution $(X_s)_{0 \leq s \leq T}$ vérifie

$$\mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s|^2 \right] < \infty \quad (3.9)$$

(voir preuve dans [?] page 60).

Un processus de diffusion est *Markovien*, c'est à dire que le comportement futur de ce processus après s dépend uniquement de X_s et non pas de ce qui s'est passé avant s . Cela se traduit mathématiquement par :

$$\mathbb{E} [f(X_t)|\mathcal{F}_s] = \mathbb{E} [f(X_t)|X_s], \quad \forall s \leq t$$

pour toute fonction f borélienne bornée.

3.3.1 Formule d'Ito

Pour les solutions de l'équation différentielle stochastique (3.5), la règle usuelle de différentiation n'est pas valide. Elle doit être remplacée par la *formule d'Ito*. Cette formule est cruciale dans le calcul différentiel stochastique.

Soit $X(t)$ un processus stochastique défini par la forme différentielle :

$$dX(t) = b(t)dt + \sigma(t)dW(t), \quad X(0) = X_0$$

où $b(t)$ et $\sigma(t)$ sont des processus adaptés à \mathcal{F}^t .

Soit $v(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction deux fois différentiable par rapport à x et une fois par rapport à t . On note $Dv(t, x) = (\frac{\partial v}{\partial x_1}, \dots, \frac{\partial v}{\partial x_n})$ son gradient par rapport à x au point (t, x) , et $D^2v(t, x) = (\frac{\partial^2 v}{\partial x_i \partial x_j})_{ij}$ sa matrice hessienne (matrice symétrique $n \times n$ des dérivées secondes par rapport à x).

On a la formule de développement suivante :

$$\begin{aligned} v(t, X(t)) &= v(0, X_0) + \int_0^t \left(\frac{\partial v}{\partial t}(s, X(s)) + Dv(s, X(s)) \cdot b(s) \right) ds \\ &+ \frac{1}{2} \int_0^t \text{tr} D^2v(s, X(s)) \sigma \sigma^T(s) ds + \int_0^t Dv(s, X(s)) \cdot (\sigma(s) dW(s)) \end{aligned}$$

où σ^T désigne la matrice transposée de σ et $\text{tr}M$ désigne la trace de la matrice M , c'est à dire la somme de ses éléments diagonaux. Si A et B sont 2 matrices symétriques $n \times n$, $\text{tr}BA = \sum_{i,j=1}^n B_{ij}A_{ij}$.

Des extensions de la formule de Ito pour des fonctions v moins régulières sont données dans [7]. La formule d'Ito est également souvent utilisée sous forme différentielle : notons $\eta(t) = v(t, X(t))$. On a

$$d\eta(t) = \frac{\partial v}{\partial t}(t, X(t))dt + Dv(t, X(t))dX + \frac{1}{2} \text{tr} D^2v(t, X(t)) \sigma \sigma^T(t) dt.$$

3.3.2 Générateur infinitésimal d'une diffusion

A une solution d'une équation différentielle stochastique est associée naturellement un opérateur du deuxième ordre qui caractérise sa loi. On le voit apparaître de la manière suivante :

Théorème 3.3.2. *Soit f une fonction de classe $\mathcal{C}^2(\mathbb{R}^n)$ à dérivées bornées et soit X_t une solution de*

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s$$

où W_s est un \mathcal{F}^t mouvement Brownien p -dimensionnel. Alors :

$$M_t = f(X_t) - f(X_0) - \int_0^t (Af)(X_s) ds$$

est une \mathcal{F}_t -martingale où

$$Af = \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{j=1}^n b_j(x) \frac{\partial f}{\partial x_j}, \quad (3.10)$$

$a = \frac{1}{2} \sigma \sigma^T$, c'est à dire $a_{ij} = \frac{1}{2} \sum_{k=1}^p \sigma_{ik} \sigma_{jk}$.

L'opérateur A est appelé le *générateur infinitésimal de la diffusion* (3.3).

Preuve : Utilisant la formule d'Ito, on prouve que

$$\begin{aligned} f(X_t) - f(X_0) - \int_0^t Af(X_s)ds &= \int_0^t \sum_j \frac{\partial f}{\partial x_j}(X_s) \left(\sum_k \sigma_{jk}(X_s) dW_s^k \right) \\ &= \sum_{j,k} \int_0^t \frac{\partial f}{\partial x_j}(X_s) \sigma_{jk}(X_s) dW_s^k \end{aligned}$$

Rappelons que le processus $(\int_0^t H_s dW_s)_{0 \leq t \leq T}$ est une martingale si $\mathbb{E} \left[\int_0^T H_s^2 ds \right] < \infty$.

Utilisant le fait que f est à dérivées bornées ainsi que les propriétés (3.7) et (3.9), on obtient :

$$\mathbb{E} \left[\int_0^T \left| \frac{\partial f}{\partial x_j} \sigma_{jk}(X_s) \right|^2 ds \right] < K(1 + \mathbb{E} \left[\sup_{s \leq T} |X_s|^2 \right]) < \infty.$$

□

Ce théorème se généralise au cas dépendant du temps et aux quantités actualisées :

Théorème 3.3.3. Soit $u(t, x)$ une fonction de classe $\mathcal{C}^{1,2}$ à dérivées en x bornées, $r(t, x)$ est une fonction continue bornée sur $\mathbb{R}^+ \times \mathbb{R}$, et (X_t) une diffusion solution de

$$dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, \quad X_0 = x.$$

Si A_s est l'opérateur associé à cette diffusion, alors

$$M_t = e^{-\int_0^t r(s, X_s)ds} u(t, X_t) - \int_0^t e^{-\int_0^s r(\theta, X_\theta)d\theta} \left(\frac{\partial u}{\partial t} + A_s u - r u \right) (s, X_s) ds$$

est une martingale.

Preuve : On différencie le produit $e^{\int_0^t r(s, X_s)ds} u(t, X_t)$ puis on applique la formule d'Ito. □

Propriétés du générateur infinitésimal

– **Ellipticité** Soit A un opérateur du deuxième ordre qui à une fonction f de classe $\mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ associe

$$Af = \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i},$$

où les coefficients a_{ij} et b_j sont des fonctions réelles. On suppose, sans perte de généralité, que $a_{ij} = a_{ji}$.

Definition 3.3.1. On dit que l'opérateur A est elliptique si et seulement si

$$\sum_{i=1}^n a_{ij}(x) \eta_i \eta_j \geq 0, \quad \forall x \in \mathbb{R}^n, \quad \eta \in \mathbb{R}^n.$$

Ceci est équivalent à dire que la matrice $(a_{ij}(x))_{i,j}$ est symétrique semi-définie positive dans \mathbb{R}^n , ou encore que les valeurs propres de la matrice $(a_{ij}(x))_{i,j}$ sont positives. S'il existe $c > 0$ tel que

$$\sum_{i=1}^n a_{ij}(x)\eta_i\eta_j \geq c|\eta|^2, \forall x \in \mathbb{R}^n, \eta \in \mathbb{R}^n, \quad (3.11)$$

l'opérateur est dit uniformément elliptique et toutes les valeurs propres de la matrice $(a_{ij}(x))_{i,j}$ sont strictement positives. Lorsqu'une ou plusieurs valeurs propres sont nulles, la relation (3.11) n'est pas satisfaite et l'équation est dégénérée.

Le générateur infinitésimal A défini en (3.10) d'une diffusion est elliptique (éventuellement dégénérée) puisque $a = \frac{1}{2}\sigma\sigma^T$.

- **Principe du Maximum.** On dit qu'un opérateur A satisfait au *Principe du Maximum* si et seulement si :

$\forall v \in \mathcal{D}^2(\mathbb{R}^n)$, si x_0 est un point de maximum local de v , alors $Av(x_0) \leq 0$.

Le générateur A d'une diffusion satisfait le Principe du maximum.

3.4 Equations de Kolmogorov

Nous allons montrer dans ce qui suit qu'il y a une forte connection entre les processus de diffusion et les équations différentielles du 2ème ordre parabolique ou elliptiques.

3.4.1 Problèmes à horizon fini

Plus généralement, considérons la fonctionnelle

$$v(t, x) = \mathbb{E} \left[\int_t^T e^{-\lambda s} u(s, X_s) ds + e^{-\lambda(T-t)} g(X_T) | X_t = x \right]. \quad (3.12)$$

Sous certaines hypothèses de régularité, la fonction v est solution de l'équation de Kolmogorov rétrograde :

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + Av(t, x) - \lambda v(t, x) + u(t, x) = 0 & \text{dans } [0, T[\times \mathbb{R}^n, \\ v(T, x) = g(x) & \text{dans } \mathbb{R}^n. \end{cases} \quad (3.13)$$

Le coefficient d'actualisation peut dépendre de t et x ou bien être identiquement nul. Il existe des théorèmes d'existence et d'unicité des solutions de l'équation de Kolmogorov permettant d'affirmer que la solution de l'équation est bien égale au prix. Ces résultats s'obtiennent en général sous une condition d'uniforme ellipticité de l'opérateur A . Lorsqu'il n'y a pas de solution forte, il faut définir une notion de solution faible. L'approche usuelle est *l'approche variationnelle* : on définit une solution *faible* en utilisant les espace de Sobolev, puis on établit l'existence et l'unicité de cette solution par un théorème de type Lax-Milgram, enfin on démontre des résultats de régularité sur la solution (voir [27]).

Application à l'évaluation des options européennes.

Soit $u(t, x)$ une fonction de classe $\mathcal{C}^{1,2}$ à dérivées en x bornées, solution de :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + Au(t, x) - ru = 0 & \text{dans } [0, T[\times \mathbb{R}^n, \\ u(T, x) = f(x) & \text{dans } \mathbb{R}^n \end{cases} \quad (3.14)$$

où A est le générateur de la diffusion

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

b et σ vérifient les conditions (3.6) et (3.7), $r(x, t)$ est une fonction continue, positive et bornée et f est une fonction continue à croissance polynomiale. Alors

$$u(t, x) = \mathbb{E} \left[e^{-\int_t^T r(s, X_s)ds} f(X_T) | X_t = x \right].$$

Le prix d'une option européenne à l'instant t peut s'écrire dans le cas de marchés complets

$$V_t = \mathbb{E} \left[e^{-\int_t^T r(s, X_s)ds} f(X_T) | \mathcal{F}_t \right]$$

et $V_t = u(t, X_t)$.

En effet, si u est solution de (3.14), le théorème 3.3.3 permet d'assurer que $M_t \equiv e^{-\int_t^T r(s, X_s)ds} u(t, X_t)$ est une martingale, donc

$$\mathbb{E} [M_T | \mathcal{F}_t] = M_t,$$

soit en utilisant la condition finale

$$\mathbb{E} \left[e^{-\int_0^T r(s, X_s)ds} f(X_T) | \mathcal{F}_t \right] = e^{-\int_0^t r(s, X_s)ds} u(t, X_t)$$

ce qui implique

$$\mathbb{E} \left[e^{-\int_t^T r(s, X_s)ds} f(X_T) | \mathcal{F}_t \right] = u(t, X_t).$$

Comme le processus $(X_t)_{t \geq 0}$ est markovien, $V_t = u(t, X_t)$.

Exemples :

– $X_t = W_t, r = 0$

$$\mathbb{E} [f(X_t) | \mathcal{F}_t] = u(t, X_t)$$

$u(t, x)$ est solution de l'équation de la chaleur (qui s'écrit habituellement avec une condition initiale) :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} = 0 & \text{dans } \mathbb{R}^n \times [0, T[, \\ v(T, x) = f(x). & \text{dans } \mathbb{R}^n \end{cases} \quad (3.15)$$

– **modèle de Black-Scholes**

$$dS_t = S_t(rdt + \sigma dW_t)$$

$$\mathbb{E} \left[e^{-r(T-t)} f(S_T) | \mathcal{F}_t \right] = u(t, S_t)$$

où $u(t, x)$ est solution de

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + rx \frac{\partial u}{\partial x} - ru = 0 & \text{dans } [0, T[\times \mathbb{R}^n, \\ v(T, x) = f(x). & \text{dans } \mathbb{R}^n \end{cases} \quad (3.16)$$

Dans le cas d'un *call*, $f(x) = (x - K)^+$, dans le cas d'un *put*, $f(x) = (K - x)^+$, où $x^+ = \sup(0, x)$.

– **modèle de Black-Scholes logarithmique**

On considère le changement de variable

$$Y_t = \log S_t.$$

$$dY_t = (r - \sigma^2)dt + \sigma dW_t$$

Le prix de l'option s'écrit $u(t, x) = v(t, \log x)$ où

$$v(t, y) = \mathbb{E} \left[e^{-r(T-t)} f(e^{Y_T}) | Y_t = y \right]$$

$v(t, y)$ est solution de

$$\begin{cases} \frac{\partial v}{\partial t}(t, y) + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial y^2} + (r - \frac{1}{2}\sigma^2) \frac{\partial v}{\partial y} - rv = 0 & \text{dans } [0, T] \times \mathbb{R}^n, \\ v(T, y) = f(e^y) & \text{dans } \mathbb{R}^n. \end{cases} \quad (3.17)$$

3.4.2 Problèmes à horizon infini

Soit $X(t)$ un processus de diffusion défini par

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \quad (3.18)$$

Considérons le coût actualisé sur un horizon infini

$$v(x) = \mathbb{E} \left[\int_0^\infty e^{-\lambda t} u(X_t) dt | X_0 = x \right] \quad (3.19)$$

où $\lambda > 0$ est le taux d'actualisation (qui peut être une fonction de X_t) et la fonction $u : \mathbb{R}^n \rightarrow \mathbb{R}^+$ est parfois appelé le coût instantané. La fonction v définie par (3.19) est solution (sous certaines hypothèses de régularité des fonctions b, σ et u) de l'équation de Kolmogorov :

$$-\lambda v(x) + Av(x) + u(x) = 0 \quad \text{dans } \mathbb{R}^n. \quad (3.20)$$

où A est le générateur infinitésimal de la diffusion $X(t)$. Réciproquement, une solution régulière v de (3.20) possédant en plus la propriété

$$\lim_{t \rightarrow \infty} e^{-\lambda t} \mathbb{E}_x [v(X(t))] = 0 \quad (3.21)$$

peut s'interpréter en utilisant un théorème de vérification comme le critère (3.19) ([?]).

3.4.3 Conditions aux bords

On a supposé dans ce qui précède que l'état X_t du système pouvait prendre n'importe quelle valeur de \mathbb{R}^n . Dans le cas où l'état du système est contraint de rester dans un domaine \mathcal{O} de \mathbb{R}^n , nous devons préciser le comportement du système à la frontière du domaine.

Processus arrêté à la frontière

On peut décider d'arrêter le système lorsqu'il sort du domaine. On définit le temps d'arrêt τ comme étant le premier instant où le processus X_t sort du domaine \mathcal{O} :

$$\tau(\omega) = \inf\{t, X_t(\omega) \notin \mathcal{O}\}$$

et le processus est dit arrêté à la frontière (frontière absorbante). Dans ce cas, il y a en général un coût (où un gain) d'arrêt f qu'il faut prendre en compte dans la définition du critère.

- **le cas d'un horizon fini** T . On considère

$$\begin{aligned} v(t, x) &= \mathbb{E}_{x,t} \left[\int_t^{\tau \wedge T} e^{-\lambda s} u(s, X_s) ds + e^{-\lambda(\tau-t)} f(\tau, X_\tau) \mathbb{I}_{\{\tau < T\}} \right] \\ &+ \mathbb{E}_{x,t} \left[e^{-\lambda(T-t)} g(X_T) \mathbb{I}_{\{T < \tau\}} \right] \end{aligned}$$

où \mathbb{I}_A désigne la fonction caractéristique de l'ensemble A . On suppose que f est à croissance polynomiale.

La fonction v satisfait l'équation de Kolmogorov (3.13) dans $[0, T[\times \mathcal{O}$ avec la condition au limite de **Dirichlet**

$$v(t, x) = f(t, x) \quad \text{sur } [0, T[\times \partial\mathcal{O}. \quad (3.22)$$

Exemple : Option "extinguishable" On suppose que dès que le cours de l'option sort de l'ouvert $\mathcal{O} =]a, b[$, la valeur de l'option est nulle. Le prix de cette option est donnée par :

$$u(t, x) = \mathbb{E} \left[\mathbb{I}_{\{\forall s \in [t, T[, X_s \in \mathcal{O}\}} e^{-\int_t^T r(s, X_s) ds} g(X_T) | X_t = x \right]$$

et satisfait :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + Au - ru = 0 & \text{dans } [0, T[\times]a, b[, \\ u(t, a) = u(t, b) = 0 & \text{dans }]0, T[, \\ u(T, x) = g(x) & \text{dans }]a, b[. \end{cases}$$

- **le cas d'un horizon infini**. On considère

$$v(x) = \mathbb{E} \left[\int_0^\tau e^{-\lambda t} u(X_t) dt + e^{-\lambda \tau} f(X_\tau) | X_0 = x \right] \quad (3.23)$$

Dans ce cas, la fonction $v(x)$ est solution de l'équation de Kolmogorov (3.20) dans \mathcal{O} avec la condition au limite de Dirichlet

$$v(x) = f(x) \quad \text{sur } \partial\mathcal{O}. \quad (3.24)$$

Processus réfléchi à la frontière.

On peut également considérer des processus réfléchis sur la frontière de \mathcal{O} . Dans ce cas, l'équation d'évolution du processus s'écrit

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t - n_{X_t} d\xi_t \quad (3.25)$$

ξ_t est un processus strictement croissant lorsque $X_t \in \partial\mathcal{O}$ et $d\xi_t = 0$ si $X_t \in \mathcal{O}$, n_x est la normale extérieure à la frontière $\partial\mathcal{O}$ de \mathcal{O} en x . En d'autres termes, lorsque le processus X_t

atteint la frontière du domaine, il est réfléchi à l'intérieur du domaine dans la direction de la normale intérieure.

Exemples en dim. 1 :

Il y a en général un coût (ou un gain) associé à la réflexion du processus :

- **le cas d'un horizon fini** T . La fonction

$$v(t, x) = \mathbb{E}_{x,t} \left[\int_t^T e^{-rs} u(s, X_s) ds + \int_t^T e^{-rs} f(s, X_s) d\xi_s + e^{-r(T-t)} g(X_T) \right]$$

satisfait l'équation de Kolmogorov (3.13) dans $[0, T[\times \mathcal{O}$ avec la condition au limite de **Neumann**

$$\frac{\partial v(t, x)}{\partial n} = f(t, x) \quad \text{sur } [0, T[\times \partial \mathcal{O}. \quad (3.26)$$

- **le cas d'un horizon infini** on définit

$$v(x) = \mathbb{E}_x \left[\int_0^{+\infty} e^{-\lambda t} u(X_t) dt + \int_0^{+\infty} e^{-\lambda t} f(X_t) d\xi_t \right].$$

La fonction v est solution de l'équation de Kolmogorov (3.20) dans \mathcal{O} avec la condition au limite de Neumann :

$$\frac{\partial v}{\partial n} = f \quad \text{sur } \partial \mathcal{O}. \quad (3.27)$$

3.4.4 Evaluation d'un cout moyen - le cas ergodique

3.5 Les Equations de Fokker-Planck

3.5.1 Exemple : Concentration de particules dans un décanteur

Equation de densité des particules

3.6 Processus de diffusion avec sauts

3.6.1 Equation intégrro-différentielle

3.6.2 Exemple

3.7 Exercices

Exercice 3.7.1. Option sur moyenne *On considère une option dont le strike est la moyenne sur la période de temps considérée. On suppose que le prix S_t du sous-jacent évolue selon l'équation*

$$dS_t = S_t(rdt + \sigma dW_t), \quad S_0 = s.$$

où r, σ et $s > 0$. Il s'agit d'évaluer

$$V(t, s) = \mathbb{E} \left[e^{-\delta(T-t)} (S_T - \frac{1}{T} \int_0^T S_\theta d\theta)_+ | S_t = s \right] \quad \text{où } \delta > 0 \text{ et } T > 0.$$

1. On pose $X_t^1 = S_t$ et $X_t^2 = \int_0^t S_\theta d\theta$. Montrer que $X_t = (X_t^1, X_t^2)$ satisfait une équation différentielle stochastique dans \mathbb{R}^2 que l'on explicitera. Ecrire le générateur infinitésimal associé à la diffusion X_t .
2. Montrer que $V(t, s) = u(t, x_1, x_2)$ avec

$$u(t, x_1, x_2) = \mathbb{E} \left[e^{-\delta(T-t)} \Phi(X_1^T, X_2^T) | X_1^t = x_1, X_2^t = x_2 \right]$$

et Φ est une fonction que l'on explicitera.

3. Ecrire l'équation de Kolmogorov satisfaite par la fonction u .
4. Posons à présent $\delta = 0$. On veut résoudre numériquement l'équation :

$$\begin{cases} \frac{\partial u}{\partial t} + Au(t, x_1, x_2) = 0 & \text{dans } [0, T] \times \Omega \\ u(T, x_1, x_2) = (x_1 - \frac{x_2}{T})_+ & \text{dans } \Omega \\ u(t, x_1, x_2) = 0 & \text{dans } [0, T] \times \partial\Omega \end{cases} \quad (3.28)$$

avec $\Omega =]0, 1]^2$ et $Au = \frac{1}{2} \sigma^2 x_1^2 \frac{\partial^2 u}{\partial x_1^2} + r x_1 \frac{\partial u}{\partial x_1} + x_1 \frac{\partial u}{\partial x_2}$.

- (a) Soit $h > 0$ un pas de discrétisation en espace. Proposer une méthode d'approximation de l'opérateur A en utilisant des discrétisations de type différences finies. Quelle est la condition à imposer à h pour que cette discrétisation soit stable ?
 - (b) Soit Δt un pas de discrétisation en temps. Ecrire une approximation de l'équation (3.28) en utilisant un schéma explicite pour la discrétisation en temps. Donner une condition liant les pas h et Δt assurant la stabilité du schéma numérique.
 - (c) Donner l'interprétation probabiliste de l'équation discrétisée en termes de chaînes de Markov.
5. On s'intéresse à présent au cas d'une option américaine sur moyenne. Ecrire l'inéquation variationnelle satisfaite par la fonction

$$\tilde{u}(t, x_1, x_2) = \sup_{t \leq \tau(\text{temps d'arrêt}) \leq T} \mathbb{E} \left[e^{-\delta(\tau-t)} \Phi(X_1^\tau, X_2^\tau) | X_1^t = x_1, X_2^t = x_2 \right].$$

Chapitre 4

Etude numérique des équations de Kolmogorov

4.1 Introduction

Sauf dans quelques cas particuliers, il n'est pas possible de résoudre explicitement les problèmes d'évaluation d'espérance et de contrôle stochastique et il faut se tourner vers des méthodes numériques.

On peut aborder ces problèmes par

- **des méthodes probabilistes.** Considérons la diffusion

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0. \quad (4.1)$$

On approche la diffusion X_t par une chaîne de Markov $(S_n, n \geq 0)$ de matrice de transition M telle que $S_0 = x_0$. Sous certaines hypothèses sur M , S_n converge en loi vers X_t .

- **des méthodes d'analyse numérique.** On part de l'équation aux dérivées partielles satisfaite par la fonctionnelle recherchée et on la résout numériquement. Pour cela, on la discrétise de manière adéquate et on applique les algorithmes de résolution des équations de la Programmation-Dynamique en temps et espaces discrets.

On décrit ici ce deuxième type de méthodes sur des exemples particuliers.

4.2 Équations de Kolmogorov elliptiques

On veut résoudre l'équation :

$$\begin{cases} Av(x) - \lambda v(x) + c(x) = 0 & \text{dans } \Omega \\ v = \Phi \text{ ou } \frac{\partial v}{\partial n} = \Phi & \text{sur } \partial\Omega \end{cases} \quad (4.2)$$

où Ω est un ouvert de \mathbb{R}^d et A est un opérateur du deuxième ordre elliptique

$$Av(x) \equiv \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(x) \frac{\partial v}{\partial x_i}(x)$$

avec

$$\sum_{i,j=1}^d a_{ij}(x)\eta_i\eta_j \geq 0, \quad \forall x \in \Omega, \eta \in \mathbb{R}^d, \quad \lambda > 0,$$

et $\frac{\partial v}{\partial n}$ désigne la dérivée le long de la normale extérieure de l'ouvert Ω . On rappelle que ces équations sont liées à l'évaluation d'espérance de fonctionnelle de diffusion sur un horizon infini. Par exemple la solution de l'équation (4.2) avec conditions de Dirichlet homogène s'interprète comme (équation 3.23 page 67)

$$v(x) = \mathbb{E} \left[\int_0^\tau e^{-\lambda t} c(X_t) dt + e^{-\lambda \tau} \Phi(X_\tau) | X_0 = x \right]$$

en supposant que $\tau < \infty$ p.s où τ est le premier instant où le processus X_t sort du domaine Ω , $\tau(\omega) = \inf\{t, X_t(\omega) \notin \Omega\}$.

4.2.1 Réduction à un domaine borné

Si le domaine Ω est non borné, il est tout d'abord nécessaire, en vue d'une résolution numérique de l'équation (4.2), de se ramener à un domaine borné. Pour cela, deux procédures peuvent être utilisées :

– **Localisation du problème**

Cela consiste à se restreindre à un domaine borné et à imposer des conditions aux limites adéquates. Pour cela il est nécessaire d'étudier le comportement à l'infini de la solution. Cette méthode induit une erreur qu'il faudra estimer.

Par exemple, si $\Omega = \mathbb{R}$, la localisation consiste à se restreindre à un domaine du type $\mathcal{O}_l =]-l, l[$, l étant une constante à choisir soigneusement pour que l'algorithme soit efficace. Il faut de plus imposer des conditions aux limites au bord. Par exemple, dans le cas où on impose de conditions de Dirichlet homogène, on obtient l'EDP :

$$\begin{cases} \tilde{L}v(t, x) = 0 & \text{dans } \mathcal{O}_l, \\ v(l) = v(-l) = 0 & . \end{cases} \quad (4.3)$$

On peut envisager d'autres conditions aux limites (conditions de Neumann, conditions mixtes, ou conditions imposées par le problème lui-même).

Il faudra étudier en détail la façon dont le problème avec condition au limite converge vers le problème sur \mathbb{R} .

– **Changement de variables**

On peut également procéder à un changement de variables adéquat qui transforme le domaine initial en un domaine borné. Cette méthode a l'avantage de ne pas induire d'erreur supplémentaire.

Supposons par exemple que $\Omega =]0, +\infty[$. Le changement de variable $y = \frac{x}{1+x}$ transforme Ω en $]0, 1[$.

4.2.2 Méthode des différences finies

Trouver une bonne discrétisation du problème est un point très important de la méthode numérique. Pour discrétiser l'équation, on peut utiliser des méthodes d'approximation de type éléments finis ou différences finies. La méthode des éléments finis est généralement associée à

la recherche d'une solution faible de l'équation mise sous forme variationnelle dans un espace de Sobolev adéquat.

On s'intéresse ici à la méthode d'approximation de type différences finies qui consiste à trouver une approximation de la solution d'une EDP aux noeuds d'une grille régulière $h\mathbb{Z}^d \equiv \{hz, z_i \in \mathbb{Z}, i = 1, \dots, d\}$ (le pas de la grille ici h peut bien sur ne pas être le même dans chaque coordonnée d'espace, dans ce cas h désignera le vecteur $h = (h_1, \dots, h_d)$). On s'attend à ce que la qualité de l'approximation soit d'autant meilleure que le pas h de la grille est petit. On définit :

$$\Omega_h = \Omega \cap h\mathbb{Z}^d \quad \text{et} \quad \partial\Omega_h = \{\mathbb{R}^d \setminus \Omega\} \cap h\mathbb{Z}^d$$

et on notera (e_1, \dots, e_d) la base canonique de \mathbb{R}^d .

Un opérateur différentiel Lv sera remplacé par un opérateur discret $L^h v$ qui n'utilise que les valeurs de v sur la grille $h\mathbb{Z}^d$. Soit donc v^h une fonction définie sur la grille $h\mathbb{Z}^d$, $L^h v^h(x)$ pour $x \in \Omega_h$ utilisera les valeurs de v^h sur un sous ensemble du bord $\partial\Omega_h$ que l'on notera $\partial^+\Omega_h$. Par exemple, la Figure 4.1 montre $\partial\Omega_h$ et $\partial^+\Omega_h$ pour une discrétisation du Laplacien en utilisant les schémas (4.8).

Donnons maintenant les approximations usuelles utilisées pour construire L^h .

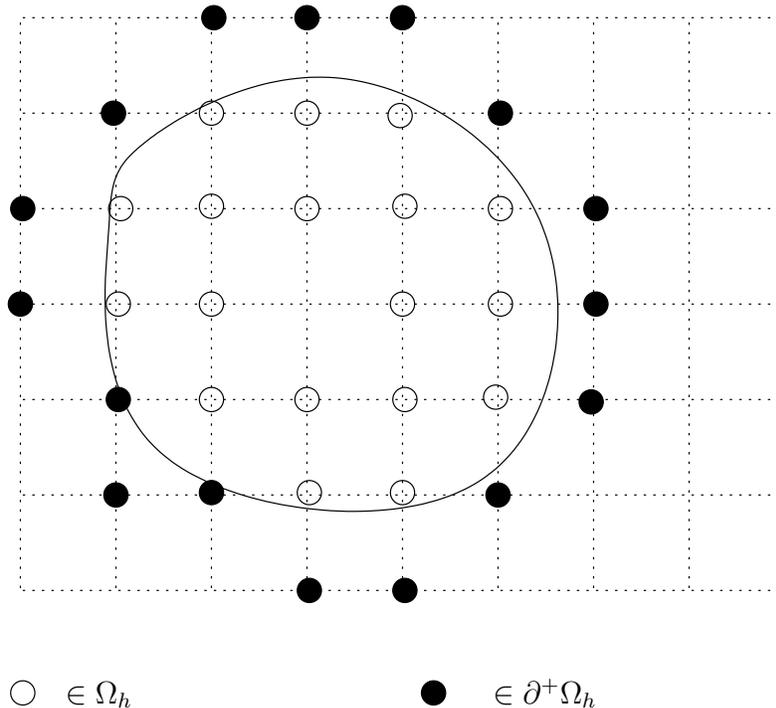


FIG. 4.1 – Exemple de grille utilisable pour $a(x)$ sans termes croisés

Approximation des dérivées premières

On cherche à approcher la valeur des dérivées de v en $x \in \Omega_h$ en fonction des valeurs de v aux points voisins. Selon les cas que l'on explicitera plus tard, on fait les approximations suivantes des dérivées premières

– schéma symétrique ou saute-mouton

$$\frac{\partial v}{\partial x_i}(x) \sim \frac{v(x + he_i) - v(x - he_i)}{2h} \equiv \partial_i^h v \quad (4.4)$$

– schéma décentré positif ou négatif

$$\frac{\partial v}{\partial x_i}(x) \sim \frac{v(x + he_i) - v(x)}{h} \quad (4.5)$$

$$\frac{\partial v}{\partial x_i}(x) \sim \frac{v(x) - v(x - he_i)}{h} \quad (4.6)$$

Approximation de la dérivée normale ($\Omega = (0, 1)^d$)

$$\frac{\partial v}{\partial n} \sim \partial_n^h \equiv \begin{cases} \partial_i^h v & \text{sur } \{x_i = 1\} \\ -\partial_i^h v & \text{sur } \{x_i = 0\}. \end{cases} \quad (4.7)$$

Approximation des dérivées secondes

$$\frac{\partial^2 v}{\partial x_i^2}(x) \sim \frac{v(x + he_i) - 2v(x) + v(x - he_i)}{h^2}, \quad (4.8)$$

– Approximation des dérivées secondes croisées ($i \neq j$) :
si $a_{ij}(x) \geq 0$

$$\begin{aligned} \frac{\partial^2 v}{\partial x_i \partial x_j}(x) &\sim \frac{2v(x) + v(x + he_i + he_j) + v(x - he_i - he_j)}{2h^2} \\ &- \left[\frac{v(x + he_i) + v(x - he_i) + v(x + he_j) + v(x - he_j)}{2h^2} \right] \end{aligned} \quad (4.9)$$

si $a_{ij}(x) < 0$

$$\begin{aligned} \frac{\partial^2 v}{\partial x_i \partial x_j}(x) &\sim - \frac{[2v(x) + v(x + he_i - he_j) + v(x - he_i + he_j)]}{2h^2} \\ &+ \frac{v(x + he_i) + v(x - he_i) + v(x + he_j) + v(x - he_j)}{2h^2} \end{aligned} \quad (4.10)$$

– D'autres approximations des dérivées secondes croisées sont possibles comme :

$$\begin{aligned} \frac{\partial^2 v}{\partial x_i \partial x_j}(x) &\sim \frac{v(x + he_i + he_j) - v(x + he_i - he_j)}{4h^2} \\ &+ \frac{v(x - he_i - he_j) - v(x - he_i + he_j)}{4h^2} \end{aligned} \quad (4.11)$$

Conditions au bord Sur le bord $\partial^+\Omega_h$, on utilise l'une des deux équations :

$$v(x) = \Phi(x) \text{ sur } \partial^+\Omega_h \text{ (Problème de Dirichlet)}$$

$$\partial_n^h v(x) = \Phi(x) \text{ sur } \partial^+\Omega_h \text{ (Problème de Neumann).}$$

Si nous reprenons le problème de départ (4.2) en posant :

$$Lv \equiv Av - \lambda v$$

Le problème de Dirichlet

$$Lv(x) + c(x) = 0 \text{ pour } x \in \Omega \text{ et } v = \Phi \text{ pour } x \in \partial\Omega \quad (4.12)$$

conduit au problème discret :

$$L_h v^h + c_h = 0 \text{ pour } x \in \Omega_h \text{ et } v^h(x) = \Phi(x) \text{ pour } x \in \partial^+\Omega_h \quad (4.13)$$

En éliminant les valeurs connues au bord, on obtient ainsi un système de N_h équations linéaires à N_h inconnues où $N_h = \text{card}(\Omega_h)$. Sur

$$V_h = \{v^h | v^h(x) = \Phi(x), x \in \partial^+\Omega_h\}$$

cela revient à considérer l'opérateur $\bar{L}_h v^h(x) = L_h v(x) - L_h \bar{\Phi}$, avec $\bar{\Phi}(x) = \mathbb{I}_{\{x \in \partial^+\Omega_h\}} \Phi(x)$. $\bar{L}_h v(x)$ pour $x \in \Omega_h$ ne fait intervenir que les valeurs de v dans Ω_h . On notera que si $e(x) = 0$ sur $\partial^+\Omega_h$ alors $\bar{L}_h e = L_h e(x)$. On confondra dans la suite l'opérateur \bar{L}_h et la matrice associée.

On notera que les points de $\partial^+\Omega_h$ ne sont pas forcément sur $\partial\Omega$, il faut donc pouvoir prolonger la valeur de Φ sur un voisinage du bord de Ω et supposer h suffisamment petit pour que les points de $\partial^+\Omega_h$ restent dans ce voisinage, cette opération exigera une certaine régularité de Φ et de $\partial\Omega$.

Le problème de Neumann

$$Lv(x) + c(x) = 0 \text{ pour } x \in \Omega \text{ et } \frac{\partial v}{\partial n} = \Phi \text{ pour } x \in \partial\Omega \quad (4.14)$$

peut par exemple conduire au problème discret :

$$\begin{cases} L_h v^h + c_h = 0 & \text{pour } x \in \Omega_h \cup \partial^+\Omega_h \\ B_h v^h(x) + \Phi^h(x) = 0 & \text{pour } x \in \partial^+\Omega_h \end{cases} \quad (4.15)$$

où ici B_h désigne l'opérateur discret utilisé pour la condition de Neumann (par exemple $B_h = -\partial_n^h$ (4.7)). On notera que la discrétisation des conditions au bord avec des schémas symétriques (∂_n^h) fait intervenir des points fictifs qui sont dans $\partial^+(\Omega_h \cup \partial^+\Omega_h)$. En général on pourra éliminer les points fictifs en trouvant une combinaison linéaire des opérateurs $L_h v^h + \alpha(h)B_h v$ qui n'utilise plus que les valeurs de v dans $\Omega_h \cup \partial^+\Omega_h$. Il faudra donc résoudre le système linéaire à N_h équations et N_h inconnues ou cette fois $N_h = \text{card}(\Omega_h \cup \partial^+\Omega_h)$:

$$\begin{cases} \bar{L}^h v^h + \bar{c}^h = 0 & \text{pour } x \in \Omega_h \cup \partial^+\Omega_h \\ \bar{L}^h = L_h \text{ et } \bar{c}^h = c_h & \text{sur } \Omega_h \\ \bar{L}^h = L_h + \alpha(h)B_h \text{ et } \bar{c}^h = c_h + \alpha(h)\Phi^h & \text{sur } \partial^+\Omega_h \end{cases} \quad (4.16)$$

À nouveau, il faut pouvoir éventuellement prolonger Φ sur un voisinage du bord de Ω . On associe aussi à l'opérateur L l'opérateur \bar{L} qui suit la définition de $\bar{L}_h : \bar{L} = L$ dans Ω et $\bar{L}v = Lv + \alpha(x)(-\frac{\partial v}{\partial n})$ sur $\partial\Omega$. Un exemple de calcul de $\alpha(x)$ est donné dans l'exercice 4.2.1.

4.2.3 Conditions de stabilité de la méthode d'approximation

Exprimons à présent les conditions de consistance et de stabilité de la méthode d'approximation qui assurent que le problème discret (4.15 ou 4.13) a une unique solution v_h et que cette solution converge vers la solution v du problème continu aux points du maillage lorsque le pas de discrétisation h tend vers 0.

Considérons ici pour fixer les idées le problème continu de Dirichlet (4.12) et le schéma numérique aux différences finies (4.13).

Definition 4.2.1. *Le schéma numérique (4.13) est dit consistant avec l'équation (4.12) si pour toute fonction C^∞ donnée v on a :*

$$|Lv(x) - L_h v(x)| \leq C|h|^p \quad \text{pour } x \in \Omega_h \quad (4.17)$$

où $p > 0$. p est l'ordre du schéma numérique et C est une constante qui ne dépend ni de h ni de x .

Cette propriété traduit l'idée naturelle que pour une fonction régulière l'équation discrétisée est bien une approximation de l'équation initiale sur l'ensemble discret Ω_h . En général les solutions de l'équation (4.12) ne seront pas C^∞ et il faudra établir des majorations (4.17) pour des fonctions moins régulières. L'ordre du schéma indique donc la borne que l'on peut espérer dans le meilleur des cas.

Dans le cas de conditions de Neumann il faudra regarder la consistance de \bar{L}_h avec \bar{L} pour $x \in \Omega_h \cup \partial^+ \Omega_h$ (4.16), nous illustrons ce cas dans l'exercice 4.2.1. Les formules de Taylor sont les clefs des démonstrations de consistance et on pourra par exemple vérifier à titre d'exercice que l'approximation (4.4) est consistante avec $Lv = \partial_x^2 v$ d'ordre 2 (Les approximations (4.5 et 4.6) sont d'ordre 1 seulement). Les approximations des dérivées secondes (4.8, 4.9 et 4.10) s'obtiennent de façon analogue par combinaisons linéaires de développements de Taylor à l'ordre 4, et sont d'ordre 2 (Voir exercice 4.2.2).

Exercice 4.2.1. *Soient $d = 1$ et $\Omega = (0, 1)$, on considère l'opérateur $L = a(x)v''(x) - \lambda v(x)$ et sa version discrétisée obtenue avec le schéma (4.8). Au bord $\{x = 0\}$ on considère la condition de Neumann $\frac{\partial v}{\partial n} = \Phi$.*

- *On discrétise la condition au bord avec le schéma (4.4). Trouver $\alpha(x)$ pour que l'opérateur \bar{L}_h (4.16) n'utilise que des valeurs de v en des points discrets de $[0, 1]$ et montrer qu'au point $x = 0$ pour une fonction v régulière :*

$$|\bar{L}_h v(0) - \bar{L}v(0)| \leq K|h|$$

- *On discrétise la condition au bord avec le schéma (4.5). Calculer le nouvel opérateur \bar{L} et montrer que cette fois pour une fonction v régulière solution de $Lv + c = 0$ on a la majoration :*

$$|\bar{L}_h v(0) - \bar{L}v(0) + \frac{1}{2}(c(0) + \lambda v(0))| \leq K|h|$$

On perd la propriété de consistance si $c(0) + \lambda v(0) \neq 0$.

Exercice 4.2.2. Soient $d = 1$ et $\Omega = (0, 1)$, On considère l'opérateur $Lv = v''(x)$ et le schéma (4.8). Montrer que pour $v \in C^4([0, 1])$:

$$|L_h v - Lv| \leq \frac{h^2}{12} \max_{[0,1]} |v^{(4)}(x)|$$

Montrer que si l'on suppose seulement $v \in C^{2,\alpha}$ on a alors :

$$|L_h v - Lv| \leq K|h|^\alpha$$

Definition 4.2.2. Le schéma numérique (4.13) est dit stable au sens de la norme l^∞ si la norme l^∞ de la matrice \bar{L}_h^{-1} est majorée indépendamment de h :

$$\|\bar{L}_h^{-1}\|_\infty \leq C.$$

Cela implique en particulier que la matrice $(\bar{L}_h)^{-1}$ est inversible.

Nous donnons tout d'abord une condition suffisante de stabilité :

Proposition 4.2.1. Le schéma est stable si la matrice \bar{L}_h est à diagonale fortement dominante de paramètre λ (definition 1.3.2) et on a l'estimation $\|\bar{L}_h^{-1}\|_\infty \leq 1/\lambda$.

Preuve : On cherche à montrer que $\|\bar{L}_h^{-1}\|_\infty \leq C$ où C ne dépend pas de h . Si \bar{L}_h est a diagonale fortement dominante alors (proposition 1.3.4) elle satisfait le *Principe du Maximum Discret strictement positif*. On sait qu'alors \bar{L}_h est inversible. Soit donc v^h solution de $\bar{L}_h v^h + c_h = 0$, En utilisant les propriétés de \bar{L}_h et en notant l_{ij} les éléments de \bar{L}_h on obtient :

$$-\sum_{j \neq i} l_{ij} \|v\|_\infty \leq -\sum_{j \neq i} l_{ij} v_j$$

d'où :

$$-\sum_j l_{ij} \|v\|_\infty + l_{ii} \|v\|_\infty - l_{ii} v_i \leq -\sum_j l_{ij} v_j \leq \|c\|_\infty$$

et donc :

$$\lambda \|v\|_\infty + l_{ii} (-v_i + \|v\|_\infty) \leq \|c\|_\infty$$

S'il existe un indice i_0 pour lequel $v_{i_0} = \|v\|_\infty$ on obtient $\|v\|_\infty \leq \frac{1}{\lambda} \|c\|_\infty$ sinon il existe un indice i_0 tel que $-v_{i_0} = \|v\|_\infty$ et en reprenant le raisonnement avec $w = -v$ on obtient à nouveau $\|v\|_\infty \leq \frac{1}{\lambda} \|c\|_\infty$. Ce qui termine la preuve. \square

Si \bar{L}_h est a diagonale fortement dominante nous verrons dans la suite que le problème discrétisé peut avoir une interprétation probabiliste ce qui nous donnera une autre preuve de la stabilité.

Remarque : La condition énoncée plus haut est non nécessaire. Le schéma (4.8) pour l'équation $u''(x) + c(x) = 0$ ($\Omega = (0, 1)$) avec condition de Dirichlet $u = 0$ est stable sans que la matrice associée soit à diagonale fortement dominante (On trouvera une démonstration de ce résultat dans [28] page 20–23). \square

Proposition 4.2.2. *Supposons le schéma (4.13) stable et consistant avec l'équation (4.12). Soient u et u_h les solutions respectives de (4.12) et (4.13). Supposons que la solution u ait une régularité suffisante pour que la majoration (4.17) ait lieu alors :*

$$\max_{x \in \Omega_h} |u(x) - u_h(x)| \leq C|h|^p.$$

C étant une constante indépendante de h .

Preuve : L'erreur $e_h = u(x) - u_h(x)$ vérifie $e_h(x) = 0$ sur $\partial^+ \Omega_h$. On a vu que dans ce cas $\bar{L}_h e = L_h e$, d'autre part $L_h e_h = L_h u - L_h u_h = L_h u - Lu + Lu - L_h u_h = L_h u - Lu$ ($Lu = L_h u_h = -v_h$ sur Ω_h). d'où en utilisant la consistance, $|\bar{L}_h e^h(x)| \leq C_1 |h|^p$ pour tout $x \in \Omega_h$ et en utilisant la stabilité $\|e_h\|_\infty \leq C_2 |h|^p$. \square

Remarque : On peut avoir convergence d'un schéma sans qu'il soit consistant. Un exemple est donné à nouveau par $u''(x) + c(x) = 0$ dans $\Omega = (0, 1)$ et condition mixte $u(0) = 0$ et $u'(1) = 0$. Le schéma n'est pas consistant si la condition de Neumann est discrétisée avec le schéma (4.5) mais il est convergent d'ordre 1 ($\|e_h\|_\infty \leq C|h|$) [28] page 29-33. Le manque de consistance est compensé par une majoration plus fine dans la propriété de stabilité. \square

Nous illustrons ici la remarque précédente en proposant un exercice. Nous verrons dans un paragraphe suivant que la discrétisation des conditions de Neumann conduit aux hypothèses faites dans cet exercice.

Exercice 4.2.3. *On considère le système linéaire sur $E = \mathbb{R}^n$*

$$L^h v^h + c^h = 0 \quad \text{avec} \quad L^h = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}$$

On appelle L_1^h la matrice (L_{11}, L_{12}) et L_2^h la matrice (L_{11}, L_{12}) et on suppose que les propriétés suivantes sont vérifiées $L_{i,j}^h \geq 0$ pour $i \neq j$, $\sum_j (L_1^h)_{i,j} \leq -\lambda < 0$, $\sum_j (L_2^h)_{i,j} = 0$ et enfin L_{22} est une matrice inversible et $L_{22}^{-1} \leq 0$.

- *Montrer que le système admet une unique solution qui vérifie $\|v^h\|_\infty \leq C_2 \|L_{22}^{-1}\|_\infty \|c_2^h\|_\infty + C_1 \|c_1^h\|_\infty$*
- *Soit v solution de $Lv + c^h = 0$ et supposons que $\|L_{22}^{-1}\|_\infty \leq Kh^\alpha$ en déduire que $e^h = v - v^h$ vérifie :*

$$\|e^h\|_\infty \leq C_2 h^\alpha \|L_2 v - L_2^h v\|_\infty + C_1 \|L_1 v - L_1^h v\|_\infty$$

Appliqué à un schéma de discrétisation on voit que la non consistance de L_2^h peut être compensée (ici par une propriété de L_{22}) pour assurer la convergence du schéma.

- *Montrer aussi la propriété suivante :*

$$\|e_1^h\|_\infty \leq \frac{1}{\lambda} \| -L_{12} L_{22}^{-1} d_2 + d_1 \|_\infty, \quad d_2 = L_2 v - L_2^h v \quad \text{et} \quad d_1 = L_1 v - L_1^h v$$

4.2.4 Discrétisation d'un opérateur elliptique sans dérivées croisées et uniformément elliptique.

Supposons qu'il n'y ait pas de dérivées croisées dans l'équation (4.2), i.e $a_{ij}(x) = 0$ pour $i \neq j$, que l'opérateur A est uniformément elliptique, i.e. $a_{ii}(x) \geq \alpha > 0$ et que les b_i sont bornés. Alors l'opérateur A_h obtenu en discrétisant les dérivées secondes par (4.8) et les dérivées premières par des schémas saute-mouton (4.4) satisfait les conditions de la définition 1.3.2 pour un pas de discrétisation h vérifiant (si le pas de discrétisation n'est pas le même dans toutes les directions h_i désignera le pas dans la direction i et h le vecteur (h_1, \dots, h_n)).

$$0 < h_i \leq \min_{x \in \Omega_h} \frac{2a_{ii}(x)}{|b_i(x)|}. \quad (4.18)$$

Nous développons les calculs dans le cas de la dimension 1 avec des conditions de Dirichlet ou de Neumann. Sous les mêmes conditions, nous montrons que le problème discrétisé admet une interprétation probabiliste ce qui permet une autre démonstration de la stabilité (4.25).

Conditions de Dirichlet

On considère tout d'abord le problème de Dirichlet en dimension 1 :

$$\begin{cases} Av(x) - \lambda v(x) + c(x) = 0 & \text{dans } \Omega = (0, 1) \\ Av \equiv a(x) \frac{d^2 v}{dx^2}(x) + b(x) \frac{dv}{dx}(x) \\ v(0) = \Phi(0) \quad \text{et} \quad v(1) = \Phi(1) \end{cases} \quad (4.19)$$

avec $a(x) \geq \alpha > 0 \forall x \in]0, 1[$.

Soit $h = \frac{1}{n}$ le pas de discrétisation. Il s'agit d'approcher la valeur de v aux points $x_i = ih$, $i = 0, \dots, n$. Utilisant les schémas aux différences finies (4.4,4.8), on introduit l'opérateur A_h définie pour tout x sur la grille $\Omega_h = \{ih, i = 1, \dots, n-1\}$ par :

$$A_h v \equiv \frac{a(x)}{h^2} (v(x+h) - 2v(x) + v(x-h)) + \frac{b(x)}{2h} (v(x+h) - v(x-h)) \quad (4.20)$$

Le problème (4.19) est alors remplacé par le problème discrétisé :

$$\begin{cases} A_h v^h(x) - \lambda v^h(x) + c_h(x) = 0 & \text{pour } x \in \Omega_h \\ v^h(x_0) = \Phi(x_0) \quad \text{et} \quad v^h(x_n) = \Phi(x_n) & (\partial^+ \Omega_h = \{x_0, x_n\}) \end{cases} \quad (4.21)$$

En introduisant les notations :

$$\mathcal{D}(x) \equiv -\frac{2a(x)}{h^2} \quad \text{et} \quad \mathcal{A}^\pm \equiv \frac{a(x)}{h^2} \pm \frac{b(x)}{2h}$$

L'opérateur A_h s'écrit :

$$A_h v(x_i) \equiv A_{(\mathcal{A}^-, \mathcal{D}, \mathcal{A}^+)}^h v(x_i) \equiv \mathcal{A}^-(x_i) v(x_{i-1}) + \mathcal{D}(x_i) v(x_i) + \mathcal{A}^+(x_i) v(x_{i+1})$$

Pour toute fonction $v(x)$ définie sur $\bar{\Omega} = [0, 1]$, on note v^h le vecteur colonne de composantes $v(x_i)_{i=0,n}$ et \bar{v}^h le vecteur colonne de composantes $v(x_i)_{i=1,n-1}$. L'équation (4.21) s'écrit sous forme Matricielle :

$$-\lambda \bar{v}^h + \mathcal{A}^h v^h + \bar{c}^h = 0 \quad \text{avec} \quad v_0^h = \Phi(x_0), v_n^h = \Phi(x_n)$$

avec :

$$\mathcal{A}^h \equiv \begin{pmatrix} \mathcal{A}^-(x_1) & D(x_1) & \mathcal{A}^+(x_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \mathcal{A}^-(x_{n-2}) & D(x_{n-2}) & \mathcal{A}^+(x_{n-2}) & 0 \\ 0 & \dots & 0 & \mathcal{A}^-(x_{n-1}) & D(x_{n-1}) & \mathcal{A}^+(x_{n-1}) \end{pmatrix}$$

La matrice \mathcal{A}^{h1} associée à l'opérateur A_h est de taille $(n-1 \times n+1)$ et on notera $\overline{\mathcal{A}}^h$ la matrice obtenue en retirant à \mathcal{A}_h sa première et sa dernière colonne.

On obtient alors :

$$-\lambda \overline{v}^h + \overline{\mathcal{A}}^h \overline{v}^h + \overline{d}^h = 0$$

avec cette fois : $\overline{d}^h = \overline{c}^h + \mathcal{A}^h(\Phi(x_0), 0, \dots, 0, \Phi(x_n))^T$. On doit donc numériquement résoudre le système linéaire dans \mathbb{R}^{n-1} donné par :

$$\overline{L}_h \overline{v}_h + \overline{d}_h = 0 \quad \text{avec} \quad \overline{L}_h \equiv -\lambda I + \overline{\mathcal{A}}^h \quad (4.22)$$

et la solution du problème (4.21) sera donnée par \overline{v} sur Ω_h et par $v(x) = \Phi(x)$ sur $\partial\Omega_h = \{0, 1\}$.

Nous laissons au lecteur le soin de vérifier la consistance du schéma proposé et regardons ici la stabilité.

Proposition 4.2.3. *Soit $h > 0$ tel que $\mathcal{A}^\pm(x) \geq 0$ pour tout $x \in \Omega_h$, alors \overline{L}^h est à diagonale fortement dominante, elle satisfait donc (proposition 1.3.4 page 14) le principe du maximum discret strictement positif.*

La condition $\mathcal{A}^\pm(x) \geq 0$ est satisfaite pour h vérifiant :

$$h \leq \min_{x \in \Omega_h} \frac{2a(x)}{|b(x)|} \quad \text{ou} \quad h_i \leq \min_{x \in \Omega_h} \frac{2a_{ii}(x)}{|b_i(x)|} \quad \text{en dimension } d \quad (4.23)$$

Preuve : Nous donnons la preuve dans le cas de la dimension 1. On remarque tout d'abord que $\mathcal{A}^-(x) + \mathcal{D}(x) + \mathcal{A}^+(x) = 0$ pour tout $x \in \Omega_h$. On a donc :

$$\sum_{j=1}^{n-1} \mathcal{A}_{i,j}^h = 0 \quad \text{pour} \quad i = 1, \dots, n-1.$$

Soit aussi pour la matrice \overline{L}^h :

$$\sum_{j=1}^{n-1} \overline{L}_{i,j}^h = -\lambda - \mathcal{A}^-(x_1)\delta_{i,1} - \mathcal{A}^+(x_{n-1})\delta_{i,n-1} \quad \text{pour} \quad i = 1, \dots, n-1.$$

L'hypothèse $\mathcal{A}^\pm(x) \geq 0$ assure alors qu'il existe $\lambda > 0$ tel que $\sum_{j=1}^{n-1} \overline{L}_{i,j}^h \leq -\lambda < 0$. D'autre part les termes hors diagonaux de la matrice A^h sont les $\mathcal{A}^\pm(x_i)$ qui sont par hypothèse positifs ou nuls. La

¹Dans la suite on utilisera parfois la notation $\mathcal{A}_{[x_i, x_j]}^h$ pour préciser les lignes de début et de fin de la matrice. Par exemple, avec cette notation la matrice \mathcal{A}^h décrite ici se noterait $\mathcal{A}_{[x_1, x_{n-1}]}^h$.

matrice \bar{L}^h est donc bien à diagonale fortement dominante. La deuxième partie de la proposition se vérifie aisément.

Le cas de la dimension d n'est pas très différent puisqu'en absence de terme de dérivé croisées $A_h v = \sum_{i=1}^d A_{h_i}^i v$ avec $A_{h_i}^i$ donnée par la formule (4.20) où a est remplacé par a_{ii} et b par b_i et nous laissons la preuve en exercice. \square

Considérons d'autre part pour une fonction $\Delta t^h(x)$ donnée sur Ω_h , la matrice \mathcal{M}^h associée à l'opérateur $M^h v = \Delta t(x) A_h v(x) + v(x)$:

$$\mathcal{M}^h \equiv \mathcal{A}_{(\Delta t^h \mathcal{A}^-, \Delta t^h \mathcal{D}+1, \Delta t^h \mathcal{A}^+)}^h$$

(Par abus de notation on écrira $\mathcal{M}^h = \Delta t^h \mathcal{A}^h + I$).

Le problème discrétisé (4.21) s'écrit donc aussi :

$$\begin{cases} v^h(x) = \frac{1}{1+\lambda \Delta t^h} (\mathcal{M}^h v^h(x) + \Delta t^h c(x)) & \text{pour } x \in \Omega_h \\ v^h(x) = \Phi(x) & \text{pour } x \in \partial^+ \Omega_h \end{cases} \quad (4.24)$$

Proposition 4.2.4. *Soit $h > 0$ tel que $\mathcal{A}^\pm(x) \geq 0$ pour tout $x \in \Omega_h$ et $\Delta t(x)$ une fonction sur Ω_h vérifiant :*

$$0 < \Delta t(x) \leq \frac{1}{2} \left(\sum_{i=1}^n a_{ii}(x)/h_i^2 \right)^{-1} \quad \left(\frac{\bar{h}^2}{2 \text{trace}(a)} \quad \text{si } h_i = \bar{h} \quad \forall i \right)$$

la matrice \mathcal{M}^h est une matrice stochastique². La solution du problème (4.24) s'interprète alors comme la fonction valeur du problème suivant :

$$v^h(x) = \mathbb{E}_{x,0} \left[\sum_{n=0}^{\nu_{\Omega_h}-1} \left(\prod_{l=0}^{n-1} \psi(X^l) \right) \Delta t^h(X^n) c(X^n) + \left(\prod_{l=0}^{\nu_{\Omega_h}-1} \psi(X^l) \right) \Phi(X^{\nu_{\Omega_h}}) \right]$$

avec

$$\psi(x) \equiv \frac{1}{1 + \lambda \Delta t^h(x)}$$

où X^n est une chaîne de Markov sur $\mathcal{E} = \Omega_h \cup \partial \Omega_h$ de matrice de transition M^h qui coïncide avec \mathcal{M}^h sur Ω_h (quelconque ailleurs) et ν_{Ω_h} est le temps de sortie de Ω_h .

Preuve : On a vu dans la proposition précédente que $\sum_{j=1}^{n-1} \mathcal{A}_{i,j}^h = 0$ on a donc $\sum_{j=1}^{n-1} \mathcal{M}_{i,j}^h = 1$. Les termes hors diagonaux de la matrice \mathcal{M}^h sont $\Delta t^h(x) \mathcal{A}^\pm(x)$ qui sont positifs par hypothèse. Les termes diagonaux sont $1 - \Delta t^h(x) 2a(x)/h^2$ et l'hypothèse sur $\Delta t^h(x)$ nous assure qu'ils sont positifs. \mathcal{M}^h est bien une matrice stochastique. Si $\Delta t^h(x)$ ne dépend pas de x la fin de la proposition résulte de la proposition 1.19 page 11 sinon la preuve suit cette même proposition et nous la laissons en exercice au lecteur. Le cas de la dimension d se montre tout aussi simplement les termes diagonaux s'écrivant alors $1 - \Delta t^h(x) (\sum_{i=1}^d 2a_{ii}(x)/h_i^2)$. \square

²On notera que \mathcal{M}^h n'est pas une matrice carrée ($(n-1) \times (n-1)$) on veut donc dire par là que c'est une matrice extraite d'une matrice stochastique

Nous montrons maintenant que l'interprétation probabiliste du problème discret permet une majoration directe de l'erreur $e_h(x) = v(x) - v^h(x)$ sur Ω_h . Nous raisonnons ici comme dans en donnant une interprétation probabiliste de l'erreur.

Soit h vérifiant la condition (4.23) de la Proposition 4.2.3. On suppose ici que $\text{trace}(a(x))$ est majorée sur Ω et on peut alors choisir :

$$\Delta t^h(x) = h_m^2 / (2\text{trace}(a(x))) \quad \text{avec} \quad h_m = \min_{i \in [1, d]} h_i$$

qui vérifie les hypothèses de la Proposition 4.2.4. Il existe alors α_1 et α_2 , $0 < \alpha_1 \leq \alpha_2 < \infty$, tel que l'on ait $\alpha_1 h^2 \leq \Delta t^h(x) \leq \alpha_2 h^2$. Soit v la solution du problème continu (4.19), (ou de sa version d -dimensionnelle) un calcul immédiat montre que v vérifie la même équation que v^h en remplaçant $c(x)$ par $c + Av - A_h v$, on a donc :

$$\begin{cases} v(x) = \frac{1}{1+\lambda\Delta t^h} (\mathcal{M}^h v(x) + \Delta t^h (c(x) + Av - A_h v)) & \text{pour } x \in \Omega_h \\ v(x) = \Phi(x) & \text{pour } x \in \partial\Omega_h \end{cases}$$

La différence $e_h = v - v^h$ est donc solution de :

$$\begin{cases} e_h(x) = \frac{1}{1+\lambda\Delta t^h} (\mathcal{M}^h e_h(x) + \Delta t^h (Av - A_h v)) & \text{pour } x \in \Omega_h \\ e_h(x) = 0 & \text{pour } x \in \partial\Omega_h \end{cases}$$

On peut à nouveau utiliser la proposition 4.2.4 pour construire une interprétation probabiliste pour $e_h(x)$ sur laquelle on obtient facilement la majoration :

$$|e_h(x)| \leq \sum_{k=0}^{\infty} \frac{\alpha_2 h_m^2}{(1 + \lambda \alpha_1 h_m^2)^k} \|Av - A_h v\|_{\infty, \Omega_h} \quad (4.25)$$

Soit aussi : $|e_h(x)| \leq C(\alpha_2/(\lambda\alpha_1)) \|Av - A_h v\|_{\infty, \Omega_h}$. Cette propriété est l'analogie de la condition de stabilité. Associée à une condition de consistance pour L_h ou pour A_h ($\|Av - A_h v\|_{\infty} = \|Lv - L_h v\|_{\infty}$) elle nous permet d'obtenir la convergence du schéma (On notera qu'il faut avoir une estimation de la régularité de la solution v pour utiliser la condition de consistance associée).

Conditions de Neumann

Considérons l'équation (4.19) mais avec la condition de Neumann $\frac{\partial v(x)}{\partial n} = \Phi(x)$ pour $x \in \partial\Omega_h$. Nous suivons ici la méthode générale qui a été décrite au paragraphe 4.2.2, on a ici

$$\Omega_h = \{x_i = ih, i \in [1, n-1]\} \quad \text{et} \quad \partial^+ \Omega_h = \{x_0 = 0, x_n = 1\}.$$

Le problème discrétisé, en utilisant l'approximation (4.8) en tout point de $\Omega_h \cup \partial^+ \Omega_h$ et l'approximation (4.4) en tout point de $\partial^+ \Omega_h$, s'écrit :

$$\begin{cases} -\lambda v^h(x) + A_h v^h(x) + c_h(x) = 0 & \text{pour } x \in \Omega_h \cup \partial^+ \Omega_h \\ B_h v(x_n) + \Phi(x_n) = B_h v(x_0) + \Phi(x_0) = 0 \\ B_h v(x_n) \equiv -(v^h(x_{n+1}) - v^h(x_{n-1}))/2h, \\ B_h v(x_0) = -(v^h(x_{-1}) - v^h(x_1))/2h \end{cases} \quad (4.26)$$

avec A_h défini par l'équation (4.20). On rappelle, comme il a été vu dans le cas général, que l'on introduit dans l'écriture précédente la valeur de v^j en deux points fictifs x_{n+1} et x_{-1} . On peut alors éliminer les variables fictives $v^h(x_{n+1})$ et $v^h(x_{-1})$ des équations précédentes en utilisant $\bar{A}_h = A_h + \alpha_h(x)B_h(x)$ avec

$$\alpha_h(x) = 2h(\mathcal{A}^-(x_0)\delta_{x,0} + \mathcal{A}^+(x_n)\delta_{x,x_n})$$

pour obtenir :

$$\left\{ \begin{array}{l} -\lambda v^h(x) + \bar{A}_h v^h(x) + d^h(x) = 0 \quad \text{pour } x \in \Omega_h \cup \partial^+ \Omega_h \\ \bar{A}_h(x) = A_h(x) \quad \text{pour } x \in \Omega_h \\ \bar{A}_h v^h(x_0) = \mathcal{D}(x_0)v^h(x_0) - \mathcal{D}(x_0)v^h(x_1) \\ \bar{A}_h v^h(x_n) = -\mathcal{D}(x_n)v^h(x_{n-1}) + \mathcal{D}(x_n)v^h(x_n) \\ d^h(x) = c_h(x) + 2h\mathcal{A}^-(x_0)\Phi(x_0)\delta_{x,0} + 2h\mathcal{A}^+(x_n)\Phi(x_n)\delta_{x,x_n} \end{array} \right.$$

On a à présent $n+1$ inconnues : $(v_h(ih))_{i=0,\dots,n}$. Notons par \bar{A}_h la matrice associée à l'opérateur \bar{A}_h , c'est maintenant une matrice carrée $(n+1) \times (n+1)$ qui s'écrit :

$$\bar{A}_h \equiv \left(\begin{array}{cccccc} \mathcal{D}(x_0) & -\mathcal{D}(x_0) & 0 & \dots & 0 & \\ \hline & & \mathcal{A}_{[x_1, x_{n-1}]}^h & & & \\ \hline 0 & \dots & 0 & -\mathcal{D}(x_n) & \mathcal{D}(x_n) & \end{array} \right)$$

Comme $-\mathcal{D}(x) \geq 0$, il est facile de vérifier que la proposition 4.2.3 est encore vérifiée sous la même condition (4.23), $\bar{L}_h \equiv -\lambda I + \bar{A}_h$ est à diagonale fortement dominante sous la condition (4.23).

Soit $\mathcal{M}_h = \Delta t^h \bar{A}_h + I$. Le problème discret précédent s'écrit :

$$\left\{ \begin{array}{l} v^h(x) = \frac{1}{1+\lambda\Delta t^h} (\mathcal{M}_h v^h(x) + \Delta t^h d(x)) \quad \text{pour } x \in \Omega_h \cup \partial^+ \Omega_h \end{array} \right. \quad (4.27)$$

L'analogue de la proposition 4.2.4 devient :

Proposition 4.2.5. *Soit $h > 0$ tel que $\mathcal{A}^\pm(x) \geq 0$ pour tout $x \in \Omega^h$ et $\Delta t(x)$ une fonction sur $\bar{\Omega}^h$ vérifiant*

$$0 < \Delta t(x) \leq \frac{1}{2} \left(\sum_{i=1}^d a_{ii}(x)/h_i^2 \right)^{-1} \quad \left(\frac{\bar{h}^2}{2\text{trace}(a)} \quad \text{si } h_i = \bar{h} \quad \forall i \right)$$

la matrice \mathcal{M}_h est une matrice stochastique. La solution du problème (4.27) s'interprète alors comme la fonction valeur du problème suivant :

$$v(x) = \mathbb{E}_{x,0} \left[\sum_{n=0}^{\infty} \prod_{l=0}^{n-1} \frac{1}{1 + \lambda\Delta t^h(X^l)} \Delta t^h(X^n) d(X^n) \right]$$

où X^n est une chaîne de Markov sur $\mathcal{E} = \Omega_h \cup \partial\Omega_h$ de matrice de transition \mathcal{M}_h .

Preuve :

La preuve simple est laissée en exercice (on utilisera les résultats de la proposition 1.2.3 page 10).

□

Il n'est pas nécessaire pour obtenir une interprétation probabiliste d'éliminer $v^h(x_{n+1})$ et $v^h(x_{-1})$ on peut en effet considérer une chaîne de Markov sur $\Omega^h \cup \partial^+ \Omega_h \cup \{x_{-1}, x_{n+1}\}$ (cette approche étant plus simple en dimension supérieure à 1). On considère alors directement le système linéaire (4.26) à $n + 3$ inconnues. On notera que la matrice de ce système linéaire n'est plus à diagonale fortement dominante, mais qu'en considérant la partition de $\mathcal{E} = (\Omega_h \cup \partial^+ \Omega_h) + \{x_{-1}, x_{n+1}\}$ le système linéaire (4.26) a la même forme que dans l'exercice 4.2.3 avec $L_{22} = -I/2h$. Il vérifie donc bien les hypothèses de l'exercice 4.2.3

Soit $\Delta t^h(x)$ telle que $\Delta t^h(x_{-1}) = \Delta t^h(x_{n+1}) = 0$ et $\Delta t^h(x) \neq 0$ sinon. La résolution de l'équation (4.26) est équivalente à la résolution du système :

$$\begin{cases} v^h(x) = \frac{1}{1+\lambda\Delta t^h(x)} (M^h v^h(x) + c_h(x)) & \text{pour } x \in \Omega^h \cup \partial^+ \Omega_h \\ v^h(x_{n+1}) = \frac{1}{1+\lambda\Delta t^h(x_{n+1})} (v^h(x_{n-1}) + 2h\Phi(x_n)) \\ v^h(x_{-1}) = \frac{1}{1+\lambda\Delta t^h(x_{-1})} (v^h(x_1) + 2h\Phi(x_0)) \end{cases}$$

Qui s'interprète cette fois comme la fonction valeur du problème suivant :

$$v(x) = \mathbb{E}_{x,0} \left[\sum_{n=0}^{\infty} \prod_{l=0}^{n-1} \frac{1}{1 + \lambda\Delta t^h(X^l)} g(X_n) \right]$$

avec :

$$g(X_n) = \Delta t^h(X^n) c(X^n) + 2h\Phi(x_0) \mathbb{I}_{X_n=x_{-1}} + 2h\Phi(x_n) \mathbb{I}_{X_n=x_{n+1}}$$

On notera que dans cette interprétation de $v(x)$ le taux d'actualisation est nul sur $\{x_{-1}, x_{n+1}\}$. Toutefois l'expression pour v a un sens car l'ensemble $\{x_{-1}, x_{n+1}\}$ est transiet. Cela est une conséquence, en reprenant les notations de l'exercice (4.2.3), de la propriété L_{22} inversible (voir Proposition 1.7.1). Le fait d'introduire des points fictifs que l'on peut éliminer dans le calcul revient d'un point de vue probabiliste à introduire ces points dans la classe transitoire.

En revenant à la formulation (4.26) nous montrons maintenant que l'interprétation probabiliste du problème discret permet une majoration directe de l'erreur $e_h(x) = v(x) - v^h(x)$ sur $\Omega_h \cup \partial^+ \Omega_h$. Nous raisonnons ici comme dans le cas des conditions de Dirichlet.

Soit v la solution du problème continu (4.19) avec condition de Neumann, un calcul immédiat montre que v vérifie la même équation que v^h (4.27) à condition de remplacer c par $c + Av - A_h v$ et Φ par $\Phi + Bv - B_h v$ ($Bv \equiv -\frac{\partial v}{\partial n}$). l'erreur $e(x) = v(x) - v_h(x)$ vérifie donc aussi la même équation (4.27) avec une fonction $d_e(x)$ qui vaut :

$$\begin{aligned} d_e(x) &= Av - A_h(v) + 2h\mathcal{A}^-(x_0)(Bv - B_h v)(x_0)\delta_{x,x_0} \\ &+ 2h\mathcal{A}^+(x_n)(Bv - B_h v)(x_n)\delta_{x,x_n} \\ &= Av + \alpha(x)Bv - A_h v - \alpha(x)B_h v(x) \\ &= \bar{L}v - \bar{L}_h v \end{aligned}$$

En utilisant l'interprétation probabiliste de e et en raisonnant comme dans le paragraphe précédent on obtient pour $x \in \Omega_h \cup \partial^+ \Omega_h$:

$$|e(x)| \leq C \|\bar{L}v - \bar{L}_h v\|_{\infty, \Omega_h \cup \partial^+ \Omega_h}$$

Cette propriété (équivalente à nouveau à une propriété de stabilité) associée à une condition de consistance pour \bar{L}_h nous permet d'obtenir la convergence du schéma.

4.2.5 Discrétisation d'un opérateur elliptique sans dérivées croisées et dégénéré.

Lorsque l'opérateur A est dégénéré, c'est à dire quand la matrice $a(x)$ n'est plus définie positive en tout point, ou lorsque les conditions (4.18) conduisent à un pas h jugé trop petit, on peut encore obtenir une discrétisation vérifiant les conditions (1.3.2) et donc le Principe du Maximum Discret strictement positif, *mais d'un ordre de convergence inférieur*, en utilisant des approximations décentrées du gradient : on utilise l'approximation (4.5) si $b_i(x) \geq 0$ et (4.6) si $b_i(x) < 0$, c'est à dire qu'il faut bien prendre en compte la direction du terme de dérive.

On peut également rajouter un terme de viscosité, c'est à dire un terme du deuxième ordre tendant vers 0 avec h , de façon à utiliser les schémas symétriques d'approximation du gradient. (voir [15]).

Supposons qu'il existe $x \in \Omega_h$ tel que $a(x) = 0$, la condition (4.23) ne peut pas être satisfaite. On utilise alors un schéma décentré pour l'approximation de la dérivée.

Soit g une fonction sur Ω , on définit $g_+(x) \equiv \max(0, g(x))$ et $g_- \equiv -\min(g(y), 0)$. On considère à nouveau le problème discrétisé (4.21) mais avec cette fois :

$$\begin{aligned} A_h v &\equiv \frac{a(x)}{h^2} (v(x+h) - 2v(x) + v(x-h)) + \frac{b_+(x)}{h} (v(x+h) - v(x)) \\ &+ \frac{b_-(x)}{h} (v(x-h) - v(x)) \end{aligned}$$

L'opérateur A_h s'écrit à nouveau :

$$A_h v(x_i) \equiv A_{(\mathcal{A}^-, \mathcal{D}, \mathcal{A}^+)}^h v(x_i) \equiv \mathcal{A}^-(x_i) v(x_{i-1}) + \mathcal{D}(x_i) v(x_i) + \mathcal{A}^+(x_i) v(x_{i+1})$$

avec les nouvelles définitions :

$$\mathcal{D}(x) \equiv -\frac{2a(x)}{h^2} - \frac{|b(x)|}{h} \quad \text{et} \quad \mathcal{A}^\pm \equiv \frac{a(x)}{h^2} + \frac{b_\pm(x)}{h}$$

et en remarquant que $|b(x)| = b_+(x) + b_-(x)$.

On notera que $\mathcal{A}^\pm \geq 0$, la proposition 4.2.3 est donc vérifiée sans condition sur h .

Pour assurer l'analogie de la proposition 4.2.4 il faut une condition sur $\Delta t^h(x)$ qui assure $\Delta t^h(x) \mathcal{D}(x) + 1 \geq 0$. Il suffit de choisir :

$$0 < \Delta t^h(x) \leq \frac{h^2}{2a(x) + h|b(x)|}$$

($\Delta t^h(x)$ pouvant être choisit arbitrairement la ou $2a(x) + h|b(x)| = 0$). Le cas de la dimension d se calcule tout aussi simplement et conduit au choix :

$$0 < \Delta t^h(x) \leq \left(\sum_{i=1}^d 2 \frac{a_{ii}(x)}{h_i^2} + \frac{|b_i(x)|}{h_i} \right)^{-1}$$

4.2.6 Discrétisation d'un opérateur elliptique avec dérivées croisées.

Supposons que la matrice a est non diagonale. Si la condition

$$a_{ii}(x) \geq \sum_{j,j \neq i} |a_{ij}(x)| \quad (4.28)$$

est satisfaite pour $i = 1 \dots d$, alors l'opérateur A_h obtenu en utilisant les discrétisations (4.9) et (4.10) pour les dérivées croisées, satisfait le Principe du Maximum Discret strictement positif. (pour h assez petit). On suppose ici $a(x)$ symétrique.

Considérons ici le cas de la dimension 2

$$Av = \sum_{i,j=1}^2 a_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j}$$

Supposons que la condition (4.28) est satisfaite, c'est à dire

$$a_{11}(x) \text{ et } a_{22}(x) \geq |a_{12}(x)|, \quad \forall x \in \Omega.$$

On peut alors écrire :

$$\begin{aligned} Av &= (a_{11} - |a_{12}|) \frac{\partial^2 v}{\partial x_1^2} + (a_{22} - |a_{12}|) \frac{\partial^2 v}{\partial x_2^2} \\ &+ |a_{12}| \left(\frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + 2 \text{sign}(a_{12}) \frac{\partial^2 v}{\partial x_1 \partial x_2} \right) \end{aligned} \quad (4.29)$$

On remarque que

$$\frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + 2 \text{sign}(a_{12}) \frac{\partial^2 v}{\partial x_1 \partial x_2}$$

est la dérivée seconde de v dans une des deux directions diagonales.

Supposons pour fixer les idées que $a_{12} > 0$ et considérons le changement de coordonnées :

$$y_1 = \frac{x_1 + x_2}{\sqrt{2}}, \quad y_2 = \frac{x_2 - x_1}{\sqrt{2}}.$$

On a

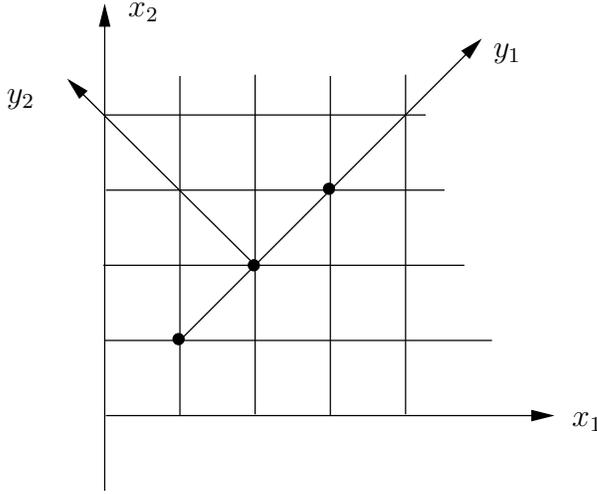
$$\frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + 2 \frac{\partial^2 v}{\partial x_1 \partial x_2} = 2 \frac{\partial^2 v}{\partial y_1^2}$$

On peut discrétiser le terme $2 \frac{\partial^2 v}{\partial y_1^2}$ en utilisant un schéma centré (4.8) dans les nouvelles coordonnées, à savoir :

$$\frac{\partial^2 v}{\partial y_1^2} \simeq \frac{1}{2h^2} (v(x_1 + h, x_2 + h) + v(x_1 - h, x_2 - h) - 2v(x_1, x_2)). \quad (4.30)$$

Il est alors facile de montrer que discrétiser Av en utilisant les schémas (4.8,4.9,4.10) donne la même expression que de discrétiser le membre droit de l'équation () en utilisant le schéma (4.8) pour $\frac{\partial^2 v}{\partial x_i^2}$ et le schéma centré dans les nouvelles coordonnées (4.30) pour le terme $\frac{\partial^2 v}{\partial y_1^2}$.

On vérifie que cette discrétisation vérifie le Principe du Maximum discret strictement positif. Si on rajoute à l'opérateur Av un terme du premier ordre, on utilisera une discrétisation symétrique du gradient dans le cas où $a_{ii} > |a_{12}|$ et décentrée si $a_{ii} = |a_{12}|$.



La condition (4.28) peut être relaxée (voit Kushner [10]) en considérant des pas de discrétisation différents selon les directions. Cette technique nécessite plus de travail dans la programmation de l'algorithme que pour des pas constants, mais elle est très utilisée.

Nous montrons ici sous quelles conditions la discrétisation de l'équation 4.2 conduit à un opérateur discret à diagonale fortement dominante dans le cas de la dimension d .

Lemme 4.2.1. *L'opérateur A_h obtenu en discrétisant l'opérateur A de l'équation (4.2) en utilisant les discrétisations (4.4, 4.8, 4.9, 4.10) et en remplaçant he_i par $h_i e_i$ et les termes h^2 par $h_i h_j$ s'écrit :*

$$\begin{aligned} A_h v(x) &= v(x) \left\{ \sum_i \frac{-2a_{ii}(x)}{h_i^2} + \sum_{j \neq i} \frac{|a_{ij}(x)|}{h_i h_j} \right\} \\ &+ \sum_{i,k=\pm 1} v(x + kh_i e_i) \left\{ \frac{a_{ii}(x)}{h_i^2} - \sum_{j,j \neq i} \frac{|a_{ij}|}{h_i h_j} + k \frac{b_i(x)}{2h_i} \right\} \\ &+ \sum_{i \neq j, k=\pm 1, l=\pm 1} v(x + ke_i h_i + le_j h_j) \frac{a_{ij}^{[kl]}}{h_i h_j} \end{aligned}$$

Preuve : En combinant les deux expressions (4.9) et (4.10) et en utilisant la notation :

$$a_{ij}^{[kl]}(x) = \begin{cases} a_{ij}^+(x) \equiv \max(0, a_{ij}(x)) & \text{pour } k.l = 1 \\ a_{ij}^-(x) \equiv -\min(0, a_{ij}(x)) & \text{pour } k.l = -1 \end{cases} ,$$

le terme $a_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j}$ s'écrit aussi :

$$\begin{aligned} a_{ij} \frac{\partial^2 v}{\partial x_i \partial x_j} &= \frac{|a_{ij}|}{2h_i h_j} \left(2v(x) - \sum_{k=\pm 1, l \in \{i,j\}} v(x + ke_l) \right) \\ &+ \frac{1}{2h_i h_j} \sum_{k=\pm 1, l=\pm 1} a_{ij}^{[kl]} v(x + ke_i h_i + lh_j e_j) \end{aligned}$$

on conclut alors facilement en utilisant (4.4) et (4.8). \square

Pour x fixé, on partitionne l'ensemble des indices en $I_1 + I_2$ tels que :

$$\begin{aligned} a_{ii}(x) - \sum_{j,j \neq i} |a_{ij}(x)| &> 0 \quad \text{pour } i \in I_1 \\ a_{ii}(x) - \sum_{j,j \neq i} |a_{ij}(x)| &\leq 0 \quad \text{pour } i \in I_2 \end{aligned}$$

Soient d'autre part $A_i \equiv \sum_{j \in I_1, j \neq i} |a_{ij}(x)|$ et $B_i \equiv \sum_{j \in I_2, j \neq i} |a_{ij}(x)|$

Lemme 4.2.2. *Sous les hypothèses suivantes supposées vérifiées pour tout $x \in \Omega_h$:*

$$(a_{ii} - B_i) > 0 \quad \text{pour } i \in I_2 \quad (4.31)$$

$$\underline{\rho} \equiv \max_{i \in I_2} \frac{-(a_{ii} - A_i - B_i)}{a_{ii} - B_i} < \bar{\rho} \equiv \min_{i \in I_1} \frac{a_{ii} - A_i - B_i}{B_i} \quad (4.32)$$

et le choix des h_i suivants :

$$h_i = \begin{cases} \alpha > 0 & \text{pour } i \in I_1 \\ \beta > 0 & \text{pour } i \in I_2 \end{cases}$$

avec $\alpha/\beta = 1 + \rho$ et ρ pris dans l'intervalle $(\underline{\rho}, \bar{\rho})$, on a pour tout i :

$$\hat{a}_i i(x) \equiv \frac{a_{ii}(x)}{h_i^2} - \sum_{j,j \neq i} \frac{|a_{ij}|}{h_i h_j} > 0 \quad (4.33)$$

Preuve : Avec les notations précédentes, on obtient facilement

$$\begin{aligned} \alpha^2 \hat{a}_i i(x) &= a_{ii}(x) - A_i - B_i - \rho B_i \quad \text{pour } i \in I_1 \\ \beta^2 \hat{a}_i i(x) &= a_{ii}(x) - A_i - B_i + \rho(a_{ii} - B_i) \quad \text{pour } i \in I_2 \end{aligned}$$

On vérifie alors facilement que les hypothèses (4.31) et (4.32) assurent $\hat{a}_i i(x) > 0$. \square

Remarque :

- si $B_i = 0$ alors la contrainte de droite disparaît ($\bar{\rho} = \infty$).
- Dans le cas particulier $I_2 = \emptyset$ alors $\rho = 0$ convient. On peut choisir le même h dans toutes les direction de l'espace. On retrouve les résultats du paragraphe précédent.
- On notera aussi que $I_1 = \emptyset$ est incompatible avec l'hypothèse (4.31).

\square

Proposition 4.2.6. *Sous les hypothèses (4.31) et (4.32) du lemme 4.2.2, on peut trouver des pas α (pour $i \in I_1$) et β (pour $i \in I_2$) tels que l'opérateur $L_h v = -\lambda v + A_h v$ défini au lemme 4.2.1 soit à diagonale fortement dominante.*

Preuve : En utilisant les équations de départ on voit facilement que $\sum_j A_{i,j}^h = 0$. Pour assurer que les éléments hors diagonaux de A_h sont positifs il faut avoir :

$$\hat{a}_{ij}(x) \pm \frac{b_i(x)}{2h_i} \geq 0$$

Sous les hypothèses du lemme (4.2.2) et en choisissant ρ dans l'intervalle $(\rho, \bar{\rho})$ les éléments $\hat{a}_{ij}(x)$ sont strictement positifs. Sur I_1 $\alpha^2 \hat{a}_i(x) = \mu_i(x, \rho)$ et sur I_2 $\beta^2 \hat{a}_i(x) = \mu_i(x, \rho)$. ρ étant fixé on choisit β :

$$0 < \beta < \min \left(\min_{i \in I_2} \frac{\mu_i(x, \rho)}{|b_i(x)|}, \min_{i \in I_1} \frac{(1 + \rho)\mu_i(x, \rho)}{|b_i(x)|} \right)$$

puis $\alpha = \beta(1 + \rho)$ pour assurer la positivité des éléments hors diagonaux. \square

Exercice 4.2.4. *Considérons l'équation (4.2) sur $\Omega = (0, 1)^2$ avec conditions de Dirichlet au bord. On suppose que a est donnée par $a_{11}(x) = 1$, $a_{12}(x) = a_{21}(x) = 2$, $a_{22}(x) = 5$. Vérifier que la condition (4.28) n'est pas satisfaite. Traduire les résultats de la proposition précédente sur cet exemple.*

4.2.7 Interprétation probabiliste d'un problème discret

On considère ici le problème d'algèbre linéaire $Lv + c = 0$, où L est une matrice $n \times n$ qui est supposée à diagonale fortement dominante (de paramètre λ). Nous montrons que la solution de ce système linéaire s'interprète comme la solution d'une équation de Kolmogorov.

On peut trouver $k > 0$ tel que la matrice :

$$M = I_n + k(L + \lambda I_n)$$

soit sous-Markovienne. Tout d'abord, L vérifie la propriété $L_{ii} + \lambda \leq 0$ pour tout $i \in [1, n]$. En effet, pour tout i on a $\sum_j L_{ij} \leq -\lambda \leq 0$ d'où $L_{ii} + \lambda \leq -\sum_{j \neq i} L_{ij}$ et donc le résultat puisque $L_{ij} \geq 0$ pour $i \neq j$.

Soit donc k tel que $0 < k \leq -1/\beta$ avec $\beta = \min_i L_{ii} + \lambda$ (on notera que si $\beta = 0$, tous les $k > 0$ conviennent). Les éléments hors diagonaux de M , $M_{ij} = kL_{ij}$, sont positifs. Les propriétés de L et le choix de $0 < k \leq -1/\beta$ assure la positivité des éléments diagonaux. D'autre part :

$$\sum_j M_{ij} = 1 + k(\lambda + \sum_j L_{ij}) \leq 1$$

La matrice M est donc sous-Markovienne ou Markovienne si $\rho_i \equiv \lambda + \sum_j L_{ij} = 0$ pour tout $i \in [1, n]$. Par contre, la matrice :

$$\bar{M} = \begin{pmatrix} 1 & 0 \\ -k\rho & M \end{pmatrix}$$

où $\rho = (\rho_1, \dots, \rho_n)^T$ est toujours Markovienne.

Soit $\bar{c} = \begin{pmatrix} 0 \\ c \end{pmatrix}$, une solution \bar{v} du système linéaire :

$$(\bar{M} - I - \lambda k I)\bar{v} + k\bar{c} = 0$$

sécrit sous la forme $\bar{v} = \begin{pmatrix} 0 \\ v \end{pmatrix}$ où v est solution du système linéaire de départ $Lv + c = 0$.

En utilisant les résultats du chapitre 1, \bar{v} s'interprète comme la solution de l'équation de Kolmogorov correspondant à un problème discret avec taux d'actualisation λk , coût instantané kc et matrice de transition \bar{M} :

$$\bar{v}(x) = \mathbb{E} \left[\sum_{n=0}^{\infty} \frac{k}{(1 + \lambda k)^{n+1}} \bar{c}(X_n) | X_0 = x \right]$$

On notera que lorsque la chaîne atteint l'état 0 elle y reste et comme la fonction \bar{c} vérifie $\bar{c}(0) = 0$, la fonction coût sécrit aussi :

$$\bar{v}(x) = \mathbb{E} \left[\sum_{n=0}^{\tau} \frac{k}{(1 + \lambda k)^{n+1}} c(X_n) | X_0 = x \right]$$

ou $\tau = \inf\{n | X_n \notin [1, n]\}$. On remarque que si la matrice M est Markovienne on a alors $\tau = \infty$.

Les problèmes de Dirichlet discrétisé conduisent à des matrices sous Markoviennes et les problème de Neumann à des matrices Markoviennes.

4.3 Discrétisation du processus

Nous nous plaçons ici à nouveau sous les hypothèse de la Proposition 4.2.4 et regardons la chaîne de Markov X_p de matrice de transition $\mathcal{M}^h \equiv \mathcal{A}_{(\Delta t^h \mathcal{A}^-, \Delta t^h \mathcal{D} + I, \Delta t^h \mathcal{A}^+)}^h$. Pour $x_n \in \Omega_h$:

$$\mathbb{E}[X_{n+1} - X_n | X_n = x_n] = h\Delta_t(x_n)\mathcal{A}^+(x_n) - h\Delta_t(x_n)\mathcal{A}^-(x_n) = \Delta_t(x_n)b(x_n)$$

Soit $Y_n = X_{n+1} - X_n - \Delta_t(x_n)b(x_n)$,

$$\begin{aligned} \mathbb{E}[(Y_n)^2 | X_n = x_n] &= (b\Delta_t)^2 \mathcal{D}\Delta_t + (h - b\Delta_t)^2 \Delta_t \mathcal{A}^+ + (h + b\Delta_t)^2 \Delta_t \mathcal{A}^- \\ &= (b\Delta_t)^2 \Delta_t (\mathcal{D} + \mathcal{A}^+ + \mathcal{A}^-) + h^2 \Delta_t (\mathcal{A}^+ + \mathcal{A}^-) \mp hb\Delta_t^2 \mathcal{A}^\pm \\ &= 2a\Delta_t - 2b^2 \Delta_t^2 \\ &= 2a(x_n)\Delta_t(x_n) + o(\Delta_t) \end{aligned}$$

La chaîne de Markov a la forme suivante :

$$X_{n+1} = X_n + \Delta_t(X_n)b(X_n) + M_n(X_n) \quad (4.34)$$

ou M_n vérifie $\mathbb{E}[M_n | X_n = x] = 0$ et $\mathbb{E}[M_n^2 | X_n = x] = 2a(x)\Delta_t(x) + o(\Delta_t)$

Considérons les temps aléatoires $t_n \equiv \sum_{k=0}^{n-1} \Delta_t^i$ avec $t_0 \equiv 0$ et $\Delta_t^i = \Delta_t(X_i)$. On associe à tout processus discret Z_n le processus continu $Z^h(t)$ tel que :

$$Z^h(t) \equiv Z_p \quad \text{pour } t \in [t_p, t_{p+1}).$$

Proposition 4.3.1. *Avec les notations précédentes et en posant $\beta_i \equiv \log(1 + \lambda \Delta_t^i)$, le critère de la proposition 4.2.4 sécrit aussi :*

$$\begin{aligned} v_h(x) &= \mathbb{E}_{x,0} \left[\int_0^{t_{\nu_{\Omega_h}}} \exp\left(-\int_0^s \beta^h(s) ds\right) c(X^h(s)) ds \right] \\ &+ \mathbb{E}_{x,0} \left[\exp\left(-\int_0^{t_{\nu_{\Omega_h}}} \beta^h(s) ds\right) \Phi(X^h(t_{\nu_{\Omega_h}})) \right] \end{aligned}$$

le processus $X^h(t)$ associé à X_n (4.34) vérifiant quant à lui :

$$X^h(t) = \int_0^t b(X^h(s))ds + M_t^h$$

Preuve : Pour $Z^h(s)$ associé à Z_i , on obtient facilement la propriété suivante :

$$\int_0^{t_{i+1}} Z^h(s)ds = \sum_{k=0}^i Z_i \Delta_t^i$$

En utilisant cette propriété on trouve :

$$\xi_k \equiv \prod_{i=0}^{k-1} \frac{1}{1 + \lambda \Delta_t^i} = \exp\left(-\sum_{i=0}^{k-1} \beta_i\right) = \exp\left(-\int_0^{t_k} \beta^h(s)ds\right)$$

et on conclut la première partie de la preuve en appliquant à nouveau cette propriété à $\sum_i \xi_i c(X_i) \Delta_t^i$. La seconde partie de la preuve se traite de la même façon. \square

4.4 Une section à déplacer dans le futur ?

On veut résoudre l'équation :

$$\begin{cases} \lambda v(x) = \inf_{u \in U} A^u v(x) + c^u(x) & \text{dans } \Omega \\ v = \Phi \text{ ou } \frac{\partial v}{\partial n} = \Phi & \text{sur } \partial\Omega \end{cases} \quad (4.35)$$

où Ω est un ouvert de \mathbb{R}^d et A^u est un opérateur du deuxième ordre elliptique

$$A^u v(x) \equiv \sum_{i,j=1}^d a_{ij}(x, u) \frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(x, u) \frac{\partial v}{\partial x_i}(x)$$

On rappelle que ces équations sont liées au problème de commande de processus de diffusion. Par exemple la solution de l'équation (4.35) avec conditions de Dirichlet homogène s'interprète comme :

$$v(x) = \inf_{U \in \mathcal{Z}^{\text{ZZZ}}} \mathbb{E} \left[\int_0^\tau e^{-\lambda t} c^{U_t}(X_t) dt + e^{-\lambda \tau} \Phi(X_\tau) \right]$$

τ est le premier instant où le processus X_t sort du domaine Ω , $\tau(\omega) = \inf\{t, X_t(\omega) \notin \Omega\}$.

On peut alors reprendre tous les cas déjà vu dans le cas des équations elliptique. Considérons par exemple le cas avec conditions de Dirichlet et avec l'opérateur $A^u(x)$ sans dérivées croisées et uniformément elliptique. On utilise la discrétisation (4.20) et en raisonnant comme dans le paragraphe (4.2.4) on obtient l'analogie de l'équation (4.24) :

$$\begin{cases} v^h(x) = \frac{1}{1 + \lambda \Delta t^h} \inf_{u \in U} (\mathcal{M}_h^u v^h(x) + \Delta t^h c_h^u(x)) & \text{pour } x \in \Omega_h \\ v^h(x) = \Phi(x) & \text{pour } x \in \partial^+ \Omega_h \end{cases} \quad (4.36)$$

on peut éliminer les termes $v^h(x)$ pour $x \in \partial^+\Omega_h$. Si on peut choisir Δt^h et h comme dans les propositions 4.2.3 et 4.2.4

$$h_i \leq \min_{x \in \Omega_h, u \in U} \frac{2a_{ii}(x, u)}{|b_i(x, u)|} \quad \text{et} \quad 0 < \Delta t(x) \leq \min_{u \in U} \frac{1}{2} \left(\sum_{i=1}^d a_{ii}(x, u)/h_i^2 \right)^{-1} \quad (4.37)$$

la matrice \mathcal{M}^u obtenue alors est sous stochastique.

Soit w la solution du problème continu supposée suffisamment régulière et prolongée pour être définie sur $\partial^+\Omega_h$. w vérifie :

$$\begin{cases} w(x) = \frac{1}{1+\lambda\Delta t^h} \inf_{u \in U} (\mathcal{M}_h^u w(x) + \Delta t^h (c_h^u(x) + A^u w(x) - \mathcal{M}_h^u w(x))) & \text{pour } x \in \Omega_h \\ v^h(x) = \Phi(x) & \text{pour } x \in \partial^+\Omega_h \end{cases} \quad (4.38)$$

On suppose que $\min_{x \in \Omega_h, u \in U} \frac{1}{2} (\sum_{i=1}^d a_{ii}(x, u)/h_i^2)^{-1}$ est atteint pour une valeurs strictement positive. On peut alors trouver α_1 et α_2 strictement positifs tels que $K = \alpha_2/\alpha_1$ soit indépendant de h et $\alpha_1 \leq \Delta t(x) \leq \alpha_2$. En utilisant la proposition 2.3.1 on obtient la stabilité du schéma numérique. En effet l'erreur sur Ω_h s'écrit :

$$\|w(x) - v^h(x)\|_\infty \leq \frac{K}{\lambda} \sup_{u \in U} \|A^u w(x) - \mathcal{M}_h^u w(x)\|_\infty$$

avec K constante indépendante de h .

4.5 Equations de Kolmogorov paraboliques

On veut résoudre ici

$$\begin{cases} \frac{\partial v(t, x)}{\partial t} + Av(t, x) - \lambda v(t, x) + c(t, x) = 0 & \text{dans } [0, T] \times \Omega \\ V(T, x) = f(x) & \forall x \in \Omega \\ v(t, x) = \Phi(t, x) \text{ ou bien } \frac{\partial v(t, x)}{\partial n} = \Phi(t, x) & \text{sur } [0, T] \times \partial\Omega \end{cases} \quad (4.39)$$

où A est un opérateur du deuxième ordre elliptique

$$Av(t, x) \equiv \sum_{i,j=1}^d a_{ij}(t, x) \frac{\partial^2 v}{\partial x_i \partial x_j}(t, x) + \sum_{i=1}^d b_i(t, x) \frac{\partial v}{\partial x_i}(t, x)$$

avec

$$\sum_{i,j=1}^d a_{ij}(t, x) \eta_i \eta_j \geq 0, \quad \forall x \in \Omega, t \in [0, T], \eta \in \mathbb{R}^d.$$

On rappelle que ces équations sont associées à l'évaluation d'espérances sur un horizon fini. Par exemple la solution de l'équation (4.39) avec conditions de Dirichlet $v = f$ sur $\partial\Omega$ s'interprète, sous certaines hypothèses de régularité, comme

$$\begin{aligned} v(t, x) &= \mathbb{E}_{x,t} \left[\int_t^{\tau \wedge T} e^{-\lambda s} c(s, X_s) ds \right] \\ &+ \mathbb{E}_{x,t} \left[e^{-\lambda(T-t)} \mathbb{I}_{\{T \leq \tau\}} f(X_T) + e^{-\lambda(\tau-t)} \mathbb{I}_{\{\tau < T\}} \Phi(\tau, X_\tau) \right] \end{aligned}$$

où $dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$, $a = \frac{1}{2}\sigma\sigma^T$, et τ est le premier instant de sortie de Ω . On suppose que l'on s'est ramené à $\Omega =]0, 1[^d$.

4.5.1 Méthode d'approximation

On doit cette fois approximer un opérateur ou interviennent des dérivées en temps et en espace. Pour ce qui concerne la partie discrétisation en espace, il est naturel d'utiliser les mêmes méthodes qu'au paragraphe 4.2.2 puisque les solutions stationnaires de l'équation parabolique (dans le cas homogène) sont solutions des équations elliptiques vues précédemment. Les opérateurs discrets associés à A ou L seront donc notés A_h et L_h et seront conformes à ceux décrits dans les paragraphes précédents.

On discrétise l'équation en temps par ce que l'on appelle les θ -schémas. Cela signifie que l'on se donne un paramètre $\theta \in [0, 1]$ et $\delta = \frac{T}{N_\delta}$ un pas de discrétisation en temps ($N_\delta \in \mathbb{N}^*$) et on approxime l'opérateur $\frac{\partial v}{\partial t} + Lv$ par le schéma suivant :

$$\mathcal{P}_{h,\delta}^\theta v(t, x) \equiv \partial_\delta v(t, x) + \theta A_h v(t, x) + (1 - \theta) A_h v(t + \delta, x) - \lambda v(t, x) \quad (4.40)$$

où $\partial_\delta v(t, x) = (v(t + \delta, x) - v(t, x))/\delta$. On utilise cette approximation pour estimer $v(t, x)$ aux points :

$$(t, x) \in Q_{\delta,h} \equiv \{k\delta, k = 0, \dots, N_\delta\} \times \Omega_h$$

($\Omega_h \cup \partial^+ \Omega_h$ pour la condition de Neumann). Pour tout $(t, x) \in Q_{\delta,h}$ on notera l'approximation discrète de $v(t, x)$ par $v_h^t(x)$ et v_h^t , pour t fixé, pourra être confondu avec un vecteur de dimension N_h .

Par exemple, pour le problème avec conditions de Dirichlet, l'approximation précédente conduit à remplacer l'EDP (4.39) par le système discret,

$$\begin{cases} (I + \delta\lambda I - \delta\theta A_h)v_h^t = (I + \delta(1 - \theta)A_h)v_h^{t+\delta} + \delta c_h^t & \text{pour } x \in \Omega_h \\ v_h^t(x) = \Phi(x) & \text{pour } x \in \partial^+ \Omega_h \text{ et } t \in \{k\delta, k = 0, \dots, N_\delta\} \\ v_h^T(x) = f(T, x) & \text{pour } x \in \Omega_h \end{cases} \quad (4.41)$$

la première équation se déduisant de $\mathcal{P}_{h,\delta}^\theta v_h^t + c_h^t = 0$. Le choix de A_h doit alors être fait en se reportant aux sections précédentes : Par exemple, si l'opérateur A est uniformément elliptique et sans dérivées croisées on peut utiliser pour A_h l'opérateur défini par l'équation (4.20) ou sa version d -dimensionnelle. Nous traiterons ce cas dans la suite du paragraphe. La méthode est la même pour tous les autres cas que nous laissons en exercice au lecteur.

Soit $e_h(t, x) = v_h^t(x) - v(x, t)$ pour $(t, x) \in Q_{\delta,h}$ où $v_h^t(x)$ est solution de l'équation (4.41) et $v(t, x)$ est solution de l'équation (4.39). Il est facile de vérifier que l'erreur $e_h(t, x)$ vérifie pour $x \in \Omega_h$ et $t \in \{k\delta, k = 0, \dots, N_\delta - 1\}$:

$$\mathcal{P}_{h,\delta}^\theta e(t, x) + (\mathcal{P}_\theta^h v - \partial_t v - Lv) = 0$$

En conclusion, $e_h(t, x)$ sur $Q_{\delta,h}$ est solution de la même équation que v_h^t (4.41) avec $\Phi \equiv 0$, $f \equiv 0$, $c_h(t, x) = c_h^e(t, x)$ et avec la définition :

$$c_h^e(t, x) \equiv \mathcal{P}_\theta^h v - \partial_t v - Lv$$

On suppose que sont vérifiées les deux hypothèses suivantes :

$$(H1) \quad 0 < h_i \leq \min_{(t,x) \in Q_{\delta,h}} \frac{2a_{ii}(t, x)}{|b_i(t, x)|}$$

$$(H2) \quad 0 < \delta(1 - \theta) \leq \frac{1}{2} \left(\sum_{i=1}^n a_{ii}(t, x)/h_i^2 \right)^{-1} \quad \text{pour tout } (t, x) \in Q_{\delta,h}$$

On notera que la condition (H2) n'impose aucune restriction sur δ quand $\theta = 1$.

Sous les hypothèses (H1) et (H2) et en notant comme au paragraphe 4.2.4 par $\bar{\mathcal{A}}_h$ la matrice carrée associée à l'opérateur A_h , on peut déduire (voir Propositions 4.2.3 et 4.2.4) les propriétés suivantes :

- La matrice $\mathcal{B}_h = -(I + \delta\lambda I - \delta\theta\bar{\mathcal{A}}_h)$ est diagonale fortement dominante de paramètre $(1 + \delta\lambda)$ pour tout $\theta \in (0, 1)$. La matrice \mathcal{B}_h est donc inversible avec $\|\mathcal{B}_h^{-1}\|_\infty \leq 1/(1 + \delta\lambda)$.
- La matrice $\mathcal{C}_h = I + \delta(1 - \theta)\bar{\mathcal{A}}_h$ est sous Markovienne.

On en conclut que le problème discret (4.41) est bien posé et se résout de façon rétrograde, pour tout $t = k\delta$, $k = N, \dots, 0$. Chaque étape demande la résolution d'un système linéaire inversible de taille $N_h = \text{card}(\Omega_h)$. Quand l'opérateur A ne dépend pas du temps, la matrice \mathcal{B}_h est constante et on est amené à résoudre plusieurs systèmes linéaires faisant intervenir la même matrice \mathcal{B}_h . On aura intérêt dans ce cas à factoriser la matrice (méthode LU ou QR) avant les itérations.

On peut aussi résoudre chaque système linéaire par une méthode itérative. En effet, pour $\bar{\delta}$ vérifiant :

$$0 < \bar{\delta} \leq \frac{1}{2} \left(\sum_{i=1}^n a_{ii}(t, x, u)/h_i^2 \right)^{-1} \quad \text{pour tout } (t, x, u) \in Q_{\delta, h} \times \mathcal{U}$$

la matrice $\mathcal{M}_h = I + \bar{\delta}\mathcal{A}_h$ est sous Markovienne. La première équation du système (4.41) s'écrit encore :

$$v_h^t = \frac{\delta\theta/\bar{\delta}}{1 + \lambda\bar{\delta} + \delta\theta/\bar{\delta}} \left\{ M_h v_h^t + \bar{\delta}/(\delta\theta) \left((I + \delta(1 - \theta)A_h)v_h^{t+\delta} + \delta c_h \right) \right\} \quad (4.42)$$

v_h^t apparaît comme l'unique point fixe d'un opérateur contractant ($v^x t_h$ est la fonction valeur d'un problème de Kolmogorov en horizon infini) ce qui conduit à une méthode numérique itérative pour calculer v_h^t .

En utilisant maintenant l'équation vérifiée par l'erreur :

$$\begin{cases} e_h^t = (-\mathcal{B})_h^{-1} \mathcal{C}_h e_h^{t+\delta} + (-\mathcal{B})_h^{-1} \delta c_h^e \\ e_h(T, x) = 0 \quad \text{pour } x \in \Omega_h \end{cases} \quad (4.43)$$

on obtient facilement pour tout $k \in [0, N_\delta]$:

$$e_h(T - k\delta, \cdot) = \sum_{l=1}^k ((-\mathcal{B})_h^{-1} \mathcal{C}_h)^{k-l} (-\mathcal{B})_h^{-1} \delta c_h^e(T - l\delta, \cdot) \quad (4.44)$$

d'où l'on déduit, avec la notation $\|f(x, t)\|_{\infty, Q_{\delta, h}} = \sup_{k \in [0, N_\delta]} \|f(\cdot, \delta k)\|_{\infty, \Omega_h}$, la majoration suivante pour $\lambda \neq 0$:

$$\|e_h(T - k\delta, \cdot)\|_{\infty, \Omega_h} \leq \sum_{l=1}^k \frac{1}{(1 + \delta\lambda)^{(k-l)}} \delta \|c_h^e\|_{\infty, Q_{\delta, h}} \leq \frac{1 + \delta\lambda}{\lambda} \|c_h^e\|_{\infty, Q_{\delta, h}}$$

et pour $\lambda = 0$ la majoration :

$$\|e_h(T - k\delta, \cdot)\|_{\infty, \Omega_h} \leq k\delta \|c_h^e\|_{\infty, Q_{\delta, h}} \quad (4.45)$$

En combinant les deux cas on conclut que :

$$\|e_h\|_{\infty, Q_{\delta, h}} \leq (T\mathbb{I}_{\lambda=0} + \frac{1 + \delta\lambda}{\lambda}\mathbb{I}_{\lambda \neq 0}) \|c_h^e\|_{\infty, Q_{\delta, h}} \quad (4.46)$$

On conclut sur la stabilité du schéma numérique en utilisant la définition de c_h^e :

$$\|e_h(\cdot)\|_{\infty, Q_{\delta, h}} \leq (T\mathbb{I}_{\lambda=0} + \frac{1 + \delta\lambda}{\lambda}\mathbb{I}_{\lambda \neq 0}) \|\mathcal{P}_\theta^h v - \partial_t v + Lv\|_{\infty, Q_{\delta, h}} \quad (4.47)$$

Nous donnons maintenant une interprétation probabiliste de l'erreur. On peut noter que la matrice $\mathcal{M}_{\delta, h}^\theta \equiv (1 + \delta\lambda)(-\mathcal{B})_h^{-1} \mathcal{C}_h$ est une matrice sous Markovienne. On déduit des propriétés de \mathcal{B}_h et de \mathcal{C}_h énoncée précédemment que $\|\mathcal{M}_{\delta, h}^\theta\|_\infty \leq 1$. D'autre part les éléments de \mathcal{C}_h sont positifs puisque \mathcal{C}_h est sous Markovienne et il en est de même des éléments de la matrice $(-\mathcal{B})_h^{-1}$ en utilisant la proposition 1.3.3. L'équation de l'erreur peut donc aussi s'écrire :

$$\begin{cases} e_h(t, \cdot) = \frac{1}{1 + \lambda\delta} (\mathcal{M}_h^\theta e_h(t + \delta, \cdot) + \delta d_h^e) \\ e_h(T, x) = 0 \quad \text{pour } x \in \Omega_h \end{cases} \quad (4.48)$$

avec $d_h^e(t, x) = (-\mathcal{B})^{-1}(1 + \lambda\delta)c_h^e(t, x)$. Quitte à augmenter l'espace d'état \mathcal{E} d'un point cimetière Δ , l'équation de l'erreur s'interprète comme la solution du problème de Kolmogorov :

$$e^n(x) = \mathbb{E}_{n, x} \left[\sum_{j=n}^{N_\delta \wedge \nu^n - 1} \frac{1}{(1 + \lambda\delta)^{j+1-n}} \delta d_h^e(j\delta, X^j) \mathbb{I}_{N \wedge \nu^n > n} \right] \quad (4.49)$$

où $X(t)$ est une chaîne de Markov d'espace d'état $\mathcal{E} \cup \{\Delta\}$ et $\nu_n = \inf\{k \geq n : X^k \notin \Omega_h\}$ est le temps de sortie de Ω_h . Il est alors facile de retrouver les résultats 4.47 à partir de cette nouvelle expression de l'erreur.

Nous passons maintenant quelques cas importants en revue :

– **Schéma explicite** : $\theta = 0$. Dans ce cas particulier v_h^t se calcule explicitement en fonction de $v_h^{t+\delta}$:

$$\begin{cases} v_h^t &= \frac{1}{1 + \delta\lambda} \left((I + \delta A_h) v_h^{t+\delta} + \delta c_h^t \right), \quad t = k\delta, k = N - 1, \dots, 0, \\ v_h^t(x) &= \Phi(x) \quad \text{pour } x \in \partial^+ \Omega_h \\ v_h^T(x) &= f_h(x) \quad \text{pour } x \in \Omega_h \end{cases} \quad (4.50)$$

les résultats précédent peuvent se résumer sous la forme suivante

Proposition 4.5.1. *Pour $\theta = 0$, le schéma \mathcal{P}_h^θ est stable au sens de la norme L_∞ si le pas de discrétisation en espace h vérifie la condition (H1) et si le pas de discrétisation en temps vérifie la condition (H2).*

On parle de schéma conditionnellement convergent car il y a une contrainte entre les pas de discrétisation en temps et en espace.

Comme dans le cas général on a une interprétation probabiliste du problème discret ou de l'erreur, mais elle est ici un peu plus naturelle. La solution du problème discret vérifie :

$$v^n(x) = \mathbb{E}_{n,x} \left[\sum_{j=n}^{N \wedge \nu^n - 1} \frac{1}{(1 + \lambda\delta)^{j+1-n}} \delta c_h(j\delta, X^j) \mathbb{I}_{N \wedge \nu^n > n} \right. \\ \left. + \frac{1}{(1 + \lambda\delta)^{N-n}} f(N, X^N) \mathbb{I}_{N < \nu^n} + \frac{1}{(1 + \lambda\delta)^{\nu^n - n}} \Phi(\nu^n, X^{\nu^n}) \mathbb{I}_{N \geq \nu^n} \right]$$

où $X(t)$ est une chaîne de Markov d'espace d'état $\Omega_h \cup \partial^+ \Omega_h$ et de matrice de transition $M = I + \delta A_h$ sur Ω_h . $\nu_n = \inf\{k \geq n : X^k \notin \Omega_h\}$ est le temps de sortie de l'ensemble Ω_h pour la chaîne démarrant en x à la date n .

- **Schéma totalement implicite : $\theta = 1$.** Comme nous l'avons précédemment vu, la particularité est que la condition (H2) n'impose aucune contraintes sur δ .

Proposition 4.5.2. *Pour $\theta = 1$, le schéma \mathcal{P}_h^θ est stable au sens de la norme L_∞ si le pas de discrétisation en espace h vérifie la condition (H1).*

Ce schéma converge toujours lorsque que h et δ tendent vers 0. Contrairement au schéma explicite, il n'y a pas de condition liant les pas de discrétisation en espace et en temps. On dit que le schéma implicite est inconditionnellement convergent.

- **Schéma de Crank-Nicholson.**

Le schéma de Crank-Nicholson consiste à prendre $\theta = \frac{1}{2}$. C'est le schéma le plus couramment utilisé. Il est surtout intéressant quand l'opérateur A ne dépend pas du temps car dans ce cas il est précis à l'ordre 2 en temps alors que pour les autres valeurs de θ on obtient un schéma précis à l'ordre 1.

$$\mathcal{P}_{h,\delta}^\theta v(t, x) = \partial_\delta v(t, x) + \frac{1}{2} A_h(v(t, x) + v(t + \delta, x)) - \lambda v(t, x)$$

On suppose que le schéma est d'ordre 2 en espace $A_h v(t, x) - A v(t, x) = O(h^2)$ on obtient alors :

$$\mathcal{P}_{h,\delta}^\theta v(t, x) = \partial_\delta v(t, x) + \frac{1}{2} A(v(t, x) + v(t + \delta, x)) + O(h^2) - \lambda v(t, x) \quad (4.51)$$

$$= \frac{\partial v}{\partial t}(t + \delta/2, x) + O(\delta^2) \frac{1}{2} A(v(t, x) + v(t + \delta, x)) - \lambda v(t, x) + O(h^2) \quad (4.52)$$

$$= \frac{\partial v}{\partial t}(t + \delta/2, x) + \frac{1}{2} A(v(t + \delta/2, x) - \lambda v(t, x) + O(\delta^2 + h^2)) \quad (4.53)$$

$$= (\partial_t + L)v(t + \delta/2, x) + O(\delta^2 + h^2) \quad (4.54)$$

On obtient donc un schéma d'ordre 2 en temps. On notera cependant que l'opérateur $\partial_t + L$ est évalué en $t + \delta/2$ et il faudra donc supposer que c ne dépend pas du temps pour avoir une erreur en ordre 2 en temps.

4.5.2 Calcul du rayon spectral de $(-\mathcal{B}_h)^{-1} \mathcal{C}_h$

Nous avons vu que pour $\theta \neq 1$ le schéma est à priori conditionnellement convergent pour la norme L_∞ . Nous montrons ici que sans la condition (H2) le rayon spectral de la matrice $(-\mathcal{B}_h)^{-1} \mathcal{C}_h$ est plus petit que 1. Cela permet dans certains cas particuliers d'obtenir la stabilité inconditionnelle au sens L^2 du schéma (par exemple pour les matrices normales $\rho(A) = \|A\|_{L^2}$).

Dans ce paragraphe l'opérateur A est supposé ne pas dépendre du temps et on suppose vérifiée l'hypothèse (H1). On se limite au cas $\theta \neq 0$ et on s'intéresse au rayon spectral de la matrice :

$$\mathcal{D}_h = (I + \delta\lambda I - \delta\theta\overline{\mathcal{A}}_h)^{-1}(I + \delta(1 - \theta)\overline{\mathcal{A}}_h)$$

Sous l'hypothèse (H1) il existe $\gamma(h)$ tel que la matrice $\mathcal{G} = \gamma(h)I + \overline{\mathcal{A}}_h$ soit positive et vérifie aussi $\sum_j \mathcal{G}_{ij} \leq \gamma(h)$, la démonstration est la même que pour la proposition 4.2.4 et $\gamma(h)$ est donné par la formule :

$$\gamma(h) = 2 \sup_{x \in \Omega_h} \left(\sum_{i=1}^n a_{ii}(x)/h_i^2 \right)$$

On en conclut que $\rho(\mathcal{G}) \leq \|\mathcal{G}\|_\infty \leq \gamma(h)$ et donc que $\rho(\gamma(h)I + \overline{\mathcal{A}}_h) \leq \gamma(h)$. Les valeurs propres de $\overline{\mathcal{A}}_h$ sont donc dans le plan complexe dans un disque de rayon $\gamma(h)$ et de centre $(-\gamma(h), 0)$.

Étudions maintenant la fonction :

$$\Psi(x) = \frac{1 + (1 - \theta)x}{1 + \alpha - \theta x}$$

avec $\alpha > 0$. Cette transformation est de la forme $(b + ax)/(d + cx)$ avec $ad - bc = 1 + \alpha(1 - \theta) \neq 0$ et en utilisant [12] page 337-338 on conclut que pour $\theta \neq 0$ elle transforme un disque ouvert ne contenant pas le point $(1/\theta, 0)$ en un disque ouvert. D'autre part (a, b, c, d) étant réels la symétrie par rapport à l'axe $\{y = 0\}$ est préservé par la transformation Ψ . Les valeurs propres de \mathcal{A}_h qui sont ds un disque de rayon $\gamma(h)$ et de centre $(-\gamma(h), 0)$ sont après avoir appliqué la transformation Ψ dans un disque centré sur l'axes des x et de coordonnée en x dans l'intervalle $(\beta, 1/(1 + \alpha))$. On cherche maintenant à préciser la valeur de β . On cherche donc l'image par l'application Ψ de $(-2\gamma(h), 0)$. On obtient facilement les majorations suivante :

$$|\Psi(x)| \leq \max \left(\frac{1}{1 + \alpha}, |\Psi(-2\gamma(h))| \right) \quad \text{pour } x \in \mathbb{R}^-$$

$$|\Psi(x)| < \left| \frac{\theta - 1}{\theta} \right| \quad \text{pour } x < -1/(1 - \theta)$$

Si $\theta > 1/2$ on a alors $|(\theta - 1)/\theta| < 1$ et donc :

$$\Psi(x) \leq \max(1/(1 + \alpha), |\theta - 1/\theta|) < 1.$$

On conclut que le rayon spectral de \mathcal{D}_h vérifie :

$$\rho(\mathcal{D}_h) \leq \max \left(\frac{1}{1 + \lambda\delta}, \frac{1}{\theta} - 1 \right) < 1$$

Si maintenant $\theta = 1/2$ alors $|\theta - 1/\theta| = 1$ on a cette fois :

$$\rho(\mathcal{D}_h) \leq \max \left(\frac{1}{1 + \lambda\delta}, \Phi(-2\delta\gamma(h)) \right) < 1.$$

On notera que pour $0 < \theta \leq 1/2$ on peut rajouter une condition sur δ pour que $-2\gamma(h)\delta \geq -1/(1 - \theta)$ ce qui permet alors d'avoir la majoration $\Psi(x) \leq 1/(1 + \alpha)$ pour $x \in (-2\gamma(h), 0)$ on conclut alors aussi sur le rayon spectral avec cette fois une relation liant δ et h mais qui est plus faible que H2.

4.6 Le cas HJB a déplacer dans un chapitre suivant

Soit ici :

$$\begin{cases} \frac{\partial v(t, x)}{\partial t} + \min_u A^u v(t, x) - \lambda v(t, x) + c^u(t, x) = 0 & \text{dans } [0, T[\times \Omega \\ V(T, x) = f(x) & \forall x \in \Omega \\ v(t, x) = \Phi(t, x) \text{ ou bien } \frac{\partial v(t, x)}{\partial n} = \Phi(t, x) & \text{sur } [0, T[\times \partial\Omega \end{cases} \quad (4.55)$$

où A^u est un opérateur du deuxième ordre elliptique

$$Av(t, x, u) \equiv \sum_{i,j=1}^d a_{ij}(t, x, u) \frac{\partial^2 v}{\partial x_i \partial x_j}(t, x) + \sum_{i=1}^d b_i(t, x, u) \frac{\partial v}{\partial x_i}(t, x)$$

avec

$$\sum_{i,j=1}^d a_{ij}(t, x) \eta_i \eta_j \geq 0, \quad \forall x \in \Omega, t \in [0, T], \eta \in \mathbb{R}^d.$$

On discrétise l'équation en temps à nouveau en utilisant des θ -schémas. Soit $\delta = \frac{T}{N_\delta}$ le pas de discrétisation en temps ($N_\delta \in \mathbb{N}^*$), on utilise l'approximation suivante $\mathcal{P}_{h,\delta}^\theta v(t, x) = 0$ avec la définition :

$$\mathcal{P}_{h,\delta}^\theta v(t, x) \equiv \partial_\delta v(t, x) + \min_u \theta A_h^u v(t, x) + (1 - \theta) A_h^u v(t + \delta, x) - \lambda v(t, x) + c^u \quad (4.56)$$

où $\partial_\delta v(t, x) = (v(t + \delta, x) - v(t, x))/\delta$. On utilise cette approximation pour estimer $v(t, x)$ aux points :

$$(t, x) \in Q_{\delta,h} \equiv \{k\delta, k = 0, \dots, N_\delta\} \times \Omega_h$$

($\Omega_h \cup \partial^+ \Omega_h$ pour la condition de Neumann). Pour tout $(t, x) \in Q_{\delta,h}$ on notera l'approximation discrète de $v(t, x)$ par $v_h^t(x)$ et v_h^t pour t fixé pourra être confondu avec un vecteur de dimension N_h .

À nouveau ici, nous nous limitons aux conditions de Dirichlet et nous supposons que l'opérateur A est uniformément elliptique et sans dérivées croisées. On peut utiliser pour A_h l'opérateur défini par l'équation (4.20) ou sa version d -dimensionnelle. Les diverses combinaisons d'hypothèses possibles étant laissées comme exercice au lecteur.

Les hypothèses (H1) et (H2) du cas sans contrôle deviennent ici :

$$(H1') \quad 0 < h_i \leq \min_{(t,x,u) \in Q_{\delta,h} \times \mathcal{U}} \frac{2a_{ii}(t, x, u)}{|b_i(t, x, u)|}$$

$$(H2') \quad 0 < \delta(1 - \theta) \leq \frac{1}{2} \left(\sum_{i=1}^n a_{ii}(t, x, u) / h_i^2 \right)^{-1} \quad \text{pour tout } (t, x, u) \in Q_{\delta,h} \times \mathcal{U}$$

On se donne aussi $\bar{\delta}$ vérifiant :

$$(H3) \quad 0 < \bar{\delta} \leq \frac{1}{2} \left(\sum_{i=1}^n a_{ii}(t, x, u) / h_i^2 \right)^{-1} \quad \text{pour tout } (t, x, u) \in Q_{\delta,h} \times \mathcal{U}$$

l'existence de δ et $\bar{\delta}$ vérifiant les hypothèses (H1) et (H2) étant assurées par l'uniforme ellipticité de a .

Des hypothèses (H1) et (H3) on déduit l'existence d'une matrice sous Markovienne \mathcal{M}_h^u (et un opérateur M^u associé) telle que $I + \bar{\delta}\mathcal{A}_h^u = \mathcal{M}$ (c'est encore comme dans les Propositions 4.2.3 et 4.24).

On suppose ici $\theta \neq 0$. Le θ -schémas s'écrit pour $t \in [0, N_\delta - 1]$ et $x \in \Omega_h$:

$$\begin{aligned} & v_h^{t+\delta} - v_h^t + \min_{u \in \mathcal{U}} \left\{ \theta \delta A_h^u v_h^t + (1 - \theta) \delta A_h^u v_h^{t+\delta} + \delta c_h^u - \lambda \delta v_h^t \right\} = 0 \\ & -v_h^t + \min_{u \in \mathcal{U}} \left\{ \delta \theta \left(\frac{M_h^u - I}{\bar{\delta}} \right) + (I + \delta(1 - \theta)A_h^u) v_h^{t+\delta} - \lambda \delta v_h^t + \delta c_h^u \right\} \\ v_h^t &= \frac{\delta \theta / \bar{\delta}}{1 + \lambda \delta + \delta \theta / \bar{\delta}} \min_{u \in \mathcal{U}} \left\{ M_h^u v_h^t + \bar{\delta} / (\delta \theta) \left((I + \delta(1 - \theta)A_h^u) v_h^{t+\delta} + \delta c_h^u \right) \right\} \end{aligned} \quad (4.57)$$

Le coefficient $\frac{\delta \theta / \bar{\delta}}{1 + \lambda \delta + \delta \theta / \bar{\delta}}$ s'écrit, même pour $\lambda = 0$ sous la forme $1/(1 + \nu(\delta))$ avec $\nu(\delta) > 0$. Le calcul de v_h^t à partir de $v_h^{t+\delta}$ apparaît donc comme un problème de contrôle de chaîne de Markov en horizon infini traité au Chapitre 2 paragraphe 2.3. Le calcul de v_h^t peut donc se faire pour chaque pas de temps en utilisant un algorithme d'itération sur les valeurs ou un algorithme de Howard.

En utilisant un raisonnement maintenant classique on vérifie que la solution $v(x, t)$ du problème continue vérifie sur $Q_{\delta, h}$ la même équation que v_h^t en remplaçant le terme c_h^u par $c_h^u + \partial_t v + A^u v - (\partial_\delta v - A_h^u v)$. On peut alors utiliser la proposition 2.3.1 du Chapitre 2 sur le problème (4.57) pour obtenir l'estimation suivante sur $e(t, x) = v_h^t - v(t, \cdot)$:

$$\|e(t, \cdot)\|_\infty \leq \frac{1}{1 + \lambda \delta} \|(I + \delta(1 - \theta)A_h^u)e(t + \delta, \cdot) + \delta(\partial_t v + A^u v - (\partial_\delta v - A_h^u v))\|_\infty \quad (4.58)$$

Jusqu'ici nous n'avons utilisé que les hypothèses (H1) et (H3) en rajoutant maintenant l'hypothèse (H2) la matrice $(I + \delta(1 - \theta)A_h^u)$ est sous markovienne et on déduit de l'équation précédente :

$$\|e(t, \cdot)\|_\infty \leq \frac{1}{1 + \lambda \delta} \{ \|e(t + \delta, \cdot)\|_\infty + \delta \|\partial_t v + A_h^u v - (\partial_\delta v - A_h^u v)\|_\infty \} \quad (4.59)$$

Et donc en utilisant la condition finale $e(T, x) = 0$:

$$\|e(t, \cdot)\|_\infty \leq \left(\frac{1}{\lambda} \mathbb{I}_{\{\lambda \neq 0\}} + T \mathbb{I}_{\{\lambda = 0\}} \right) \|\partial_t v + A_h^u v - (\partial_\delta v - A_h^u v)\|_\infty \quad (4.60)$$

4.7 Résolution de systèmes linéaires

Nous avons vu que la résolution des équations de Kolmogorov nécessitait (sauf dans le cas parabolique avec schéma explicite) la résolution de systèmes linéaires : $Au = b$ où la matrice A est tri-diagonale par blocs et inversible.

Deux types de méthodes de résolution sont disponibles : les méthodes directes et les méthodes itératives (algorithmes de Jacobi, Gauss-Seidel, relaxation, multigrilles).

4.7.1 Méthodes directes

La solution d'un système linéaire $Au = b$, A matrice inversible, ne s'obtient pas en calculant la matrice A^{-1} , puis en calculant le vecteur $A^{-1}b$. Le calcul de la matrice A^{-1} est en effet équivalent à la résolution des n systèmes linéaires (n : ordre de la matrice) : $Au_j = e_j$, $1 \leq j \leq n$. Autrement dit, on remplacerait par une telle méthode la résolution d'un système linéaire par la résolution de n systèmes linéaires, suivie de la multiplication de la matrice A^{-1} par le vecteur b !

Remarquons également que si la matrice A est triangulaire inférieure (ou triangulaire supérieure), la résolution d'un système linéaire $Au = b$ est immédiate; il s'écrit en effet :

$$\begin{aligned} a_{11}u_1 &= b_1, \\ a_{21}u_1 + a_{22}u_2 &= b_2 \\ &\vdots \\ a_{n-1,1}u_1 + \dots + a_{n-1,n-1}u_{n-1} &= b_{n-1}, \\ a_{n1}u_1 + a_{n2}u_2 + \dots + a_{nn}u_n &= b_n \end{aligned}$$

et puisque $a_{11}a_{22} \dots a_{nn} = \det(A) \neq 0$, on résout le système en calculant u_1 de la première équation, puis u_2 de la deuxième, etc. Cette procédure s'appelle la *méthode de descente* et s'applique aussi aux matrices triangulaires par blocs.

Le principe des méthodes directes repose sur les 2 remarques précédentes.

La *méthode de Gauss* est une méthode générale de résolution directe d'un système linéaire en utilisant un nombre de multiplications proportionnel à n . Elle comporte 3 étapes :

- une *procédure d'élimination* (successive des inconnues) qui équivaut à déterminer une matrice M inversible telle que la matrice MA soit triangulaire.
- On calcule simultanément le vecteur Mb .
- On résout le système linéaire $MAu = Mb$.

Dans les calculs effectifs, on ne calcule pas M mais seulement la matrice MA et le vecteur Mb .

L'interprétation matricielle de la méthode de Gauss est la factorisation LU . Ce résultat montre que toute matrice inversible, (et cela est particulièrement simple dans le cas de matrice tri-diagonale par blocs), peut s'écrire comme produit d'une matrice triangulaire inférieure L par une matrice triangulaire supérieure U .

4.7.2 Méthodes itératives

On présente ici un bref aperçu des méthodes itératives de résolution des systèmes linéaires.

Etant donnée une matrice inversible A et un vecteur b , on souhaite calculer la solution u du système linéaire

$$Au = b.$$

Supposons que l'on ait trouvé une matrice B et un vecteur c tels que la matrice $(I - B)$ soit inversible, et tels que la solution unique du système linéaire

$$u = Bu + c$$

soit également la solution de

$$Au = b.$$

La forme du système

$$u = Bu + c$$

suggère la définition d'une méthode itérative de résolution du système linéaire $Au = b$. On se donne un vecteur initial u_0 arbitraire, et on définit la suite de vecteurs $(u_k)_{k \geq 0}$ par

$$u_{k+1} = Bu_k + c.$$

On dit que la méthode est convergente si

$$\lim_{k \rightarrow \infty} u_k = u, \text{ pour tout vecteur initial } u_0.$$

Le résultat suivant donne le critère fondamental de convergence des méthodes itératives. On notera qu'il ne fait intervenir que la matrice B appelée la matrice de la méthode itérative considérée.

Théorème 4.7.1. *Les propositions suivantes sont équivalentes :*

- (i) La méthode itérative est convergente
- (ii) $\rho(B) < 1$
- (iii) $\|B\| < 1$ pour au moins une norme matricielle $\|\cdot\|$

où $\rho(B)$ désigne le rayon spectral de la matrice B défini comme le maximum des valeurs absolues des valeurs propres de B .

Etant donné 2 méthodes itératives convergentes, la plus rapide est celle dont la matrice a le plus petit rayon spectral.

Supposons que l'on puisse décomposer A sous la forme :

$$A = M - N$$

où M est inversible et facile à inverser, c'est à dire, pratiquement, diagonale ou triangulaire, ou encore diagonale ou triangulaire par blocs. On a les équivalences :

$$Au = b \iff Mu = Nu + b \iff u = M^{-1}Nu + M^{-1}b.$$

La dernière équation étant de la forme $u = Bu + c$, on lui associe la méthode itérative

$$u_{k+1} = M^{-1}Nu_k + M^{-1}b, \quad u_0 \text{ arbitraire}$$

dont la matrice est

$$B = M^{-1}N = I - M^{-1}A.$$

Pratiquement on est amené à résoudre les systèmes linéaires successifs :

$$Mu_{k+1} = Nu_k + b, \quad k \geq 0.$$

Soit $A = (a_{ij})$ une matrice d'ordre n telle que $a_{ii} \neq 0$, $1 \leq i \leq n$ et décomposons A de la manière suivante :

$$A = D - E - F$$

où D est la matrice diagonale $D = \text{diag}(a_{ii})$, $-E$ et $-F$ sont respectivement les matrices avec pour termes non nuls les éléments de A situés respectivement au-dessous et au-dessus de la diagonale.

- **Algorithme de Jacobi.** La méthode de Jacobi consiste à considérer la décomposition

$$A = D - (E + F),$$

c'est à dire $M = D$ et $N = E + F$. On écrit les équivalences :

$$Au = b \iff Du = (E + F)u + b \iff u = D^{-1}(E + F)u + D^{-1}b,$$

ce qui conduit à la méthode itérative de Jacobi :

$$Du_{k+1} = (E + F)u_k + b \iff u_{k+1} = D^{-1}(E + F)u_k + D^{-1}b.$$

La matrice de cette méthode, appelée matrice de Jacobi est

$$J = D^{-1}(E + F) = I - D^{-1}A.$$

- **Algorithme de Gauss-Seidel.** On améliore la convergence en considérant la décomposition

$$A = (D - E) - F$$

qui conduit à l'algorithme

$$Du_{k+1} = Eu_{k+1} + Fu_k + b \iff u_{k+1} = (D - E)^{-1}Fu_k + (D - E)^{-1}b.$$

La matrice de cette méthode, appelée matrice de Gauss-Seidel est

$$\mathcal{L}_1 = (D - E)^{-1}F.$$

- **Algorithme de relaxation.** L'algorithme de Gauss-Seidel peut être vu comme un cas particulier de l'algorithme de relaxation. Cette méthode consiste à "faire passer une partie de la matrice D dans N ", c'est à dire à considérer la décomposition

$$A = \left(\frac{D}{\omega} - E\right) - \left(\frac{1 - \omega}{\omega}D + F\right)$$

où le paramètre de relaxation $\omega \neq 0$ permet d'accélérer la convergence s'il est bien choisi. On parle de sur-relaxation si $\omega > 1$ et de sous-relaxation si $\omega < 1$.

L'algorithme s'écrit

$$\left(\frac{D}{\omega} - E\right)u_{k+1} = \left(\frac{1 - \omega}{\omega}D + F\right)u_k + b.$$

La matrice de cette méthode, appelée matrice de relaxation est

$$\mathcal{L}_\omega = \left(\frac{D}{\omega} - E\right)^{-1}\left(\frac{1 - \omega}{\omega}D + F\right)$$

Nous référons à Ciarlet [?] pour les résultats de convergence de ces méthodes.

- **Méthodes multigrilles.** Ce sont des algorithmes rapides qui permettent d'accélérer la convergence des relaxations. On se référera par exemple à [6, 30, 31] pour une description détaillée de l'algorithme multigrille.

4.8 Un exemple numérique

Nous illustrons les paragraphes précédent par une application numérique. On veut résoudre ici :

$$\begin{cases} \frac{\partial v(t, x)}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial x^2} + b \frac{\partial v}{\partial x} + c(t, x) = 0 & \text{dans } [0, T[\times \Omega \\ V(T, x) = \sin(\pi x) & \forall x \in \Omega = (-1, 1) \\ v(t, x) = 0 & \text{sur } [0, T[\times \partial\Omega \\ c(t, x) = -e^{t-T} \cos(\pi x), \sigma = \frac{\sqrt{2}}{\pi} \text{ et } b = \frac{1}{\pi} \end{cases} \quad (4.61)$$

Le problème précédent admet une solution explicite $v(x) = \exp(t - T)\sin(\pi x)$ que nous utiliserons pour évaluer les solutions de nos algorithmes numérique.

```
-->sigma = sqrt(2)/%pi; b = 1/%pi; L=1, T=1;
-->hmax= sigma^2/abs(b);                               ← le pas en espace h ≤ hmax

-->pmin=ceil((2*L)/hmax)+1;
-->p=maxi(pmin,20);                                     ← le nombre de points de
  discrétisation associés
-->x = linspace(-L,L,p)                                 ← vecteur des points de
  discrétisation
-->x = x(2:$-1)                                         ← on retire les points du bord
h=x(2)-x(1);                                           ← le pas en espace choisit

-->alpha = sigma**2/(2*h^2) - b/(2*h);
-->beta = - sigma^2/h^2 ;
-->gama = sigma^2/(2*h^2) + b/(2*h) ;

-->n= p - 2 ;                                           ← la matrice Ah
-->Ah= diag(beta*ones(1,n))+diag(gama*ones(1,n-1),1)+...
  diag(alpha*ones(1,n-1),-1) ;

-->theta = 3/4;                                         ← choix de θ pour la θ-méthode
-->dtmax = 1/((sigma^2/h^2)*(1-theta))                 ← borne pour assurer la stabi-
lite conditionnelle

-->delta = dtmax/2;
-->q= ceil((T/delta));q=maxi(q,50); delta = T/q;
-->t = linspace(0,T,q+1);                               ← temps discrets vérifiant la borne de stabi-
lite conditionnelle

-->Bmh= eye(Ah) - delta *theta*Ah;                       ← la matrice - Bh
-->Ch= eye(Ah) + delta*(1-theta)*Ah ;                   ← la matrice Ch

-->deff(' [y]=f(x)', 'y=sin(%pi*x)')
-->fh = feval( x' , f);                                  ← la condition finale

-->deff(' [y]=g(x,t)', 'y= -exp(t-T)*cos(%pi*x)');
```

```

-->c=feval(x,t,g);                               ← la fonction c(t, x)

-->W= Bmh^(-1);                                  ← précalcul de  $B_h^{-1}$ 

-->v=zeros(n,q+1);
-->v(:,q+1)=fh ;                                 ← initialisation de v(x, T)
-->for i=(q:-1:1) , v(:,i) = W*(Ch*v(:,i+1) + delta*c(:,i)); end ← Itérations

-->plot3d(x/maxi(x),t,v/maxi(abs(v)),35,45,"X@T@V"); ← Graphique Figure 4.2
-->deff(' [y]=Vref(x,t)', 'y=exp(t-T)*sin(%pi*x)');
-->vref=feval(x,t,Vref);

-->Er=maxi(abs(vref-v))                          ← Evaluation de l'erreur

```

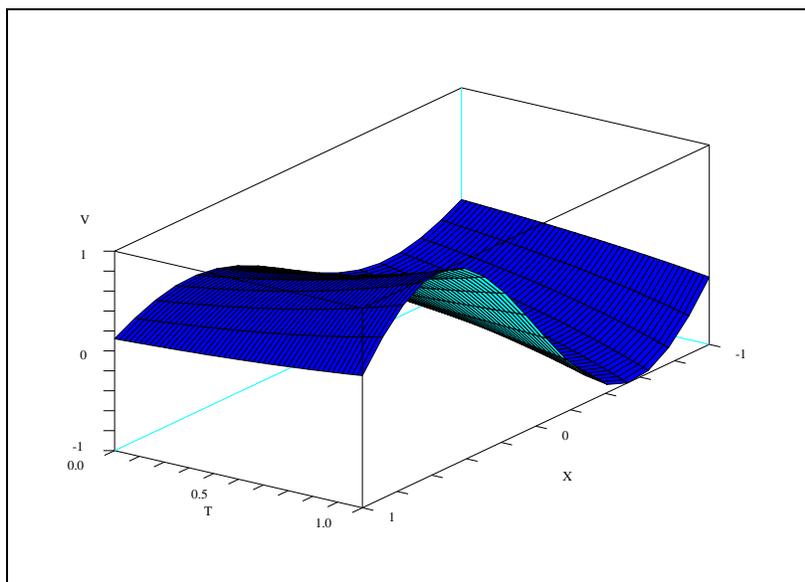


FIG. 4.2 – Fonction valeur

4.9 Corrections

Exercice 4.2.1

$$L_h v = a(x) \frac{v(x+h) - 2v(x) + v(x-h)}{h^2} - \lambda v(x)$$

– Au bord $x = 0$ on obtient avec le schéma (4.4) :

$$B_h v = \frac{v(x+h) - v(x-h)}{2h}$$

Pour ne pas utiliser $v(-h)$ on choisit $\alpha(x) = 2a(x)/h$ et on a alors :

$$\bar{L}_h v = 2a(x) \frac{v(x+h) - v(x)}{h^2} - \lambda v(x) \quad \text{pour } x = 0$$

Il faut donc regarder la consistance de \bar{L}_h avec l'opérateur $\bar{L}v = a(x)v'' + 2(a(x)/h)v'(x) - \lambda v(x)$ au point $x = 0$.

$$\begin{aligned} |\bar{L}_h v - \bar{L}v| &= \left| 2a(0) \frac{v(x+h) - v(x)}{h^2} - a(0)v'' - 2(a(0)/h)v'(x) \right| \\ &= 2(|a(0)|/h^2)|v(h) - v(0) - (h^2/2)v''(0) - hv'(0)| \\ &\leq K|h| \end{aligned}$$

– Au bord $x = 0$ on obtient avec le schéma (4.5) :

$$B_h v = \frac{v(x) - v(x-h)}{h}$$

On choisit cette fois $\alpha(x) = a(x)/h$ et on a alors :

$$\bar{L}_h = a(x) \frac{v(x+h) - v(x)}{h^2} - \lambda v(x) \quad \text{pour } x = 0$$

Pour $\bar{L}v = a(x)v'' + (a(x)/h)v'(x) - \lambda v(x)$, on obtient cette fois avec un développement de Taylor, en tenant compte du fait que v est solution de $Lv + c = 0$ dans Ω :

$$|\bar{L}_h v - Lv + \frac{1}{2}(c(0) + \lambda v(0))| \leq K|h|$$

Exercice 4.2.2 La première partie se déduit par un calcul direct de la formule des valeurs intermédiaires utilisée sur une formule de Taylor à l'ordre 4 :

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u^{(3)}(x) + \frac{h^4}{24}u^{(4)}(x + \theta h)$$

avec $\theta \in [0, 1]$. Pour $u \in C^{2,\alpha}$ on considère la formule de Taylor à l'ordre un avec reste intégral :

$$v(x+h) = v(x) + hv^{(1)}(x) + \int_x^{x+h} (x+h-u)v^{(2)}(u)du$$

Pour $v \equiv x^2/2$ on trouve que $\int_x^{x+h} (x+h-u)du = h^2/2$ on a donc aussi :

$$v(x+h) = v(x) + hv'(x) + h^2/2v^{(2)}(x) + \int_x^{x+h} (x+h-u)(v^{(2)}(u) - v^{(2)}(x))du$$

On obtient donc avec $L_h v = (v(x+h) - 2v(x) + v(x-h))/h^2$:

$$\begin{aligned} L_h v - v^{(2)}(x) &= \frac{1}{h^2} \left(\int_x^{x+h} (x+h-u)(v^{(2)}(u) - v^{(2)}(y))du \right. \\ &\quad \left. - \int_{x-h}^x (u - (x-h))(v^{(2)}(u) - v^{(2)}(y))du \right) \end{aligned}$$

ce qui conduit à $|L_h v - v^{(2)}(x)| \leq K|h|^\alpha$ chaque intégrale se majorant facilement. Par exemple :

$$\begin{aligned} \left| \int_x^{x+h} (x+h-u)(v^{(2)}(u) - v^{(2)}(x))du \right| &\leq \int_x^{x+h} (x+h-u)K|u-x|^\alpha du \\ &\leq \frac{K}{(\alpha+1)(\alpha+2)}|h|^{2+\alpha} \end{aligned}$$

Exercice 4.2.3 On a tout d'abord $v_2 = -L_{22}^{-1}(L_{21}v_1^h + c_2^h)$, la matrice $A = (-L_{22}^{-1}L_{21})$ est à coefficients positifs et vérifie $Ae_1 = e_2$ compte tenu de l'hypothèse $\sum_j (L_2^h)_{i,j} = 0$. On a donc $\|A\|_\infty = 1$ et :

$$\|v_2^h\|_\infty \leq \|v_1^h\|_\infty + \|-L_{22}^{-1}\|_\infty \|c_2^h\|_\infty$$

En remplaçant v_2^h par sa valeur dans $L_{11}v_1^h + L_{12}v_2^h + c_1^h = 0$ on obtient :

$$Bv_1^h + L_{12}(-L_{22}^{-1}c_2^h) + c_1^h = 0 \quad \text{avec} \quad B \equiv L_{11} + L_{12}A \quad (4.62)$$

et B est a diagonale fortement dominante de paramètre λ . En effet, on montre facilement que les éléments hors diagonaux sont positifs et la propriété $Ae_1 = e_2$ donne $Be_1 = L_1^h e_1$ ce qui permet de conclure puisque $L_1^h e_1 \leq -\lambda < 0$ par hypothèse . On a donc finalement :

$$\|v_1^h\|_\infty \leq C(\|L_{12}(-L_{22}^{-1})c_2^h\|_\infty + \|c_1^h\|_\infty) \leq C'(\|L_{22}^{-1}\|_\infty \|c_2^h\|_\infty + \|c_1^h\|_\infty)$$

ce qui permet de conclure la première question. Soit maintenant v solution de $Lv + c^h = 0$ on a alors : $L^h e_h = L^h v - L^h v_h = L^h v - Lv + Lv - L^h v_h = L^h v - Lv$ et en utilisant la question précédente on trouve facilement :

$$\|e_h\|_\infty \leq C_2 h^\alpha \|L_2 v - L_2^h v\|_\infty + C_1 \|L_1 v - L_1^h v\|_\infty$$

Pour montrer la dernière question, on utilise directement l'équation (4.62) qui nous donne :

$$\|v_1^h\|_\infty \leq \|B^{-1}\|_\infty \|L_{12}(-L_{22}^{-1}c_2^h) + c_1^h\|_\infty$$

et l'équation pour l'erreur s'en déduit immédiatement.

Chapitre 5

Contrôle de processus de diffusion

5.1 Introduction

La théorie du contrôle stochastique a de nombreuses applications en gestion et en finance. En effet, dans ces domaines, on considère des systèmes dynamiques (c'est à dire évoluant au cours du temps) en avenir incertain et sur lesquels on peut agir en vue d'optimiser un certain critère économique.

Pour décrire un problème de contrôle stochastique, il est important de préciser quelle est l'information disponible à tout instant. Plusieurs situations sont possibles :

1. Le “contrôleur” n'a aucune information pendant l'opération du système. Dans ce cas, il choisit un contrôle qui est fonction du temps. Ces contrôles sont appelés en “boucle ouverte” ou “open-loop” en anglais.
2. Le contrôleur connaît l'état du système à chaque instant. C'est le cas de l'observation (ou information) complète.
3. Le contrôleur a une connaissance partielle de l'état du système. C'est le cas de l'observation incomplète.

Pour les problèmes de contrôle déterministe, les contrôles peuvent être indifféremment choisis en boucle ouverte ou en boucle fermée (appelés alors “feedback”). Les contrôles en feedback ne donneront pas un plus petit minimum. En effet, l'état du système à tout instant t peut être déduit des données initiales et du contrôle appliqué jusqu'à l'instant t par la résolution de l'équation différentielle satisfaite par l'état du système. Donc l'observation de l'état courant du système à tout instant t ne donne pas davantage d'information que les données initiales.

Par contre, pour les problèmes de contrôle stochastique, à partir d'une donnée initiale et d'un contrôle, le système peut suivre différentes trajectoires. Dans le cas stochastique la trajectoire optimale dépend de l'information disponible au contrôleur à tout instant t .

Dans le cas de l'information complète, auquel nous nous restreindrons ici, la méthode de la Programmation Dynamique conduit à des équations aux dérivées partielles du deuxième ordre non linéaires (équations d'Hamilton-Jacobi-Bellman) dont la résolution permet d'obtenir la fonction valeur et des contrôles en boucle fermée. Dans le cas de l'observation incomplète, la méthode de la Programmation Dynamique conduit à des équations de dimension infinie.

5.2 Contrôle de diffusions

Soit (Ω, \mathcal{F}, P) un espace de Probabilité complet et soit $(\mathcal{F}_t)_{t \geq 0}$ une filtration. On considère des systèmes dynamiques dont l'état X_t est modélisé par des processus stochastiques de diffusion dans \mathbb{R}^n . Les variables de décision, appelés variables de contrôle, sont des processus stochastiques dont la valeur peut être décidée à tout instant en fonction des informations disponibles à cet instant, et ces variables de contrôle apparaissent dans l'équation d'évolution du système. On note P_t les contrôles et l'on suppose qu'ils sont à valeurs dans une partie fermée \mathcal{P}_{ad} de \mathbb{R}^k . Les contrôles P_t sont des processus adaptés à \mathcal{F}_t , selon la terminologie usuelle.

On suppose que l'état X_t du système dynamique est régi par une équation différentielle stochastique de la forme

$$dX_t = b(X_t, P_t)dt + \sigma(X_t, P_t)dW_t, \quad X_0 = x \quad (5.1)$$

où W_t est un processus de Wiener p -dimensionnel. Le terme de dérive $b(X_t, P_t)$ est un vecteur de dimension n et $\sigma(X_t, P_t)$ est une matrice $n \times p$.

L'équation (5.1) peut se réécrire sous la forme d'un système de n équations :

$$dX_t^i = b^i(X_t, P_t)dt + \sum_{j=1}^p \sigma_{ij}(X_t, P_t)dW_t^j, \quad X_0^i = x^i, \quad i = 1, \dots, n. \quad (5.2)$$

On suppose que b et σ sont continues sur $\mathbb{R}^n \times \mathcal{P}_{ad}$, avec $b(\cdot, P), \sigma(\cdot, P)$ de classe $C^1(\mathbb{R}^n)$ et qu'il existe une constante $C > 0$ telle que

$$|D_x b(x, P)| \leq C, \quad |D_x \sigma(x, P)| \leq C \quad (5.3)$$

$$|b(x, P)| \leq C(1 + |x| + |P|), \quad |\sigma(x, P)| \leq C(1 + |x| + |P|). \quad (5.4)$$

D_x désigne la dérivée en x . Sous ces hypothèses, l'équation (5.1) admet une solution unique et le processus X_t est bien défini (voir Théorème 3.3.1).

On suppose tout d'abord qu'il n'y a pas de contraintes sur l'état, c'est à dire que l'espace d'état est \mathbb{R}^n tout entier. Posons le problème d'optimisation. Il s'agit de déterminer la stratégie qui optimise un certain critère sur un horizon de gestion.

- Dans le cas d'un **horizon infini**, le processus X_t est stationnaire (b et σ sont indépendants du temps t) et on définit le coût actualisé par

$$J(x, P(\cdot)) = E \left(\int_0^\infty e^{-\lambda t} u(X_t, P_t) dt \mid X_0 = x \right) \quad (5.5)$$

où $\lambda > 0$ est le taux d'actualisation et la fonction $u : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^+$ est parfois appelé le coût instantané. On suppose que u satisfait une condition de croissance polynomiale, c'est à dire $\exists C$ et k tels que

$$|u(x, P)| \leq C(1 + |x|^k + |P|^k). \quad (5.6)$$

Le problème est de déterminer le processus $P(\cdot)$ qui minimise la fonctionnelle J et de calculer la fonction valeur

$$v(x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} J(x, P(\cdot)). \quad (5.7)$$

- Dans le cas d'un **horizon fini** T , les coefficients b et σ dans la dynamique du système peuvent dépendre du temps. On les suppose de classe C^1 par rapport au temps et tels que

$$\left| \frac{\partial b}{\partial t} \right|, \left| \frac{\partial \sigma}{\partial t} \right| \leq C. \quad (5.8)$$

Partant d'un état x à l'instant t , on définit, pour tout processus de contrôle $P(\cdot)$, le coût

$$J(t, x, P(\cdot)) = E \left(\int_t^T u(s, X_s, P_s) ds + g(X_T) | X_t = x \right), \quad (5.9)$$

où g est appelé le coût sur l'état final. On suppose g à croissance polynomiale. La fonction valeur dépend alors du temps :

$$v(t, x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} J(t, x, P(\cdot)) \quad (5.10)$$

Remarque : Lorsque l'on cherche à maximiser un gain (au lieu de minimiser un coût), alors on écrira que

$$v(x) = \max_{P(\cdot)} J(x, P(\cdot)) = - \min_{P(\cdot)} (-J(x, P(\cdot))),$$

ceci pour se conformer à la littérature où les problèmes d'optimisation sont en général écrits sous forme de minimisation.

5.3 Problème de temps d'arrêt optimal

Un des problèmes de contrôle les plus simples est celui du temps d'arrêt optimal pour lequel à tout instant t , seulement deux actions de contrôle sont possibles : arrêter le processus ou le laisser évoluer librement.

Definition 5.3.1. *On dit qu'une variable aléatoire positive τ est un temps d'arrêt admissible si c'est un \mathcal{F}_t temps d'arrêt.*

Soit $k(\cdot)$ et $g(\cdot)$ des fonctions régulières. Pour un temps d'arrêt admissible $t \leq \tau \leq T$, on définit le coût

$$J(t, x, \tau) = E \left(\int_t^\tau k(s, X_s) ds + g(\tau, X_\tau) | X_t = x \right)$$

et le coût optimal où l'infimum est considéré sur l'ensemble des temps d'arrêt admissibles :

$$V(t, x) = \inf_{t \leq \tau \leq T} J(t, x, \tau). \quad (5.11)$$

Exemple des Options Américaines : Un exemple typique pour illustrer le problème de temps d'arrêt optimal est l'évaluation des options américaines. Le détenteur d'une option américaine peut l'exercer à tout moment avant la date d'échéance T et le prix s'écrit souvent

$$V_t = \sup_{t \leq \tau \leq T} E \left(e^{-\int_t^\tau r(s, X_s) ds} f(X_\tau) | X_t \right). \quad (5.12)$$

Comme pour les options européennes, il n'y a pas de terme intégral mais seulement une contribution actualisée à l'état final.

5.4 Programmation Dynamique

Le Principe d'optimalité dû à Bellman [25](p.83) selon lequel "an optimal policy has the property that, whatever the initial state and control are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" est une des bases du contrôle optimal.

Il s'énonce ainsi pour un problème stationnaire :

Proposition 5.4.1. *La fonction valeur définie en (5.7) satisfait :*

$$v(x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} E \left(\int_0^\theta e^{-\lambda t} u(X_t, P_t) dt + e^{-\lambda \theta} v(X_\theta) | X_0 = x \right) \quad (5.13)$$

pour tout temps d'arrêt θ par rapport à la filtration \mathcal{F}_t .

Ce résultat permet, en particulier, de relier la valeur de v au point x avec ses valeurs en des points voisins, et donc d'établir l'équation satisfaite par v .

Nous nous contentons ici d'expliquer brièvement ce principe et nous référons à [32] pour la démonstration.

Supposons que l'on exerce un contrôle p_t sur l'intervalle de temps $[0, \theta]$. A l'instant θ , l'état du système est X_θ et nous pouvons l'observer. Supposons connue la politique optimale \tilde{p}_t à partir de l'instant θ , autrement dit, supposons connu le contrôle \tilde{p}_t , $t \geq \theta$, tel que

$$v(X_\theta) = E \left(\int_\theta^{+\infty} e^{-\lambda(t-\theta)} u(X_t, \tilde{p}_t) dt | X_\theta \right).$$

Considérons le contrôle P_t tel que

$$P_t = \begin{cases} p_t & \text{sur } [0, \theta] \\ \tilde{p}_t & \text{pour } t \geq \theta. \end{cases} \quad (5.14)$$

Le coût associé à ce contrôle est

$$J(x, P(\cdot)) = E \left(\int_0^\theta e^{-\lambda t} u(X_t, p_t) dt + e^{-\lambda \theta} v(X_\theta) | X_0 = x \right). \quad (5.15)$$

Le principe d'optimalité nous dit que si l'on choisit p_t sur $(0, \theta)$ de manière à minimiser (5.15), on obtient le coût optimal.

5.5 Equations d'Hamilton-Jacobi-Bellman

La fonction valeur est solution d'une équation aux dérivées partielles non-linéaire appelée équation d'Hamilton-Jacobi-Bellman (HJB) ou équation de Bellman ou encore équation de la Programmation Dynamique.

Notons A^P l'opérateur aux dérivées partielles du second ordre qui à une fonction de classe $\mathcal{C}^2(\mathbb{R}^n)$ associe :

$$A^P v = b(x, P) \cdot Dv(x) + \frac{1}{2} \text{tr} \sigma \sigma^T(x, P) D^2 v(x) \quad (5.16)$$

$$= \sum_{i,j=1}^n a_{ij}(x, P) \frac{\partial^2 v}{\partial x_i \partial x_j} + \sum_{j=1}^n b_j(x, P) \frac{\partial v}{\partial x_j} \quad (5.17)$$

où $a = \frac{1}{2} \sigma \sigma^T$, c'est à dire $a_{ij} = \frac{1}{2} \sum_{k=1}^p \sigma_{ik} \sigma_{jk}$.

5.5.1 Equation parabolique

Dans le cas d'un horizon fini, sous les hypothèses de régularité (5.3, 5.4, 5.6, 5.8) de b, σ, u , la fonction valeur v définie par (5.10) est solution de l'équation d'HJB parabolique :

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \min_{P \in \mathcal{P}_{ad}} (A^P v(t, x) + u(t, x, P)) = 0 & \text{dans } [0, T[\times \mathbb{R}^n, \\ v(T, x) = g(x). \end{cases} \quad (5.18)$$

L'existence d'une solution forte n'est là aussi assurée que sous l'hypothèse d'uniforme ellipticité (5.22).

Si l'espace d'état est un ouvert \mathcal{O} de \mathbb{R}^n , la fonction valeur

$$v(t, x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} E \left(\int_t^{\tau \wedge T} u(s, X_s, P_s) ds + f(\tau, X_\tau) 1_{\tau < T} + g(X_T) 1_{T < \tau} \mid X_t = x \right)$$

où le coût f est à croissance polynomiale est solution de l'équation d'H.J.B

$$\begin{cases} \frac{\partial v}{\partial t}(t, x) + \min_{P \in \mathcal{P}_{ad}} (A^P v(t, x) + u(t, x, P)) = 0 & \text{dans } [0, T[\times \mathcal{O}, \\ v(T, x) = g(x). \end{cases} \quad (5.19)$$

avec la condition au limite de Dirichlet

$$v(t, x) = f(t, x) \quad \text{sur } [0, T[\times \partial \mathcal{O}. \quad (5.20)$$

Si le processus est réfléchi sur la frontière, la fonction valeur

$$v(t, x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} E \left(\int_t^T u(s, X_s, P_s) ds + \int_t^T f(s, X_s) d\xi_s + g(X_T) \mid X_t = x \right)$$

est solution de l'équation d'H.J.B (5.19) avec la condition au limite de Neumann

$$\frac{\partial v(t, x)}{\partial n} = f(t, x) \quad \text{sur } [0, T[\times \partial \mathcal{O}. \quad (5.21)$$

5.5.2 Equation elliptique

On se place dans le cas d'un horizon infini et on suppose que les fonctions b, σ , et u satisfont les hypothèses de régularité (5.3, 5.4, 5.6). On suppose de plus qu'il existe $c > 0$ tel que

$$\sum_{i=1}^n a_{ij}(x, P) \eta_i \eta_j \geq c |\eta|^2, \forall x \in \mathcal{O}, \eta \in \mathbb{R}^n, P \in \mathcal{P}_{ad}. \quad (5.22)$$

Alors, la fonction valeur v définie par (5.7) est solution de l'équation d'H.J.B. elliptique :

$$-\lambda v(x) + \min_{P \in \mathcal{P}_{ad}} (A^P v(x) + u(x, P)) = 0 \quad \text{dans } \mathbb{R}^n. \quad (5.23)$$

Si \mathcal{P}_{ad} est compact, le minimum dans (5.23) est toujours atteint. Sinon des conditions supplémentaires sont nécessaires.

Si l'espace d'état est un ouvert \mathcal{O} de \mathbb{R}^n et que le processus est arrêté sur la frontière, la fonction valeur est définie par

$$v(x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} J(x, P)$$

avec

$$J(x, P) = E \left(\int_0^\tau e^{-\lambda t} u(X_t, P_t) dt + e^{-\lambda \tau} f(X_\tau) | X_0 = x \right) \quad (5.24)$$

où

$$\tau(\omega) = \inf\{t, X_t(\omega) \notin \mathcal{O}\}$$

est le premier instant où le processus X_t sort du domaine \mathcal{O} et f est le coût d'arrêt. La fonction $v(x)$ satisfait l'équation d'H.J.B

$$-\lambda v(x) + \min_{P \in \mathcal{P}_{ad}} (A^P v(x) + u(x, P)) = 0 \quad \text{dans } \mathcal{O}. \quad (5.25)$$

avec la condition au limite de Dirichlet

$$v(x) = f(x) \quad \text{sur } \partial\mathcal{O}. \quad (5.26)$$

Si le processus est réfléchi sur la frontière, l'équation d'évolution du processus s'écrit

$$dX_t = b(X_t, P_t)dt + \sigma(X_t, P_t)dW_t - n_{X_t}d\xi_t \quad (5.27)$$

ξ_t est un processus strictement croissant lorsque $X_t \in \partial\mathcal{O}$ et $d\xi_t = 0$ si $X_t \in \mathcal{O}$, n_x est la normale extérieure à la frontière $\partial\mathcal{O}$ de \mathcal{O} en x .

La fonction valeur, définie par

$$v(x) = \min_{P(\cdot) \in \mathcal{P}_{ad}} E \left(\int_0^{+\infty} e^{-\lambda t} u(X_t, P_t) dt + \int_0^{+\infty} e^{-\lambda t} f(X_t) d\xi_t | X_0 = x \right)$$

où f est le coût associé à la réflexion du processus, satisfait l'équation d'H.J.B (5.25) avec la condition au limite de Neumann

$$\frac{\partial v}{\partial n} = f \quad \text{sur } \partial\mathcal{O}. \quad (5.28)$$

La condition (5.22) est la condition d'uniforme ellipticité. Elle implique l'existence d'une solution forte de l'équation d'HJB, c'est à dire une fonction de classe \mathcal{D}^2 qui vérifie l'équation en tout point. L'unicité de la solution est assurée dans le cas d'un ouvert \mathcal{O} borné. Il n'y a pas unicité quand l'équation est définie dans un ouvert non borné. Parmi toutes les solutions, la fonction valeur se trouve alors être celle qui ne croit pas trop rapidement lorsque $|x| \rightarrow \infty$.

Lorsque (5.22) n'est pas vérifiée, l'équation est dite dégénérée et dans ce cas, il n'y a pas de solution forte à l'équation d'HJB. Il faut alors considérer une notion de solution faible comme les solutions de viscosité (voir paragraphe 5.6).

La résolution de l'équation d'HJB permet de définir en tout point x un contrôle optimal $P(x)$ qui ne dépend que de x : c'est un contrôle en "feedback", c'est à dire que c'est le contrôle que l'on doit utiliser lorsqu'on se trouve au point x . Il ne dépend donc que de l'état du système et non pas de la variable t .

5.5.3 Interprétation probabiliste

Donnons ici une dérivation formelle de l'équation d'HJB dans le cas elliptique (5.7). On admet pour cela le *Principe d'optimalité de la Programmation dynamique* et l'on suppose toute la régularité nécessaire sur la fonction valeur pour pouvoir appliquer la formule d'Ito ordinaire.

D'après la formule de Ito, on a, sachant que X_t évolue selon l'équation (5.1)

$$\begin{aligned} e^{-\lambda\theta}v(X_\theta) &= v(x) + \int_0^\theta (-\lambda e^{-\lambda s}v(X_s) + e^{-\lambda s}Dv(X_s).b(X_s, P_s))ds \\ &+ \frac{1}{2} \int_0^\theta e^{-\lambda s} \text{tr} D^2v(X_s) \sigma \sigma^T(X_s, P_s) ds \\ &+ \int_0^\theta e^{-\lambda s} Dv(X_s). \sigma(X_s, P_s) dW_s. \end{aligned}$$

Prenant l'espérance, on obtient, en utilisant que l'espérance de l'intégrale stochastique est nulle :

$$\begin{aligned} Ee^{-\lambda\theta}v(X_\theta) &= v(x) + E\left(\int_0^\theta (-\lambda e^{-\lambda s}v(X_s) + e^{-\lambda s}Dv(X_s).b(X_s, P_s))ds\right) \\ &+ \frac{1}{2} \int_0^\theta e^{-\lambda s} \text{tr} D^2v(X_s) \sigma \sigma^T(X_s, P_s) ds. \end{aligned}$$

On reporte cette expression dans l'équation (5.13), on fait tendre θ vers 0 et on fait un développement au premier ordre en θ . On obtient :

$$\begin{aligned} v(x) &= \min_{P \in \mathcal{P}_{ad}} \{ \theta u(x, P) + v(x) + \theta(-\lambda v(x) + Dv(x).b(x, P)) \\ &+ \frac{1}{2} \text{tr} D^2v(x) \sigma \sigma^T(x, P) \} \end{aligned}$$

soit l'équation d'HJB :

$$-\lambda v + \min_{P \in \mathbb{P}_{ad}} \{ b(x, P).Dv(x) + \frac{1}{2} \text{tr} \sigma \sigma^T(x, P) D^2v(x) + u(x, P) \} = 0.$$

Dans la pratique, soit l'on démontre le principe d'optimalité (5.13) et on en déduit que la fonction valeur est bien solution de l'équation d'HJB, soit on établit de manière formelle l'équation d'HJB associée au problème et on vérifie par un théorème dit "Théorème de Vérification" qu'une solution v de cette équation, possédant certaines conditions de régularité et un bon comportement à l'infini si l'espace d'état est non borné (par exemple une croissance polynomiale), est bien la fonction valeur du problème de contrôle stochastique. (voir Fleming-Rishel [?] et Fleming-Soner [32]). Par exemple dans le cas elliptique, le théorème de vérification s'énonce comme suit :

Théorème 5.5.1. *Soit v une solution de l'équation d'HJB (5.25, 5.26) de classe $C^2(\mathcal{O})$ et à croissance polynomiale sur $\bar{\mathcal{O}}$. Alors, $\forall x \in \mathcal{O}$*

- $v(x) \leq J(x, P)$ pour tout processus de contrôle admissible $P(\cdot)$ tel que

$$\lim_{t \rightarrow \infty} e^{-\lambda t} E_x 1_{\tau \geq t} v(X(t)) \leq 0 \quad (5.29)$$

– Si $P^*(\cdot)$ est un contrôle admissible tel que pour tout x ,

$$\min_{P \in \mathcal{P}_{ad}} (A^P v(x) + u(x, P)) = A^{P^*(x)} v(x) + u(x, P^*(x))$$

et

$$\lim_{t \rightarrow \infty} e^{-\lambda t} E_x 1_{\tau^* \geq t} v(X^*(t)) = 0 \quad (5.30)$$

où $X^*(t)$ désigne la trajectoire associée au contrôle P^* , alors $v(x) = J(x, P^*)$ où J est défini en (5.24) et P^* est optimal.

Si v est bornée sur $\bar{\mathcal{O}}$, (en particulier lorsque \mathcal{O} est borné et v est continu sur $\bar{\mathcal{O}}$), les conditions (5.29,5.30) sont automatiquement satisfaites et on a le corollaire :

Corollaire 5.5.1. *Soit v une solution de l'équation d'HJB (5.25,5.26) de classe $C^2(\mathcal{O})$ et bornée sur $\bar{\mathcal{O}}$. Supposons de plus que $\lambda > 0$ ou $\tau < \infty$ avec une probabilité 1 pour tout contrôle admissible $P(\cdot)$. Alors v est la fonction valeur du problème de contrôle stochastique correspondant.*

5.5.4 Inéquation variationnelle associée au problème de temps d'arrêt optimal

Le coût optimal $V(t, x) = \inf_{t \leq \tau \leq T} E(\int_t^\tau k(s, X_s) ds + g(\tau, X_\tau) | X_t = x)$ satisfait l'inéquation variationnelle :

$$\min\{AV + \frac{\partial V}{\partial t} + k(x), g(t, x) - V(t, x)\} = 0 \text{ dans } [0, T] \times \mathbb{R}^n, \quad (5.31)$$

où A est l'opérateur associé à la diffusion X_t . Soit B_t l'ensemble d'arrêt optimal, i.e. le processus s'arrête dès que B_t est atteint. Alors $V(t, x) \leq g(t, x)$ et $V(t, x) = g(t, x)$ seulement dans B_t . L'ensemble B_t est à déterminer.

Donnons une preuve simple et heuristique de (5.31) basée sur le principe d'optimalité de la Programmation dynamique :

Supposons que le processus soit au point x à l'instant t . On a le choix d'arrêter le système ou de le laisser évoluer.

– si on arrête, le coût que l'on doit payer est $g(t, x)$. Donc

$$V(t, x) \leq g(t, x). \quad (5.32)$$

– le Principe d'optimalité nous permet d'exprimer le prix à payer pour laisser évoluer le système pendant un intervalle de temps δ et utiliser ensuite les décisions optimales :

$$E_{xt}[\int_t^{t+\delta} k(X_s) ds + V(t + \delta, X(t + \delta))]$$

Lorsque δ tend vers 0, utilisant la formule d'Ito et la propriété $E(W(t + \delta) - W(t))^2 = \delta$ des processus de Wiener, on obtient que cette quantité tend vers :

$$\delta k + \delta AV + \delta \frac{\partial V}{\partial t} + V.$$

Le coût optimal est majorée par cette quantité ce qui conduit à :

$$0 \leq AV + \frac{\partial V}{\partial t} + k \quad (5.33)$$

Comme l'une des deux décisions est optimale, l'une des inégalités (5.32) où (5.33) est une égalité, d'où l'équation (5.31). (voir [1], chapitre 3 section 2) pour une démonstration détaillée.

le cas actualisé. Si le coût est actualisé :

$$J(t, x, \tau) = E\left(\int_t^\tau e^{-\int_t^s \beta(\theta, X(\theta))d\theta} k(s, X_s) ds + e^{-\int_t^\tau \beta(X(\theta, \theta))d\theta} g(\tau, X_\tau) | X_t = x\right)$$

où $\beta(\cdot)$ est une fonction positive continue et bornée, on obtient la même équation (5.31) pour le coût optimal mais l'opérateur doit être remplacé par $A - \beta I$ où I est l'identité.

Par exemple, le prix d'une option américaine s'écrit

$$V_t = \sup_{t \leq \tau \leq T} E(e^{-\int_t^\tau r(s, X_s) ds} f(X_\tau) | \mathcal{F}_t).$$

Utilisant tout d'abord la propriété markovienne de X_t , on écrit $V_t = u(t, X_t)$ avec

$$u(t, x) = \sup_{t \leq \tau \leq T} E(e^{-\int_t^\tau r(s, X_s) ds} f(X_\tau) | X_t = x).$$

La fonction $u(t, x)$ est solution de l'inéquation variationnelle :

$$\begin{cases} \max\left(\frac{\partial V}{\partial t} + Au - ru, u(t, x) - V(t, x)\right) = 0, \\ u(T, x) = f(x) \end{cases} \quad (5.34)$$

On réfère à [11](section 3) pour une démonstration directe. Par exemple dans le modèle de Black et Scholes, le prix de l'option américaine

$$V_t = \sup_{t \leq \tau \leq T} E(e^{-\int_t^\tau r(s, X_s) ds} f(X_\tau) | X_t)$$

est solution de (5.34) avec

$$Au(t, x) = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + rx \frac{\partial u}{\partial x}.$$

5.6 Solutions de viscosité

5.6.1 Introduction

Pour démontrer le théorème de vérification qui identifie la solution de l'équation d'HJB avec celle du problème de contrôle stochastique, on utilise la formule d'Ito ce qui requiert en général que la solution v de l'équation d'HJB soit de classe \mathcal{D}^2 .

Or, il arrive souvent que l'équation d'HJB n'ait pas de solution forte (dans le cas dégénéré par exemple). Dans ce cas, il faut définir une notion de solution faible.

L'approche usuelle, pour les équations elliptiques du deuxième ordre, est *l'approche variationnelle*. Cette approche est particulièrement bien adaptée aux problèmes linéaires ou aux problèmes où le terme linéaire est dominant.

Une autre approche, plus récente, est l'introduction des solutions de viscosité. Cette notion a été introduite pour résoudre des équations du premier ordre, et elle s'étend de façon naturelle aux équations elliptiques (ou paraboliques) du deuxième ordre (voir [23, 21, 8]). La notion de solution de viscosité est bien adaptée pour des équations fortement non linéaires elliptiques dégénérées ou non.

On montre que si la fonction valeur est continue (pas nécessairement de classe C^2), et satisfait le principe d'optimalité de la programmation dynamique, alors elle est solution de viscosité de l'équation d'HJB correspondante (voir [32]).

Solutions de viscosité d'équations elliptiques du deuxième ordre :

On considère des équations elliptiques du type

$$F(D^2v, Dv, v, x) = 0 \quad \text{dans } \mathcal{O} \quad (5.35)$$

où F est une fonction numérique continue sur $S^n \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{O}$, S^n est l'espace des matrices symétriques $n \times n$, \mathcal{O} est un ouvert de \mathbb{R}^n . Si v est régulière, on note Dv son gradient ($Dv = (\frac{\partial v}{\partial x_1}, \dots, \frac{\partial v}{\partial x_n})$), et D^2v sa matrice hessienne ($D^2v = (\frac{\partial^2 v}{\partial x_i \partial x_j})_{i,j}$). L'ellipticité de (5.35) est exprimée par

$$F(A, p, v, x) \geq F(B, p, v, x) \quad \text{si } A \geq B, \quad A, B \in S^n, \quad p \in \mathbb{R}^n, \quad v \in \mathbb{R}, \quad x \in \mathcal{O}. \quad (5.36)$$

On rappelle que

$$A \geq B \iff (Ap, p) \geq (Bp, p) \quad \forall p \in \mathbb{R}^n$$

où (\cdot, \cdot) désigne le produit scalaire usuel dans \mathbb{R}^n .

Un cas particulier de (5.35) est

$$F(X, p, v, x) = \max_{P \in P_{ad}} \left\{ \sum_{i,j=1}^n a_{ij}(x, P) X_{ij} + \sum_{i=1}^n b_i(x, P) p_i - \lambda v + u(x, P) \right\}$$

où la condition d'ellipticité (5.36) est satisfaite lorsque la matrice $(a_{ij}(x, P))_{i,j}$ est symétrique semi-définie positive dans $\mathcal{O} \times P_{ad}$, c'est à dire :

$$\sum_{i=1}^n a_{ij}(x, P) \eta_i \eta_j \geq 0, \quad \forall x \in \mathcal{O}, \quad \eta \in \mathbb{R}^n, \quad P \in P_{ad}.$$

S'il existe $c > 0$ tel que

$$\sum_{i=1}^n a_{ij}(x, P) \eta_i \eta_j \geq c |\eta|^2, \quad \forall x \in \mathcal{O}, \quad \eta \in \mathbb{R}^n, \quad P \in P_{ad} \quad (5.37)$$

l'équation est uniformément elliptique.

Les équations d'HJB elliptiques de type (5.23) rentrent dans ce formalisme.

La notion de solution de viscosité repose sur le théorème suivant qui peut être vu comme une variante du *Principe du Maximum* :

Théorème 5.6.1. $v \in \mathcal{D}^2(\mathcal{O})$ est solution classique de (5.35) si et seulement si :

$\forall \phi \in \mathcal{D}^2(\mathcal{O})$, si x_0 est un point de maximum local de $v - \phi$, alors :

$$F(D^2\phi(x_0), D\phi(x_0), v(x_0), x_0) \geq 0$$

et

$\forall \phi \in \mathcal{D}^2(\mathcal{O})$, si x_0 est un point de minimum local de $v - \phi$, alors :

$$F(D^2\phi(x_0), D\phi(x_0), v(x_0), x_0) \leq 0.$$

Ce théorème fournit une définition de la solution classique qui a l'avantage de ne faire intervenir aucune régularité de v et qui va nous permettre de définir les solutions de viscosité. Donnons tout d'abord la preuve de ce théorème.

Preuve : Soit $v \in \mathcal{D}^2(\mathcal{O})$ une solution classique de (5.35). Si $\phi \in \mathcal{D}^2(\mathcal{O})$, et si x_0 est un point de maximum local de $v - \phi$, on a

$$Dv(x_0) = D\phi(x_0) \quad \text{et} \quad D^2v(x_0) \leq D^2\phi(x_0).$$

En utilisant l'ellipticité de l'équation, on obtient :

$$F(D^2\phi(x_0), D\phi(x_0), v(x_0), x_0) \geq F(D^2v(x_0), Dv(x_0), v(x_0), x_0) = 0,$$

et la première propriété est démontrée. La deuxième s'obtient de manière analogue.

Réciproquement, si v est de classe \mathcal{D}^2 , on peut prendre $\phi = v$ dans les deux propriétés et donc, comme tout point x_0 de \mathcal{O} est à la fois un point de maximum et de minimum local de $v - v$, $F(D^2v(x_0), Dv(x_0), v(x_0), x_0)$ est à la fois négatif et positif en tout point de \mathcal{O} . Donc $F(D^2v(x), Dv(x), v(x), x) = 0$ dans \mathcal{O} et la preuve est complète. \square

On peut donner à présent la définition de solution de viscosité.

Definition 5.6.1. $v \in \mathcal{D}(\mathcal{O})$ est solution de viscosité de (5.35) si et seulement si :

$\forall \phi \in \mathcal{D}^2(\mathcal{O})$, si x_0 est un point de maximum local de $v - \phi$, alors :

$$F(D^2\phi(x_0), D\phi(x_0), v(x_0), x_0) \geq 0 \tag{5.38}$$

et $\forall \phi \in \mathcal{D}^2(\mathcal{O})$, si x_0 est un point de minimum local de $v - \phi$, alors :

$$F(D^2\phi(x_0), D\phi(x_0), v(x_0), x_0) \leq 0. \tag{5.39}$$

Si v ne vérifie que (5.38) (resp. (5.39)), on dit que v est une sous-solution de viscosité (resp. sur-solution de viscosité).

Le rôle de l'ellipticité est central dans cette définition. Remarquons que les équations paraboliques sont un cas particulier des équations elliptiques telles que nous les avons définies : on remplace simplement la variable x par la variable (x, t) , Du et D^2u représentant alors le gradient et le hessien par rapport à la nouvelle variable (x, t) .

L'espace naturel pour chercher des solutions est l'espace des fonctions continues.

5.6.2 Exemple d'applications

5.7 Application à l'optimisation dynamique de portefeuille

On s'intéresse ici à l'application du contrôle stochastique à l'optimisation dynamique de portefeuilles en temps continu. On considère un modèle de type Black-Scholes pour la modélisation des cours des actifs.

Dans le cas où les transactions sur les actifs financiers sont gratuites, le problème peut être résolu explicitement. Le cas des coûts de transaction proportionnels au montant de la transaction fait appel à la théorie des contrôles singuliers.

Soit un espace de probabilité complet (Ω, \mathcal{F}, P) muni d'une filtration $(\mathcal{F}_t)_{t \geq 0}$ représentant les informations disponibles à l'instant t .

On considère un investisseur possédant un compte en banque, aussi appelé actif sans risque, (de type "caisse d'épargne") rapportant un intérêt fixe $r > 0$ et n actifs risqués de prix $S_i(t)$, $i = 1 \dots n$ à l'instant t . Les cours des actifs risqués sont modélisés par des processus de diffusion

$$dS_i(t) = S_i(t)(\alpha_i dt + \sigma_i dW_i(t)), \quad i = 1, \dots, n \quad (5.40)$$

où α_i et σ_i sont 2 constantes représentant respectivement le rendement instantané moyen et la volatilité de l'actif i , et $W_i(t)$ est un mouvement Brownien standard. En général $\alpha_i > r$ et les actions qui ont le meilleur rendement moyen ont également la plus grande volatilité. On note

- $s_0(t)$: capital détenu dans le compte en banque à l'instant t
- $s_i(t)$: capital investi dans le $i^{\text{ème}}$ actif risqué à l'instant t

et $s(t) = (s_0(t), s_1(t), \dots, s_n(t))$ le vecteur de \mathbb{R}^{n+1} représentant la position de l'investisseur à l'instant t .

On pourra considérer des modèles avec ou sans consommation. Dans le premier cas, l'investisseur consomme avec un taux $c(t)$ en retirant de l'argent du compte en banque. On suppose que les augmentations du capital proviennent des gains dûs à la variation des cours. En l'absence de transactions entre les comptes, les quantités d'argent dans les comptes évoluent selon les équations :

$$\begin{cases} ds_0(t) = (rs_0(t) - c(t))dt, \\ ds_i(t) = \alpha_i s_i(t)dt + \sigma_i s_i(t)dW_i(t), \quad i = 1, \dots, n, \end{cases} \quad (5.41)$$

avec les valeurs initiales $s_i(0) = x_i$, $i = 0, \dots, n$.

Dans le modèle sans consommation, en l'absence de transactions entre les comptes, la quantité d'argent dans le compte en banque évolue donc selon l'équation :

$$ds_0(t) = rs_0(t)dt. \quad (5.42)$$

On suppose que l'investisseur a la possibilité d'acheter ou de vendre l'actif risqué i ($i = 1, \dots, n$) à tout instant. Les transactions se font par l'intermédiaire du compte en banque et sont instantanées.

Dans le cas d'un horizon de gestion infini, l'objectif de l'investisseur est de maximiser une fonction d'utilité de la consommation de la forme

$$E \int_0^{\infty} e^{-\delta t} u(c(t)) dt \quad (5.43)$$

où E représente l'espérance, $\delta > 0$ est le taux d'actualisation et u est une fonction d'utilité.

Dans le cas d'un horizon de gestion fini T , l'objectif est de maximiser une fonction de la valeur finale du portefeuille

$$E u \rho(T) \quad (5.44)$$

où $\rho(t)$ est la fortune totale de l'investisseur à l'instant t définie par

$$\rho(t) = s_0(t) + \sum_{i=1}^n s_i(t). \quad (5.45)$$

Les fonctions d'utilité HARA. Les fonctions HARA (hyperbolic absolute risk aversion) sont les fonctions d'utilité $u(x)$ définies pour $x \geq 0$ par

$$u(x) = \frac{1}{\gamma}(x + \eta)^\gamma, \quad \gamma < 1, \gamma \neq 0, \eta \geq 0 \quad (5.46)$$

ou

$$u(x) = \log(x + \eta), \quad \eta \geq 0. \quad (5.47)$$

Les fonctions HARA permettent de modéliser le comportement de l'investisseur par rapport au risque. Le coefficient $1 - \gamma$ est l'aversion au risque relative. L'investisseur prend le maximum de risque pour $\gamma = 1$ (tout l'argent est investi dans l'action qui a le meilleur rendement et la plus grande volatilité) et le risque pris décroît avec γ (voir exercice).

La fonction (5.47) est appelée fonction HARA de paramètre 0. En effet elle peut, grâce au théorème asymptotique suivant, s'interpréter comme le cas limite $\gamma = 0$:

Théorème 5.7.1. *Pour tout variable aléatoire positive f et pour toute mesure μ , on a l'égalité :*

$$\lim_{\gamma \rightarrow 0} \left(\int f^\gamma d\mu \right)^{\frac{1}{\gamma}} = \exp \left(\int \log f d\mu \right). \quad (5.48)$$

En particulier

$$\lim_{\gamma \rightarrow 0} (E \rho^\gamma(T))^{\frac{1}{\gamma}} = \exp(E \log \rho(T)) \quad (5.49)$$

et

$$\lim_{\gamma \rightarrow 0} \left(\delta E \int_0^\infty e^{-\delta t} c^\gamma(t) dt \right)^{\frac{1}{\gamma}} = \exp \left(\delta E \int_0^\infty e^{-\delta t} \log c(t) dt \right). \quad (5.50)$$

Supposons que les transferts d'argent entre comptes n'entraînent pas de coût de transaction. Ce problème est appelé "le problème de Merton" car il fut formulé à l'origine par Merton en 1971. Considérons un problème en horizon infini avec consommation. On considère comme variable d'état la fortune totale $\rho(t)$ de l'investisseur définie en (5.45) et comme variables de contrôle

- $c(t)$: le taux de consommation
- et $y_i(t) = \frac{s_i(t)}{\rho(t)}$: la proportion de capital investi dans l'actif i , $\forall i = 1 \dots n$.

Plus précisément, une politique d'investissement-consommation est un ensemble $P = (c(t), y_i(t)_{i=1, \dots, n})$ de processus \mathcal{F}_t -adaptés t.q.

- $c(t) \geq 0$, $\int_0^t c(s) ds < \infty$, p.p. t ,
- $|y_i(t)| \leq K$, p.p. t , où K peut varier d'une politique à une autre.

La dynamique du système est donnée par l'équation différentielle stochastique

$$\begin{cases} d\rho(t) = (r\rho(t) + \rho(t) \sum_{i=1}^n (\alpha_i - r)y_i(t) - c(t))dt + \rho(t) \sum_{i=1}^n \sigma_i y_i(t) dW_i(t), \\ \rho(0) = \sum_{i=0}^n x_i \equiv \rho. \end{cases} \quad (5.51)$$

Soit ρ la fortune initiale ; une politique P est dite *admissible* si

$$\rho(t) \geq 0, \quad \forall t \geq 0$$

ce qui signifie que la quantité d'argent totale $\rho(t)$, solution de (5.51) doit rester positive. Les quantités $s_0(t)$ et $s_i(t)$, ($i = 1, \dots, n$) peuvent, elles, prendre des valeurs négatives. Dire que $s_0(t) < 0$ signifie que l'on a emprunté la quantité $|s_0(t)|$ sur le marché des placements sans risques. Dire que $s_i(t) < 0$ pour un $i > 1$ signifie qu'on a des dettes libellées en actifs risqués par suite de ventes à découvert.

Notons \mathcal{P}_ρ l'ensemble des politiques admissibles pour un état initial ρ . L'objectif est de maximiser sur l'ensemble des politiques admissibles une fonction d'utilité de la consommation de la forme

$$J(\rho, P) = E_\rho \int_0^\infty e^{-\delta t} u(c(t)) dt \quad (5.52)$$

où E_ρ représente l'espérance conditionnelle à l'état initial ρ , $\delta > 0$ est le taux d'actualisation et $u(c)$ est une fonction d'utilité définie par

$$u(c) = \frac{c^\gamma}{\gamma}, \quad \gamma < 1, \gamma \neq 0. \quad (5.53)$$

On définit la fonction valeur

$$V(\rho) = \sup_{P \in \mathcal{P}_\rho} J(\rho, P) \quad (5.54)$$

Remarque 5.7.1. – Soit $\tau = \inf\{t \geq 0, \rho(t) = 0\}$ le temps de sortie de \mathbb{R}^{+*} . Une politique est admissible si $c(t) = 0$ p.s. (presque sûrement) pour $t \geq \tau$ et donc $\rho(t) = 0$, $\forall t \geq \tau$. Cela rend compte de l'absence d'opportunité d'arbitrage, c'est à dire l'impossibilité de réaliser un profit sans prendre de risque. Une stratégie d'arbitrage pourrait faire passer une fortune initiale nulle à une valeur non nulle.

– Si $\gamma > 0$, on peut écrire

$$V(\rho) = \sup_{P \in \mathcal{P}} E_\rho \int_0^\tau e^{-\delta t} u(c(t)) dt \quad (5.55)$$

où \mathcal{P} représente l'ensemble de toutes les politiques. En effet pour toute politique P admissible, on a $J(\rho, P) = E_\rho \int_0^\tau e^{-\delta t} u(c(t)) dt$ puisque $u(c(t)) = 0$ pour $t \geq \tau$.

Si P est quelconque, on peut construire une politique admissible P_{ad} qui coïncide avec la politique P jusqu'à τ , qui fait sauter le processus $s(t)$ à l'origine à l'instant τ et telle que $c(t) = 0$ pour $t \geq \tau$.

– Si $\gamma < 0$, on peut se restreindre à considérer les politiques telles que $\rho(t) > 0 \forall t$. En effet, s'il existe $\tau \geq 0$ tel que $\rho(\tau) = 0$ alors $c(t) = 0$ pour $t \geq \tau$ et $u(c(t)) = -\infty$. Donc si pour une certaine politique P , $\rho(t)$ atteint 0, $J(\rho, P) = -\infty$.

La fonction valeur V définie en (5.54) est croissante; de plus V est concave car la dynamique est linéaire et la fonction u est concave.

Le problème de Merton peut se résoudre explicitement - fait remarquable pour un problème de contrôle stochastique - par la résolution analytique de l'équation d'HJB associée. On a le théorème :

Théorème 5.7.2. *On suppose*

$$\delta > \gamma \left(r + \frac{1}{2(1-\gamma)} \sum_{i=1}^n \left(\frac{\alpha_i - r}{\sigma_i} \right)^2 \right) \quad (5.56)$$

et l'on note

$$C = \frac{1}{1-\gamma} \left(\delta - \gamma \left(r + \frac{1}{2(1-\gamma)} \sum_{i=1}^n \left(\frac{\alpha_i - r}{\sigma_i} \right)^2 \right) \right). \quad (5.57)$$

Alors, la fonction valeur V est égale à

$$V(\rho) = \frac{1}{\gamma} C^{\gamma-1} \rho^\gamma, \quad (5.58)$$

la consommation optimale est

$$c^*(\rho) = C\rho$$

et la politique optimale d'investissement est

$$y_i^*(\rho) \equiv y_i^* = \frac{\alpha_i - r}{\sigma_i^2(1-\gamma)}. \quad (5.59)$$

Preuve : L'équation d'HJB correspondant au problème s'écrit

$$-\delta V(\rho) + \max_{c \geq 0, y_i} \{AV(\rho) + u(c)\} = 0 \quad \text{pour } \rho > 0 \quad (5.60)$$

avec

$$AV(\rho) = (r\rho + \rho \sum_{i=1}^n (\alpha_i - r)y_i - c)V'(\rho) + \frac{1}{2}\rho^2 \sum_{i=1}^n \sigma_i^2 y_i^2 V''(\rho) \quad (5.61)$$

où $V'(\rho) = \frac{dV}{d\rho}$ et $V''(\rho) = \frac{d^2V}{d\rho^2}$.

- Si $\gamma > 0$, on la condition au limite

$$V(0) = 0 \quad (5.62)$$

qui s'explique par le fait que si le capital initial est nul, alors la consommation doit être nulle pour tout $t \geq 0$ et donc $u(c(t)) = 0 \forall t \geq 0$.

- Si $\gamma < 0$,

$$\lim_{\rho \rightarrow 0} V(\rho) = -\infty. \quad (5.63)$$

On peut séparer les "max" dans l'équation (5.60) qui devient :

$$-\delta V(\rho) + u^*(V'(\rho)) + r\rho V'(\rho) + \rho \sum_{i=1}^n \max_{y_i} (\alpha_i - r)y_i V'(\rho) + \frac{1}{2}\rho \sigma_i^2 y_i^2 V''(\rho) = 0 \quad (5.64)$$

où u^* est la transformée de Fenchel de u définie par

$$\begin{aligned} u^*(\nu) &= \max_{c \geq 0} (-c\nu + u(c)) \\ &= \left(\frac{1}{\gamma} - 1 \right) \nu^{\frac{\gamma}{\gamma-1}} \quad \text{si } \nu \geq 0. \end{aligned} \quad (5.65)$$

Si $\nu < 0$, $u^*(\nu)$ est égale à $+\infty$. Pour assurer une consommation finie, il faut donc imposer $V'(\rho) \geq 0$, $\forall \rho \geq 0$. La consommation optimale vaut alors

$$c^*(\rho) = V'(\rho)^{\frac{-1}{\gamma-1}}.$$

Considérons à présent le maximum en y_i . Si $V'' < 0$, c'est à dire si V est concave, il est atteint pour tout $\rho > 0$ pour

$$y_i^*(\rho) = \frac{-(\alpha_i - r)V'(\rho)}{\rho\sigma_i^2 V''(\rho)}.$$

En reportant dans (5.64), on obtient

$$-\delta V + r\rho V' - \sum_{i=1}^n \frac{(\alpha_i - r)^2 V'^2}{2\sigma_i^2 V''} + \frac{1-\gamma}{\gamma} (V')^{\frac{-\gamma}{1-\gamma}} = 0. \quad (5.66)$$

On cherche une solution de (5.66) concave, croissante et vérifiant les conditions aux limites (5.62) ou (5.63) selon le signe de γ . On la cherche de la forme :

$$V(\rho) = \lambda\rho^\gamma \quad (5.67)$$

avec $\lambda\gamma > 0$. En reportant dans (5.66), on obtient

$$\lambda = \frac{1}{\gamma} C^{\gamma-1}$$

où C est défini en (5.57).

La dernière étape consiste à s'assurer, par le théorème de vérification, que la solution V définie en (5.58) est bien la solution du problème d'optimisation (voir [32] pages 175-176 et [20]). Pour cela, on laisse au soin du lecteur de vérifier, en appliquant la formule d'Ito à $V(x)e^{-\delta t}$ et en regardant séparément les cas $\gamma < 0$ et $0 < \gamma < 1$, les 2 points suivants :

- $V(x) \geq J(\rho, P)$ pour toute politique admissible P
- $V(x) = J(\rho, P^*)$ avec $P^* = (c^*, y_i^*)$.

□

Analyse des résultats

La politique optimale, donnée par le théorème, consiste à consommer à un taux proportionnel à la fortune totale : $c(t) = C\rho(t) \forall t$, et conserver dans chaque actif risqué une proportion constante de la fortune totale : $\forall t, y_i^*(t) = y_i^* = \text{constante}$.

La politique de transaction optimale est donc la suivante : si $y_i < y_i^*$: acheter actif i , si $y_i > y_i^*$: vendre actif i .

La proportion optimale $y_i^* \in [0, 1]$ si et seulement si $r < \alpha_i < r + \sigma_i^2(1 - \gamma)$. Si $\alpha_i > r + \sigma_i^2(1 - \gamma)$, l'actif i est tellement avantageux que l'on emprunte de l'argent pour en acheter. Par contre si $\alpha_i < r$, c'est le contraire, on vend l'actif i à découvert.

La Figure (5.1) illustre le cas $n = 1$, c'est à dire le cas d'un portefeuille composé d'un compte en banque et d'un actif risqué, avec en abscisse le capital x_0 investi dans le compte en banque et en ordonnée le capital x_1 investi dans l'actif risqué. Le domaine admissible se situe au dessus de la droite $x_0 + x_1 = 0$. La stratégie optimale d'investissement est de maintenir le portefeuille sur "la ligne de Merton" d'équation $x_1 = \frac{y_1^*}{1-y_1^*}x_0$. Dès que le portefeuille s'écarte, l'investisseur doit opérer une transaction pour s'y ramener selon les directions indiquées sur le dessin.

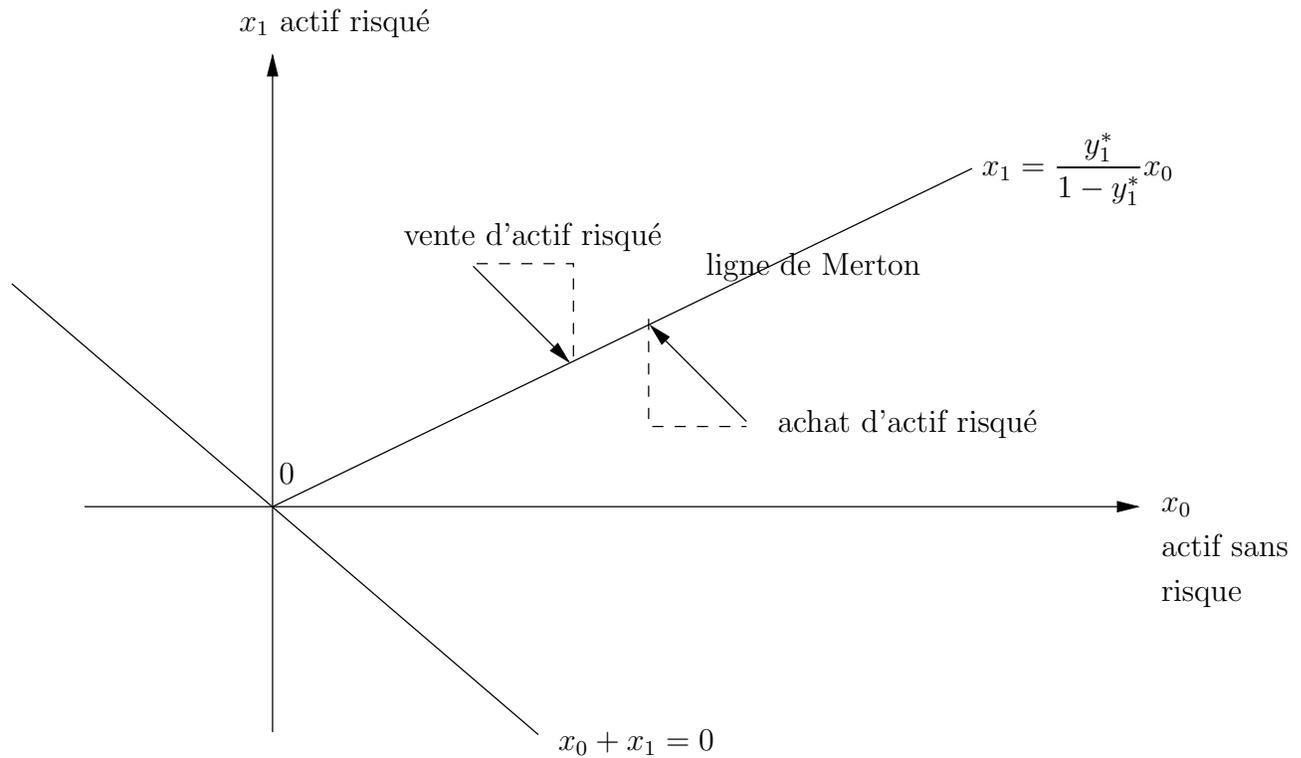


FIG. 5.1 – Espace des positions dans le cas d'un compte en banque et d'un actif risqué montrant la "ligne de Merton"

Remarque 5.7.2. Si $\gamma < 0$, la condition (5.56) est toujours satisfaite. Si $\gamma > 0$, la fonction valeur V est finie si et seulement si la condition (5.56) est satisfaite.

5.8 Contrôles singuliers et Inéquations variationnelles

Nous n'avons considéré jusque-là que des problèmes de contrôle standard. Il existe d'autres types de contrôles : en particulier les *contrôles singuliers* permettant de considérer le cas où le déplacement de l'état dû l'application du contrôle n'est pas différentiable en temps. Illustrons ce cas par un exemple de gestion de portefeuilles avec coûts de transaction proportionnels. Les équations de la programmation dynamique associées sont des inéquations variationnelles pour la fonction valeur.

5.8.1 Exemple de Gestion de portefeuille avec coûts de transaction proportionnels

On considère à présent le même modèle de gestion de portefeuille que précédemment, mais cette fois avec des coûts de transaction. On suppose que tout transfert d'argent entre les comptes entraîne un coût de transaction proportionnel au montant de la transaction : si l'on achète pour 1 Franc d'action i , on paye $(1 + \lambda_i)$ Francs ($\lambda_i \geq 0$) en déboursant l'argent du compte en banque et si l'on vend pour 1 Franc d'action i , on ne récupère que $(1 - \mu_i)$ Franc ($0 \leq \mu_i < 1$) dans le compte en banque.

La politique que l'on appliquait dans le cas de coûts de transaction nuls n'est évidemment plus réalisable car elle exige de faire sans cesse des transactions pour se maintenir dans les proportions optimales. Les coûts de transaction vont faire apparaître des régions dans lesquelles les transactions ne sont pas rentables et dans lesquelles la politique optimale est de ne faire aucune transaction.

En présence de coûts de transaction, on ne peut plus, comme dans le cas du problème de Merton, agréger les positions des différents comptes en une seule variable d'état, à savoir le capital total. Il est nécessaire de considérer l'évolution de chaque compte séparément. Notons $s_0(t), s_1(t), \dots, s_n(t)$, le montant de capital investi dans chaque action à l'instant t . On considère un modèle avec consommation sur un horizon infini. La dynamique du système est donnée par :

$$\begin{cases} ds_0(t) = (rs_0(t) - c(t))dt + \sum_{i=1}^n (-(1 + \lambda_i)d\mathcal{L}_i(t) + (1 - \mu_i)d\mathcal{M}_i(t)), \\ ds_i(t) = \alpha_i s_i(t)dt + \sigma_i s_i(t)dW_i(t) + d\mathcal{L}_i(t) - d\mathcal{M}_i(t), \quad i = 1, \dots, n, \end{cases} \quad (5.68)$$

avec les valeurs initiales $s_i(0^-) = x_i$, $i = 0, \dots, n$ où $\mathcal{L}_i(t)$ et $\mathcal{M}_i(t)$ représentent les quantités cumulées d'actif i respectivement achetées et vendues sur $[0, t]$, λ_i et μ_i sont coûts de transaction proportionnels, et $s_i(t^-)$ représente la limite à gauche du processus $s(t)$ à l'instant t .

Une *politique de consommation-investissement* est un ensemble $\mathcal{P} = (c(t), (\mathcal{L}_i(t), \mathcal{M}_i(t))_{i=1, \dots, n})$ de processus adaptés t.q.

$$- c(t) \geq 0, \quad \int_0^t c(s)ds < \infty, \text{ p.p.t,}$$

$$- \mathcal{L}_i(t) \text{ et } \mathcal{M}_i(t) \text{ continus à droite, croissants et } \mathcal{L}_i(0^-) = \mathcal{M}_i(0^-) = 0.$$

Le processus $s(t) = (s_0(t), \dots, s_n(t))$ est donc continu à droite avec limite à gauche (*cadlag*).

Les processus $\mathcal{L}_i(t)$ et $\mathcal{M}_i(t)$ ne sont pas absolument continus et peuvent avoir des dérivées infinies. Ce formalisme permet de modéliser différents modes de transactions : transactions instantanées de quantités finies ou transactions en continu à un certain taux.

La consommation est un contrôle absolument continu et les transactions $\mathcal{L}_i(t)$ et $\mathcal{M}_i(t)$ entre les actifs sont des contrôles de type *singulier*. Dans le cas du contrôle singulier, le déplacement de l'état dû à l'application du contrôle est discontinu par rapport au temps.

On définit le domaine admissible

$$\mathcal{D} = \{x = (x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1}, \quad \mathcal{W}(x) \geq 0\}$$

où

$$\mathcal{W}(x) = x_0 + \sum_{x_i > 0} (1 - \mu_i)x_i + \sum_{x_i \leq 0} (1 + \lambda_i)x_i \quad (5.69)$$

que l'on peut noter également

$$\mathcal{W}(x) = x_0 + \sum_{i=1}^n \min((1 - \mu_i)x_i, (1 + \lambda_i)x_i),$$

représente la fortune nette, c'est à dire la quantité d'argent obtenue dans le compte en banque après les transactions mettant à zero les actifs risqués (i.e. après avoir soldé les actifs risqués). Les quantités négatives d'argent dans les comptes sont donc autorisées à condition que le capital net $\mathcal{W}(x)$ soit positif.

Exemple : Lorsque $n = 1$, i.e. pour un portefeuille comprenant un actif non risqué et un actif risqué, le domaine admissible \mathcal{D} est défini par :

$$\mathcal{D} = \{(x_0, x_1) \in \mathbb{R}^2, x_0 + (1 - \mu)x_1 \geq 0, \quad x_0 + (1 + \lambda)x_1 \geq 0\}.$$

Il est représenté sur la figure (5.2). Le capital net $\mathcal{W}(x)$ est nul sur le bord du domaine.

Soit x la position initiale du portefeuille. Une politique est dite admissible si le temps de faillite $\bar{\tau}$ défini par

$$\bar{\tau} = \inf \{t \geq 0, s(t) \notin \mathcal{D}\} \quad (5.70)$$

est infini, c'est à dire $\mathcal{W}(s(t)) \geq 0$ presque sûrement sur toute la période de gestion. Soit $\mathcal{U}(x)$ l'ensemble des politiques admissibles, x étant la position initiale.

L'objectif est de maximiser sur l'ensemble des politiques admissibles une fonction d'utilité de la consommation de la forme

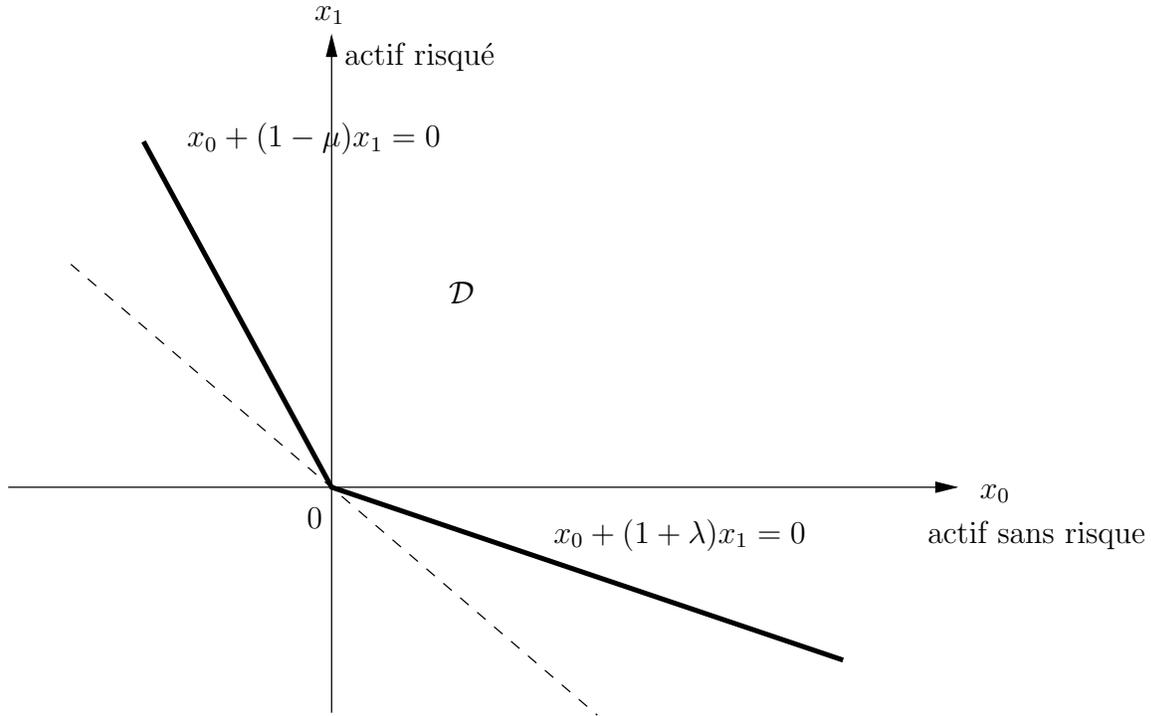
$$J_x(\mathcal{P}) = E_x \int_0^\infty e^{-\delta t} u(c(t)) dt, \quad (5.71)$$

où E_x représente l'espérance conditionnelle à l'état initial x , $\delta > 0$ est le taux d'actualisation et $u(c)$ est une fonction d'utilité de type HARA définie par

$$u(c) = \frac{c^\gamma}{\gamma}, \quad \gamma < 1, \quad \gamma \neq 0. \quad (5.72)$$

On définit la fonction valeur :

$$V(x) = \sup_{\mathcal{P} \in \mathcal{U}(x)} J_x(\mathcal{P}). \quad (5.73)$$

FIG. 5.2 – Espace admissible pour $n = 1$

Remarque 5.8.1. Soit τ le temps de sortie de l'intérieur de \mathcal{D} , défini par

$$\tau = \inf \left\{ t \geq 0, s(t) \notin \overset{\circ}{\mathcal{D}} \right\}. \quad (5.74)$$

Si $\gamma > 0$, on a, pour toute politique admissible \mathcal{P} ,

$$J_x(\mathcal{P}) = E_x \int_0^\tau e^{-\delta t} u(c(t)) dt. \quad (5.75)$$

D'autre part, pour toute politique \mathcal{P} , on peut construire une politique admissible qui coïncide avec \mathcal{P} jusqu'à τ (elle est telle que le processus $s(t)$ saute à l'origine à l'instant τ). La fonction valeur peut donc se réécrire

$$V(x) = \sup_{\mathcal{P} \in \mathcal{U}} E_x \int_0^\tau e^{-\delta t} u(c(t)) dt \quad (5.76)$$

où \mathcal{U} est l'ensemble de toutes les politiques. ■

Nous allons à présent établir l'équation de la Programmation Dynamique associée au problème de contrôle (5.68,5.71) décrit ci-dessus.

Comme les contrôles \mathcal{L}_i et \mathcal{M}_i représentant les transactions sont de type singulier, l'équation de la Programmation Dynamique est une inéquation variationnelle, c'est à dire un système d'inéquations aux dérivées partielles.

Nous nous restreignons au cas $\gamma > 0$. On obtient la même IV pour $\gamma < 0$ dans $\overset{\circ}{\mathcal{D}}$, mais V est infinie sur le bord $\partial\mathcal{D}$.

On a le théorème :

Théorème 5.8.1. *Sous les hypothèses*

$$\delta > \gamma \left(r + \frac{1}{2(1-\gamma)} \sum_{i=1}^n \left(\frac{\alpha_i - r}{\sigma_i} \right)^2 \right), \quad (5.77)$$

$$0 \leq \mu_i < 1, \quad \lambda_i \geq 0, \quad \lambda_i + \mu_i > 0, \quad \forall i = 1, \dots, n, \quad (5.78)$$

$$u(c) = \frac{c^\gamma}{\gamma}, \quad 0 < \gamma < 1, \quad (5.79)$$

la fonction valeur V est solution de viscosité de l'inéquation variationnelle (IV) elliptique :

$$\max \left\{ AV + u^* \left(\frac{\partial V}{\partial x_0} \right), \max_{1 \leq i \leq n} L_i V, \max_{1 \leq i \leq n} M_i V \right\} = 0 \quad \text{dans } \overset{\circ}{\mathcal{D}} \quad (5.80)$$

$$V = 0 \quad \text{sur } \partial \mathcal{D} \quad (5.81)$$

où

$$AV = \frac{1}{2} \sum_{i=1}^n \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \sum_{i=1}^n \alpha_i x_i \frac{\partial V}{\partial x_i} + r x_0 \frac{\partial V}{\partial x_0} - \delta V, \quad (5.82)$$

$$L_i V = -(1 + \lambda_i) \frac{\partial V}{\partial x_0} + \frac{\partial V}{\partial x_i}, \quad (5.83)$$

$$M_i V = (1 - \mu_i) \frac{\partial V}{\partial x_0} - \frac{\partial V}{\partial x_i}, \quad (5.84)$$

$$\begin{aligned} u^*(\nu) &= \max_{c \geq 0} (-c\nu + u(c)) \\ &= \left(\frac{1}{\gamma} - 1 \right) \nu^{\frac{\gamma}{\gamma-1}}. \end{aligned} \quad (5.85)$$

C'est l'unique solution de viscosité de l'inéquation variationnelle (5.80-5.81) dans la classe des fonctions continues dans \mathcal{D} qui satisfont

$$|V(x)| \leq C(1 + \|x\|^\gamma) \quad \forall x \in \mathcal{D}. \quad (5.86)$$

■

Remarque 5.8.2. *La condition au limite : Lorsque le processus $s(t)$ atteint la frontière $\partial \mathcal{D}$ à l'instant t , i.e. $s(t^-) \in \partial \mathcal{D}$, la seule politique admissible est de sauter à l'origine et d'y rester (il n'y a pas de diffusion en 0) avec une consommation nulle. Par conséquent, si l'avoir initial est sur la frontière, alors $V(x) = 0$ si $\gamma > 0$.*

Ce théorème est démontré dans [19]. Les grandes lignes de la preuve consistent à montrer que la fonction valeur V satisfait le principe de la Programmation Dynamique, c'est à dire :

$$V(x) = \sup_{\mathbb{P} \in \mathcal{U}} E \left(\int_0^{\theta \wedge \tau} e^{-\delta t} u(c(t)) dt + e^{-\delta(\theta \wedge \tau)} V(s((\theta \wedge \tau)^-)) \right) \quad \forall x \in \mathcal{D}, \quad (5.87)$$

pour tout temps d'arrêt θ .

Cela implique que V est une solution de viscosité de l'inéquation variationnelle (5.80-5.81)). Cela est un resultat standard dans le cas de processus de diffusion sans contrôle singulier. Pour les problèmes de contrôle singulier, on réfère à Fleming et Soner [32].

De plus, on peut montrer que la fonction valeur V satisfait (5.86). L'unicité est une conséquence du principe du maximum suivant :

Si v est une sous-solution de viscosité et v' est une sur-solution de viscosité de (5.80) qui satisfont (5.86) et $v \leq v'$ sur $\partial\mathcal{D}$, alors $v \leq v'$ dans \mathcal{D} .

Ceci implique bien l'unicité puisqu'une solution de viscosité de (5.80-5.81) est à la fois une sous-solution et une sur-solution avec la condition au bord $v = 0$ sur $\partial\mathcal{D}$.

Obtention formelle de l'Inéquation Variationnelle.

Pour comprendre l'origine de l'inéquation variationnelle, il est intéressant de voir comment elle peut s'obtenir formellement comme limite d'équations d'Hamilton-Jacobi-Bellman.

On peut montrer (et on l'admet ici) que les processus $\mathcal{L}_i(t)$ et $\mathcal{M}_i(t)$ peuvent s'écrire comme limites de contrôles absolument continus à dérivées bornées :

$$\mathcal{L}_i(t) = \lim_{K \rightarrow +\infty} \mathcal{L}_i^K(t)$$

$$\mathcal{M}_i(t) = \lim_{K \rightarrow +\infty} \mathcal{M}_i^K(t)$$

avec

$$\mathcal{L}_i^K(t) = \int_0^t l_i^K(s) ds, \quad \mathcal{M}_i^K(t) = \int_0^t m_i^K(s) ds, \quad 0 \leq l_i^K(s), m_i^K(s) \leq K$$

Dans le cas où l'on considère les processus $\mathcal{L}_i^K(t)$ et $\mathcal{M}_i^K(t)$ comme contrôles admissibles, les équations d'évolution des comptes peuvent s'écrire

$$\begin{cases} ds_0(t) = (rs_0(t) - c(t) + \sum_{i=1}^n (-(1 + \lambda_i)l_i^K(t) + (1 - \mu_i)m_i^K(t)))dt, \\ ds_i(t) = (\alpha_i s_i(t) + l_i^K(t) - m_i^K(t))dt + \sigma_i s_i(t) dW_i(t), \quad i = 1, \dots, n, \end{cases} \quad (5.88)$$

avec les valeurs initiales $s_i(0^-) = x_i$, $i = 0, \dots, n$ et la fonction valeur correspondante est définie par

$$V^K(x) = \sup_{\substack{c \geq 0 \\ 0 \leq l_i^K, m_i^K \leq K}} E_x \int_0^\infty e^{-\delta t} u(c(t)) dt. \quad (5.89)$$

Ce problème, de contrôle continu, est appelé *problème pénalisé*; lorsque $K \rightarrow +\infty$, il converge vers le problème initial et la fonction V^K tend vers V .

La fonction valeur V^K satisfait l'équation d'HJB :

$$\max_{\substack{c \geq 0 \\ 0 \leq l_i^K, m_i^K \leq K}} \left\{ AV^K - c \frac{\partial V^K}{\partial x_0} + u(c) + l_i^K L_i V^K + m_i^K M_i V^K \right\} = 0 \quad \text{dans } \overset{\circ}{\mathcal{D}} \quad (5.90)$$

$$V^K = 0 \quad \text{sur } \partial\mathcal{D} \quad (5.91)$$

où les opérateurs A , L_i et M_i sont définis en (5.82), (5.83), et (5.84).

Montrons que l'équation (5.90) tend vers l'IV (5.80) lorsque K tend vers l'infini.

L'équation (5.90) peut se récrire :

$$AV^K + u^* \left(\frac{\partial V^k}{\partial x_0} \right) + \max_{0 \leq l_i^K \leq K} l_i^K L_i V^K + \max_{0 \leq m_i^K \leq K} m_i^K M_i V^K = 0 \quad (5.92)$$

où u^* est définie par (5.85). Les contrôles optimaux l_i^{K*} et m_i^{K*} sont de type *bang-bang* :

$$l_i^{K*} = \begin{cases} 0 & \text{si } L_i V^K < 0 \\ K & \text{si } L_i V^K \geq 0 \end{cases} \quad (5.93)$$

$$m_i^{K*} = \begin{cases} 0 & \text{si } M_i V^K < 0 \\ K & \text{si } M_i V^K \geq 0. \end{cases} \quad (5.94)$$

L'équation (5.92) devient :

$$AV^K + u^* \left(\frac{\partial V^K}{\partial x_0} \right) + \sum_{i=1}^n K(L_i V^K)^+ + K(M_i V^K)^+ = 0 \quad (5.95)$$

où $(y)^+$ désigne la partie positive de y définie par $(y)^+ = \max(0, y)$. Comme $(L_i V^K)^+$ et $(M_i V^K)^+$ sont positifs, l'équation (5.92) implique

$$\forall K > 0, AV^K + u^* \left(\frac{\partial V^K}{\partial x_0} \right) \leq 0. \quad (5.96)$$

Admettons que V^K converge vers V lorsque $K \rightarrow +\infty$ et que AV^K converge. On déduit de (5.96) que

$$AV + u^* \left(\frac{\partial V}{\partial x_0} \right) \leq 0. \quad (5.97)$$

On a d'autre part

$$\sum_{i=1}^n (L_i V^K)^+ + (M_i V^K)^+ = -\frac{1}{K} (AV^K + u^* \left(\frac{\partial V^K}{\partial x_0} \right)) \rightarrow_{K \rightarrow +\infty} 0 \quad (5.98)$$

Par conséquent

$$(L_i V)^+ = 0 \text{ et } (M_i V)^+ = 0$$

ce qui équivalent à

$$L_i V \leq 0 \text{ et } M_i V \leq 0 \quad \forall i = 1 \dots n. \quad (5.99)$$

Le système d'inégalités (5.97, 5.99) est donc vérifié. Montrons qu'une au moins des inégalités est en fait une égalité :

- si $L_i V = 0$ ou $M_i V = 0$ on a bien une égalité
- si $L_i V < 0$ et $M_i V < 0$ pour tout $i = 1, \dots, n$, alors pour K assez grand, disons $K > K_0$, on a $L_i V^K < 0$ et $M_i V^K < 0$. Donc

$$\sum_{i=1}^n K(L_i V^K)^+ + K(M_i V^K)^+ = 0 \quad (5.100)$$

et (5.95) implique

$$AV^K + u^* \left(\frac{\partial V^K}{\partial x_0} \right) = 0, \quad \forall K > K_0. \quad (5.101)$$

D'où

$$AV + u^* \left(\frac{\partial V}{\partial x_0} \right) = 0. \quad (5.102)$$

La fonction V est donc solution de l'IV (5.80).

Analyse de la politique optimale

L'inéquation variationnelle nous indique que le domaine admissible \mathcal{D} est divisé ainsi :

$$B_i = \{x \in \mathcal{D}, L_i V(x) = 0\}, \quad (5.103)$$

$$S_i = \{x \in \mathcal{D}, M_i V(x) = 0\}, \quad (5.104)$$

$$NT_i = \mathcal{S} \setminus (B_i \cup S_i), \quad (5.105)$$

$$NT = \bigcap_{i=1}^n NT_i. \quad (5.106)$$

NT est la région de non-transaction. A l'extérieur de NT , une transaction instantanée ramène la position de l'investisseur sur la frontière de NT : achat d'actif i dans B_i , vente d'actif i dans S_i . Après la transaction initiale, la position de l'investisseur reste dans

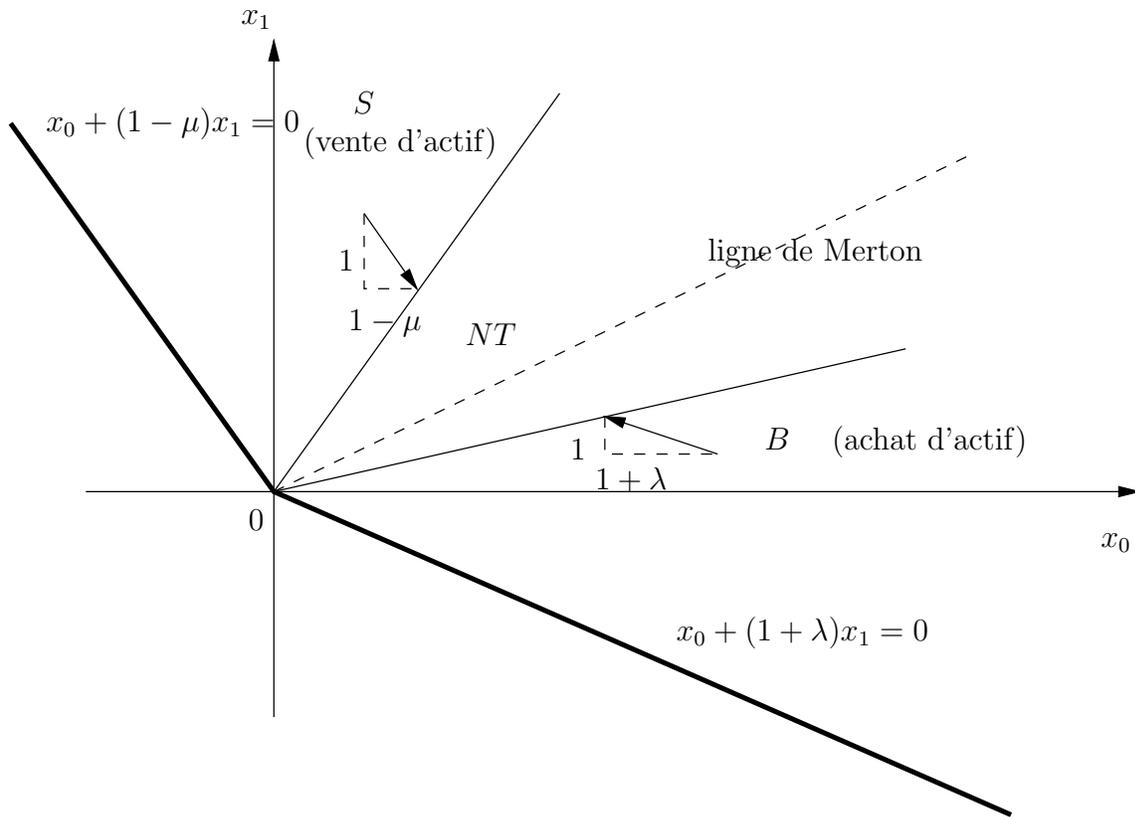
$$\overline{NT} = \{x \in \mathcal{D}, AV + u^* \left(\frac{\partial V}{\partial x_0} \right) = 0\},$$

et les transactions suivantes n'ont lieu que sur la frontière. (Le processus est réfléchi sur la frontière selon la direction appropriée donnée par $L_i V = 0$ ou $M_i V = 0$).

Nous faisons face ici à un problème de frontière libre, car la résolution de l'IV exige de déterminer non seulement la fonction V mais également les frontières des différentes régions. On ne peut pas résoudre, comme dans le cas de Merton, cette équation de manière explicite. Il faudra donc se tourner vers des méthodes numériques (voir chapitres suivants). Neanmoins, les propriétés de la fonction valeur permettent de dégager quelques heuristiques sur les régions optimales de transaction.

La figure (5.3) représente la partition du domaine dans le cas d'un compte non risqué et d'un actif risqué : la région de non-transaction est un cône NT , les régions d'achat (B) et de vente d'actif (S) risqué sont situées respectivement au-dessous et au-dessus de NT et les directions de transaction y sont indiquées. Ceci est prouvé plus loin.

5.8.2 Exemple des flux de dividendes

FIG. 5.3 – Politique optimale de transaction pour $n = 1$

Chapitre 6

Etude numérique des équations d'Hamilton-Jacobi-Bellman

Ce chapitre est consacré à l'étude numérique des équations d'Hamilton-Jacobi-Bellman elliptiques et paraboliques.

6.1 Equations d'Hamilton-Jacobi-Bellman elliptiques

On considère des équations de la forme :

$$\min_{P \in \mathcal{P}} (A^P v(x) + u(x, P)) = 0 \text{ dans } \Omega \subset \mathbb{R}^d \quad (6.1)$$

avec la condition au limite de Dirichlet

$$v = f \text{ sur } \delta\Omega \quad (6.2)$$

ou de Neumann

$$\frac{\partial v}{\partial n} = f \text{ sur } \delta\Omega. \quad (6.3)$$

A^P est un opérateur elliptique du deuxième ordre :

$$A^P v(x) = \sum_{i,j=1}^d a_{ij}(x, P) \frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(x, P) \frac{\partial v}{\partial x_i}(x) - \lambda v(x)$$

avec

$$\sum_{i,j=1}^d a_{ij}(x, P) \eta_i \eta_j \geq 0, \quad \forall x \in \Omega, \eta \in \mathbb{R}^d, P \in \mathcal{P}, \quad \lambda \geq 0.$$

On rappelle que les équations de type (6.1) sont les équations de la Programmation Dynamique associées aux problèmes de contrôle de diffusion avec critère actualisé sur un horizon infini. En effet la solution de (6.1) avec par exemple la condition au limite (6.2) peut, sous certaines conditions, s'interpréter comme :

$$v(x) = \min_{P(\cdot) \in \mathcal{P}} E \left(\int_0^\tau e^{-\lambda t} u(X_t, P_t) dt + e^{-\lambda \tau} f(X_\tau) \mid X_0 = x \right)$$

où X_t est un processus de diffusion contrôlé défini par

$$dX_t = b(X_t, P_t)dt + \sigma(X_t, P_t)dW_t, \quad X_0 = x,$$

W_t est un processus de Wiener, $a = \frac{1}{2}\sigma\sigma^T$, et

$$\tau(\omega) = \inf\{t, X_t(\omega) \notin \Omega\}$$

est le premier instant où X_t sort du domaine Ω .

Remarque :

Le traitement numérique des inéquations variationnelles (IV) est identique à celui des équations d'HJB. En effet les IV peuvent être formulées comme des équations d'HJB du type (6.1) en utilisant un contrôle discret. Considérons par exemple, l'inéquation variationnelle suivante apparaissant dans le problème de contrôle singulier modélisant les problèmes de gestion de portefeuilles avec coûts de transaction (voir chapitre 5.8.1) :

$$\max \left\{ AV + u^* \left(\frac{\partial V}{\partial x_0} \right), \max_{1 \leq i \leq n} L_i V, \max_{1 \leq i \leq n} M_i V \right\} = 0 \quad \text{dans } \overset{\circ}{\mathcal{D}}. \quad (6.4)$$

L'opérateur A est elliptique du second ordre et les opérateurs L_i et M_i sont du premier ordre (voir 5.82, 5.83, 5.84).

L'équation (6.4) peut se récrire

$$\max(A^0 W, \max_{P \in \mathcal{P}} A^P W) = 0 \quad (6.5)$$

où A^0 est un opérateur du second ordre, A^P sont des opérateurs du premier ordre et \mathcal{P} est un ensemble fini de contrôles de cardinal $2n$.

L'étude numérique de l'équation (6.1) consiste en 3 étapes : la première consiste à *se ramener à un domaine borné* (dans le cas où Ω est non borné) par les techniques décrites au chapitre 4.2.1, la deuxième à *discrétiser* l'équation et la troisième à *résoudre l'équation discrétisée*.

Supposons que l'on se soit ramené au cas où Ω est le cube unité de $\mathbb{R}^d : \Omega =]0, 1[^d$.

Soit $h = \frac{1}{n}$, ($n \in \mathbb{N}^*$) le pas de discrétisation dans chaque direction de coordonnée. On définit un maillage $\Omega_h = \Omega \cap (h\mathbb{Z})^d$ ou grille uniforme de pas h .

La méthode des différences finies décrites en 4.2.2 permet d'obtenir une approximation de la solution v de l'équation (6.1) aux nœuds du maillage. On note v_h cette approximation.

On obtient ainsi un système de N_h équations non linéaires à N_h inconnues $\{v_h(x), x \in \Omega_h\}$:

$$\min_{P \in \mathcal{P}} (A_h^P v_h + u_h(P))(x) = 0 \quad \forall x \in \Omega_h \quad (6.6)$$

où N_h est le nombre de points où il faut déterminer la valeur de la fonction.

Dans le cas de conditions aux limites de Dirichlet, la valeur de la fonction est connue aux bord du domaine et $N_h = (n-1)^d$ (nombre de points intérieurs à la grille). Dans le cas de conditions aux limites de Neumann, il faut également déterminer v_h sur les points du bord et $N_h = (n+1)^d$.

Soit \mathcal{P}_h l'ensemble des feedbacks $P : \Omega_h \rightarrow \mathcal{P}$ et \mathcal{V}_h l'ensemble des fonctions de Ω_h dans \mathbb{R} . L'équation (6.6) peut se réécrire :

$$\min_{P \in \mathcal{P}_h} (A_h^P v_h + u_h(P)) = 0, \quad v_h \in \mathcal{V}_h. \quad (6.7)$$

L'opérateur A_h^P , dépendant de P dans \mathcal{P}_h , envoie \mathcal{V}_h dans lui-même et peut donc être considéré comme une matrice $N_h \times N_h$.

Notons $N_h = m^d$ avec $m = n - 1$ ou $n + 1$. La matrice (A_h^P) est une matrice $m^d \times m^d$ tri-diagonale par blocs, composée de $m \times m$ sous-matrices $[A_h^P]_{ij}$ de dimension $m^{d-1} \times m^{d-1}$.

6.1.1 Méthode d'approximation

Conditions de stabilité

Considérons le problème discrétisé (6.7) tel que A_h^P satisfait le *Principe du Maximum discret strictement positif*. Pour cela, on utilise les discrétisations décrites au chapitre 4.2.3 qui assurent les conditions de stabilité

$$\begin{cases} (A_h^P)_{ij} \geq 0 & \forall i \neq j, \\ \sum_j (A_h^P)_{ij} \leq -\lambda < 0 & \forall i \text{ (} = -\lambda \text{ si conditions de Neumann)}. \end{cases} \quad (6.8)$$

Rappelons les brièvement :

- **Cas d'un opérateur elliptique sans dérivées croisées et uniformément elliptique.**

Supposons $a_{ij}(x, P) = 0$ pour $i \neq j, \forall x \in \Omega, P \in \mathcal{P}$ et que $a_{ii}(x, P) \geq \alpha > 0$. Alors l'opérateur A_h^P obtenu en discrétisant les dérivées secondes par (4.8) et les dérivées premières par des schémas saute-mouton (4.4) satisfait les conditions de stabilité (6.8) pour un pas de discrétisation h vérifiant

$$h \leq \min_{x \in \Omega_h, P \in \mathcal{P}_h} \frac{2a_{ii}(x, P)}{|b_i(x, P)|}. \quad (6.9)$$

- **Cas d'un opérateur elliptique sans dérivées croisées et dégénéré.**

Lorsque l'opérateur A est dégénéré, on obtient une discrétisation vérifiant les conditions (6.8) et donc le Principe du Maximum Discret strictement positif, mais d'un ordre de convergence inférieur, en utilisant des approximations décentrées du gradient : on utilise le schéma décentré positif (4.5) si $b_i(x, P) \geq 0$ et le schéma décentré négatif (4.6) si $b_i(x, P) < 0$.

- **Cas d'un opérateur elliptique avec dérivées croisées.** Si la condition

$$a_{ii}(x, P) \geq \sum_{j, j \neq i} |a_{ij}(x, P)| \quad (6.10)$$

est satisfaite $\forall i = 1 \dots d, x \in \Omega, P \in \mathcal{P}$, alors l'opérateur A_h^P obtenu en utilisant les discrétisations (4.9) et (4.10) pour les dérivées croisées, satisfait le Principe du Maximum Discret strictement positif.

Interprétation probabiliste du problème discrétisé

Si les conditions de stabilité (6.8) sont satisfaites, il existe $k > 0$ et une matrice markovienne ou sous-markovienne M_h^P tels que

$$A_h^P = -\lambda I_h + \frac{1}{k}(M_h^P - I_h) \quad (6.11)$$

où I_h désigne l'identité de \mathcal{V}_h .

On a

$$M_h^P = I_h + k(A_h^P + \lambda I_h)$$

et donc

$$(M_h^P)_{ij} = \begin{cases} 1 + k(\lambda + (A_h^P)_{ii}) & \text{si } i = j \\ k(A_h^P)_{ij} \geq 0 & \text{si } i \neq j \end{cases}$$

On choisit $k > 0$ tel que

$$k \leq \frac{1}{\lambda + |(A_h^P)_{ii}|} \quad \forall i = 1, \dots, d$$

ce qui assure que tous les coefficients de la matrice M_h^P sont positifs.

On a

$$\sum_{j=1}^d (M_h^P)_{ij} = 1 + \lambda k + k \sum_{j=1}^d (A_h^P)_{ij} \begin{cases} = 1 & \text{si Neumann} \\ < 1 & \text{si Dirichlet } (M_h^P \text{ est sous-Markovienne}) \end{cases}$$

L'équation discrète (6.6) peut se récrire :

$$\min_{P \in \mathcal{P}_h} (M_h^P - I_h - \lambda k)v_h + k u_h(P) = 0, \quad v_h \in \mathcal{V}_h. \quad (6.12)$$

Cette équation est l'équation de la programmation dynamique d'un problème de contrôle de chaîne de Markov avec taux d'actualisation λk , coût instantané $k u_h(P)$ et matrice de transition M_h^P :

$$v_h(x) = \min_{(P_n)} E \left\{ \sum_{n=0}^{\tau} \frac{k}{(1 + \lambda k)^{n+1}} u(X_n, P_n) \mid X_0 = x \right\}$$

$$P(X_{n+1} = y \mid X_n = x) = (M_h^{P_n})_{xy}$$

pour tout x de Ω_h . Dans le cas de conditions de Neumann, $\tau = \infty$ et dans le cas de conditions de Dirichlet, τ est le premier temps de sortie de la grille.

Résultats de convergence

. Si v_h désigne la solution du problème discret (6.7), et $(v)_h$ la restriction à Ω_h de la solution du problème continu (6.1), on a

$$\|v_h - (v)_h\|_{\infty} = \begin{cases} \mathcal{O}(h) & \text{si } v \in \mathcal{C}^3(\bar{\Omega}) \\ \mathcal{O}(h^2) & \text{si } v \in \mathcal{C}^4(\bar{\Omega}) \text{ et si on utilise les différences finies symétriques} \end{cases}$$

pour des données a_{ii}, b_i, c assez régulières.

Pour des fonctions v moins régulières, on a les estimations suivantes ; en notant p_h l'interpolation linéaire sur chaque coordonnée, et en supposant les contrôles assez réguliers

$$\|p_h v_h - v\|_{H^1(\Omega)} \leq Ch \|v\|_{H^2(\Omega)} \quad (6.13)$$

et si on utilise la discrétisation symétrique du gradient

$$\|p_h v_h - v\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{H^2(\Omega)}. \quad (6.14)$$

Selon que les estimations (6.13) ou (6.14) sont satisfaites, on dit que l'approximation est d'ordre h ou h^2 .

6.1.2 Algorithmes de résolution de l'équation discrétisée

On s'intéresse à présent à la résolution numérique du problème discrétisé

$$\min_{P \in \mathcal{P}_h} (A_h^P v_h + u_h(P)) = 0, \quad v_h \in \mathcal{V}_h. \quad (6.15)$$

Lorsque le problème discret s'interprète comme un problème de contrôle de chaîne de Markov, on peut le résoudre en utilisant les algorithmes présentés dans le cas discret.

Les itérations sur les valeurs

Cette méthode [26] consiste à construire une suite (v_h^n) convergeant vers v_h en faisant l'itération contractante

$$v_h^{n+1} = \frac{1}{1 + \lambda k} \min_{P \in \mathcal{P}_h} (M_h^P v_h^n + k u_h(P)) \quad (6.16)$$

où

$$M_h^P = I_h + k(A_h^P + \lambda I_h).$$

Cette méthode converge car $\|M_h^P\|_\infty \leq 1$. Le coefficient de contraction est égal à $1/(1 + \lambda k) = 1 - \mathcal{O}(h^2)$ car $k = \mathcal{O}(h^2)$.

Cet algorithme n'est plus efficace lorsque λ est proche de 0.

Remarquons que la méthode d'itération sur les valeurs est équivalente aux itérations de Jacobi dans le cas linéaire. La lenteur de l'itération sur les valeurs est due au mauvais conditionnement de l'opérateur A_h : ($\text{cond } A_h = \mathcal{O}(1/h^2)$).

Cas où le Principe du Maximum Discret n'est pas vérifié.

Lorsque l'opérateur A_h^P ne satisfait pas le Principe du Maximum Discret, l'équation (6.6) ne peut pas être interprétée comme une équation de Bellman discrète. Pourtant, la méthode d'itérations sur les valeurs (6.16) peut quand même être utilisée si l'on trouve λ et k tels que la norme L^2 de M^P (qui n'est plus une matrice de Markov) est plus petite que 1 pour tout P . Cette condition peut être obtenue par exemple lorsque le coefficient d'actualisation λ est assez grand.

Les itérations sur les politiques ou Algorithme d'Howard.

C'est un algorithme itératif sur le contrôle et la fonction valeur (commençant de P^0) :

$$\text{pour } n \geq 0 \quad v_h^n \text{ est solution de } A_h^{P_h^n} v + u_h(P_h^n) = 0 \quad (6.17)$$

$$\text{pour } n \geq 1 \quad P_h^n \in \underset{P \in \mathcal{P}_h}{\text{Argmin}}(A_h^P v_h^{n-1} + u_h(P)) \quad (6.18)$$

$$(6.19)$$

Lorsque A_h^P satisfait le Principe du Maximum discret, la suite v_h^n décroît et converge vers la solution v_h de (6.6). La convergence est en général superlinéaire (voir [25, 26, 17, 18]).

L'algorithme de Howard nécessite peu d'itérations mais chaque itération nécessite la résolution d'un système linéaire.

Algorithmes d'optimisation.

Dans les algorithmes d'itérations sur les valeurs ou sur les politiques, il faut à chaque étape faire une minimisation du type

$$\min_{P \in \mathbb{P}^h} (A_h^P v + u_h(P)).$$

Pour cela, on utilise en général des méthodes itératives d'optimisation. On décrit ci-dessous quelques algorithmes usuels. Souvent, les conditions de convergence de ces algorithmes ne sont pas vérifiées et il faut utiliser des variantes.

Rappels :

- Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est *convexe* si et seulement si

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \forall \alpha \in [0, 1], \forall x, y \in \mathbb{R}^n.$$

f est *strictement convexe* si l'inégalité est stricte.

- Si f est de classe C^1 , f est convexe Ssi :

$$f(y) \geq f(x) + Df(x) \cdot (y - x) \quad \forall x, y \in \mathbb{R}^n$$

- Si f est de classe C^2 , f est convexe Ssi :

$$D^2 f(x)(y - x) \cdot (y - x) \geq 0 \quad \forall x, y \in \mathbb{R}^n$$

où $Df(x)$ désigne le gradient de f $D^2 f(x)$ la matrice hessienne de f au point x .

Conditions d'optimalité.

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une application de classe \mathcal{D}^1 . Toute solution x^* du problème

$$\min_{x \in \mathbb{R}^n} f(x) \quad (6.20)$$

vérifie la *condition nécessaire d'optimalité du premier ordre* (appelée équation d'Euler) :

$$Df(x^*) = 0 \quad (6.21)$$

Si f est convexe, (6.21) est une condition nécessaire et suffisante.

La prise en compte des dérivées secondes permet d'une part de donner une deuxième condition nécessaire d'extremum permettant de distinguer entre minimum et maximum, et d'autre part de donner des conditions suffisantes d'extremum. Supposons que x^* est une solution du problème (6.20) et que f est de classe C^2 . Alors

1. x^* vérifie la condition nécessaire d'optimalité du premier ordre (6.21) et la condition nécessaire d'optimalité du deuxième ordre

$$D^2 f(x^*) \geq 0. \quad (6.22)$$

2. Tout point x^* tel que

$$Df(x^*) = 0, \quad D^2 f(x^*) > 0 \quad (\text{i.e. } D^2 f(x^*)x.x > 0, \forall x \in \mathbb{R}^n) \quad (6.23)$$

est un *minimum local strict* de f , c'est à dire qu'il existe un voisinage $\mathcal{V}(x^*)$ de x^* tel que $f(x) > f(x^*), \forall x \in \mathcal{V}(x^*), x \neq x^*$.

Considérons à présent le problème d'optimisation avec contrainte :

$$\min_{x \in K} f(x) \quad (6.24)$$

où K est un convexe fermé de \mathbb{R}^n .

Toute solution x^* du problème (6.20) vérifie la *condition nécessaire d'optimalité du premier ordre* (inéquation d'Euler) :

$$Df(x^*). (x - x^*) \geq 0, \quad \forall x \in K. \quad (6.25)$$

Si la fonction f est elle-même *convexe*, (6.25) est une condition nécessaire et suffisante.

Algorithme de gradient.

On considère le problème d'optimisation sans contrainte :

$$\min_{x \in \mathbb{R}^n} f(x) \quad (6.26)$$

L'algorithme de gradient consiste à faire l'itération :

$$x_{k+1} = x_k - \rho_k Df(x_k) \quad (6.27)$$

La proposition suivante assure la convergence de cet algorithme avec un taux de convergence linéaire :

Proposition 6.1.1. *Si $Df(x)$ est lipchitzienne en x de constante de Lipschitz égale à M , et vérifie*

$$(Df(x_1) - Df(x_2)).(x_1 - x_2) \geq \alpha |x_1 - x_2|^2 \quad \text{avec } \alpha > 0 \quad (6.28)$$

alors l'algorithme du gradient converge avec un pas constant $\rho_k \equiv \rho$ tel que $0 < \rho < \frac{2\alpha}{M}$, vers x^ vérifiant $Df(x^*) = 0$.*

Preuve : L'algorithme du gradient s'écrit : $x_{k+1} = h(x_k)$ avec $h(x) = x - \rho Df(x)$. L'application $h(x)$ est strictement contractante pour ρ assez petit. En effet

$$|h(x_1) - h(x_2)|^2 = |x_1 - x_2|^2 - 2\rho(x_1 - x_2).(Df(x_1) - Df(x_2)) \quad (6.29)$$

$$+ \rho^2 |Df(x_1) - Df(x_2)|^2 \quad (6.30)$$

ce qui, avec les hypothèses faites sur $Df(x)$ donne :

$$|h(x_1) - h(x_2)|^2 \leq (1 - 2\rho\alpha + \rho^2 M^2) |x_1 - x_2|^2. \quad (6.31)$$

Si $0 < \rho < \frac{2\alpha}{M^2}$, alors $1 - 2\rho\alpha + \rho^2 M^2 < 1$ et $h(x)$ est strictement contractante.

On en déduit, par le théorème du point fixe, la convergence de x_k vers x^* vérifiant :

$$x^* = h(x^*) = x^* - \rho Df(x^*)$$

donc $Df(x^*) = 0$. □

- La méthode du gradient avec projection : Dans le cas avec contrainte, on considère l'algorithme

$$x_{k+1} = P_K(x_k - \rho Df(x^k)) \quad (6.32)$$

où $P_K(x)$ désigne la projection de $x \in \mathbb{R}^n$ sur le convexe fermé K définie comme l'unique élément de K vérifiant :

$$|x - P_K(x)| \leq |x - y| \quad \forall y \in K.$$

L'algorithme (6.32) converge sous les mêmes hypothèses que dans le cas sans contrainte (voir proposition 6.1.1).

Algorithme de Newton

On s'intéresse à la résolution de l'équation

$$Df(x^*) = 0.$$

Considérons le problème sous une forme plus générale : Soit F une application définie sur \mathbb{R}^n . On cherche x^* tel que $F(x^*) = 0$.

La méthode de Newton consiste à faire l'itération

$$x_{k+1} = x_k - (DF(x_k))^{-1} F(x_k), \quad k \geq 0$$

x_0 arbitraire.

Cela suppose que F est dérivable et que $DF(x)$ est inversible en tout point.

Étudions les propriétés locales de l'algorithme : soit x^* un zéro régulier de F au sens où $DF(x^*)$ est inversible. Supposons de plus que DF est localement lipschitzien de constante L au voisinage de x^* :

$$|DF(x) - DF(y)| \leq L|y - x|, \quad \forall x, y \in \mathcal{V}(x^*)$$

où $\mathcal{V}(x^*)$ désigne un voisinage de x^* . Alors si x_0 est près de x^* , la suite x_k converge vers x^* de façon quadratique, i.e. $\exists c > 0$ tel que

$$|x_{k+1} - x^*| \leq c|x_k - x^*|^2.$$

On applique cet algorithme à la résolution de $Df(x^*) = 0$, c'est à dire que l'on prend $F = Df$. La condition $D^2 f(x^*)$ inversible est équivalente à $D^2 f(x^*) > 0$. On a le théorème

Théorème 6.1.1. *Soit f de classe \mathcal{D}^2 , avec $D^2 f$ localement lipschitzien. Soit x^* un minimum local régulier de f : $Df(x^*) = 0$ et $D^2 f(x^*) > 0$. Soit x^0 assez proche de x^* . Alors la méthode de Newton définie par*

$$Df(x_k) + D^2 f(x^k)(x_{k+1} - x_k) = 0$$

est bien définie et converge de façon quadratique vers x^ .*

6.2 Equations d'Hamilton-Jacobi-Bellman paraboliques

On veut résoudre

$$\begin{cases} \frac{\partial v(t, x)}{\partial t} + \min_{P \in \mathcal{P}} (A^P v(t, x) - rv(t, x) + u(t, x, P)) = 0 & \text{dans } [0, T[\times \Omega \\ V(T, x) = \phi(x) & \forall x \in \Omega \\ v = f \text{ ou bien } \frac{\partial v}{\partial n} = f \text{ sur } \partial\Omega \times [0, T[\end{cases} \quad (6.33)$$

où A^P est un opérateur du deuxième ordre elliptique

$$A^P v(t, x) = \sum_{i,j=1}^d a_{ij}(t, x, P) \frac{\partial^2 v}{\partial x_i \partial x_j}(t, x) + \sum_{i=1}^d b_i(t, x, P) \frac{\partial v}{\partial x_i}(t, x)$$

avec

$$\sum_{i,j=1}^d a_{ij}(t, x, P) \eta_i \eta_j \geq 0, \quad \forall x \in \Omega, t \in [0, T], \eta \in \mathbb{R}^d, P \in \mathcal{P}.$$

Les équations de type (6.33) sont les équations de la Programmation Dynamique associées aux problèmes de contrôle de diffusion sur un horizon fini T et la solution v s'interprète dans le cas de condition de Dirichlet comme

$$v(t, x) = \min_{P(\cdot) \in \mathcal{P}} E \left(\int_t^{\tau \wedge T} u(s, X_s, P_s) ds + f(\tau, X_\tau) 1_{\tau < T} + g(X_T) 1_{T < \tau} \mid X_t = x \right)$$

6.2.1 Méthodes d'approximation

On se place dans le cas $\Omega =]0, 1[^d$. On discrétise l'équation en temps par les θ -schémas décrits au paragraphe 4.5.1. Soit $\Delta t = \frac{T}{N}$ le pas de discrétisation en temps, $N \in \mathbb{N}^*$, et h un pas de discrétisation en espace. On établit un maillage Ω_h en espace et on note A_h^P l'opérateur discrétisé de $A^P - rI$ par la méthode de différences finies décrites en (4.2.2). Pour tout $t = k\Delta t$, $k = 0, \dots, N-1$, on note $v_h^t(x)$ l'approximation de $v(t, x)$, $x \in \Omega_h$. On peut considérer v_h^t comme un vecteur de dimension N_h . Pour tout $t = k\Delta t$, $k = 0, \dots, N-1$, on doit résoudre :

$$\min_{P \in \mathcal{P}_h} \{ (I - \theta \Delta t A_h^P) v_h^t - (I + (\Delta t(1 - \theta) A_h^P) v_h^{t+\Delta t} - \Delta t u_h^{P,t} \} = 0$$

– **Schéma explicite** : $\theta = 0$

On fait l'approximation :

$$\left(\frac{\partial v}{\partial t} + A^P v \right)(t, x) \sim \frac{v(t + \Delta t, x) - v(t, x)}{\Delta t} + A_h^P v(t + \Delta t, x) \quad (6.34)$$

Dans ce cas, v_h^t se calcule explicitement en fonction de $v_h^{t+\Delta t}$:

$$\begin{cases} v_h^t & = \min_{P \in \mathcal{P}_h} \{ (M^P v_h^{t+\Delta t} + \Delta t u_h^{P,t} \} \\ v_h^T & = \phi_h. \end{cases} \quad (6.35)$$

avec $M^P = I + \Delta t A_h^P$.

Le schéma est stable si

– la matrice A_h^P satisfait :

$$\begin{cases} (A_h^P)_{ij} \geq 0 & \forall i \neq j, \\ \sum_j (A_h^P)_{ij} \leq -r & \forall i \quad (= -r \text{ dans le cas Neumann}) \end{cases} \quad (6.36)$$

– et si la condition de stabilité suivante est satisfaite :

$$\frac{\Delta t}{h^2} \leq C \quad (6.37)$$

où

$$C = \frac{1}{2} \min_{x,t,P} \frac{1}{\sum_{i=1}^d a_{ii}(t,x,P)}.$$

Si les conditions (6.36) et (6.37) sont satisfaites, on a une interprétation probabiliste du problème discrétisé.

En effet,

- si $r = 0$ (pas d'actualisation), la matrice $M^P \equiv I + \Delta t A_h^P$ est markovienne (ou sous-markovienne). L'équation (6.35) est alors l'équation de la Programmation Dynamique du problème de contrôle de chaîne de Markov en horizon fini suivant :

$$v_h^t(x) = \min_{P \in \mathcal{P}_h} E \left\{ \sum_{i=0}^{\frac{T-t}{\Delta t} - 1} \Delta t u_h^P{}^{t+i\Delta t}(X(t+i\Delta t)) + \Phi_h(X(T)) | X(t) = x \right\}$$

où $X(t)$ est une chaîne de Markov définie sur l'espace Ω_h des points de discrétisation et de matrice de transition M^P .

- Si $r \neq 0$, on a l'interprétation probabiliste :

$$v_h^t(x) = \min_{P \in \mathcal{P}_h} E \left\{ \sum_{i=0}^{\frac{T-t}{\Delta t} - 1} \frac{\Delta t}{(1+r\Delta t)^{i+1}} u_h^P{}^{t+i\Delta t}(X(t+i\Delta t)) + \frac{1}{(1+r\Delta t)^{\frac{T-t}{\Delta t}}} \Phi_h(X(T)) | X(t) = x \right\}$$

où $X(t)$ est une chaîne de Markov de matrice de transition $M^P = I + \Delta t(rI + A_h^P)$.

– **Schéma implicite** : $\theta = 1$ On fait l'approximation :

$$\left(\frac{\partial v}{\partial t} + A^P v \right)(t, x) \sim \frac{v(t + \Delta t, x) - v(t, x)}{\Delta t} + A_h^P v(t, x). \quad (6.38)$$

L'équation discrétisée s'écrit :

$$\min_{P \in \mathcal{P}_h} \{ N_h^P v_h^t - v_h^{t+\Delta t} - \Delta t u_h^P{}^t \} = 0$$

avec $N_h^P = I - \Delta t A_h^P$. Le schéma est stable si $\|(N_h^P)^{-1}\|_\infty$ est borné uniformément en h , ce qui implique l'inversibilité de N_h^P .

– Si A_h^P est à diagonale dominante (i.e si A_h^P satisfait (6.36)), alors N_h^P est à diagonale fortement dominante et le schéma est stable.

– Si A_h^P ne vérifie pas le PMD, il suffit de prendre Δt assez petit pour que N_h^P soit inversible.

Ce schéma converge toujours lorsque que h et Δt tendent vers 0.

On résoud de façon rétrograde, en partant de $v_h^T = \Phi_h$, N équations stationnaires

$$\min_P \{ (I - \Delta t A_h^P) v_h^t - \Delta t u_h^P{}^t \} = 0 \text{ pour } t = k\Delta t, k = N - 1, \dots, 0.$$

– **Schéma de Crank-Nicholson** : $\theta = \frac{1}{2}$

Ce schéma est inconditionnellement convergent. On résoud de façon rétrograde, en partant de $v_h^T = \Phi_h$, les N équations :

$$\min_P \left\{ \left(I - \frac{1}{2} \Delta t A_h^P \right) v_h^t - \left(I + \frac{1}{2} \Delta t A_h^P \right) v_h^{t+\Delta t} - \Delta t u_h^t P \right\} = 0, \quad \text{pour } t = k\Delta t, k = N-1, \dots, 0.$$

6.2.2 Exemples

Le put américain dans le modèle de Black-Scholes.

Dans le cas du put américain, on peut utiliser un algorithme efficace et assez utilisé en modifiant l'algorithme décrit au paragraphe ?? pour le calcul du prix européen. Pour le modèle de Black-Scholes, la valeur d'un put américain est donnée par

$$V_t = \sup_{t \leq \tau \leq T} E(e^{-\int_t^\tau r(s, X_s) ds} g(X_\tau) | X_t)$$

avec $g(x) = (K - x)_+$.

Après changement de variable logarithmique, $V_t = v(t, \log(X_t))$ où $v(t, x)$ satisfait l'inéquation variationnelle

$$\begin{cases} \max\left\{ \frac{\partial v}{\partial t}(t, x) + \tilde{A}v(t, x), v(t, x) - f(x) \right\} = 0 & \text{dans } [0, T[\times \mathbb{R}, \\ v(T, x) = f(x) & \forall x \in \mathbb{R} \end{cases} \quad (6.39)$$

où $f(x) = (K - e^x)_+$ et \tilde{A} est défini en (??).

On procède comme dans le cas européen. On localise le problème pour se ramener à une inéquation dans $\mathcal{O}_l =]-l, l[$. Ecrivons l'IV avec des conditions aux limites de Neumann

$$\begin{cases} \max\left\{ \frac{\partial v}{\partial t}(t, x) + \tilde{A}v(t, x), v(t, x) - f(x) \right\} = 0 & \text{dans } [0, T[\times \mathcal{O}_l, \\ \frac{\partial v}{\partial x}(t, l) = \frac{\partial v}{\partial x}(t, -l) = 0 & \text{dans } [0, T[\\ v(T, x) = f(x) & \forall x \in \mathcal{O}_l. \end{cases} \quad (6.40)$$

L'erreur que l'on commet lorsqu'on passe de l'équation (6.39) à (6.40) peut être estimée de la même façon que dans le cas européen (voir [11]). On discrétise en espace par les différences finies et en temps par les θ -schémas. Soit h le pas de discrétisation en espace : $h = \frac{2l}{n}$. On considère donc le maillage : $x_i = -l + i\frac{2l}{n}$, $i = 0, \dots, n$

On note N le nombre de points du maillage où il faut approximer v ($N = n + 1$).

L'opérateur A étant uniformément elliptique, on discrétise la dérivée première en espace par un schéma saute-mouton.

$\forall t$, on note v_h^t le vecteur tel que $(v_h^t)_i$ approche $v(t, x_i)$ et f_h le vecteur où chaque f_h^i est censé approcher $f(x_i)$. On obtient, après discrétisation en espace l'opérateur A_h représenté par la matrice tri-diagonale de dimension $N \times N$:

$$A_h = \begin{pmatrix} \beta & \alpha + \gamma & 0 & \cdots & 0 \\ \alpha & \beta & \gamma & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \alpha & \beta & \gamma \\ 0 & \cdots & 0 & \alpha + \gamma & \beta \end{pmatrix}$$

avec

$$\begin{aligned} \alpha &= \frac{\sigma^2}{2h^2} - \frac{1}{2h} \left(r - \frac{\sigma^2}{2} \right) \\ \beta &= -r - \frac{\sigma^2}{h^2}, \\ \gamma &= \frac{\sigma^2}{2h^2} + \frac{1}{2h} \left(r - \frac{\sigma^2}{2} \right) \end{aligned}$$

On discrétise à présent en temps : soit $\theta \in [0, 1]$, et Δt le pas en temps tel que $\Delta t = \frac{T}{M}$, $M \in \mathbb{N}^*$.

Si u et v sont deux vecteurs de \mathbb{R}^N , on note formellement $u \leq v$ pour $\forall i = 1, \dots, n, u_i \leq v_i$.

On doit résoudre pour tout $t = k\Delta t, k = N - 1, \dots, 0$ l'inéquation en dimension finie :

$$\max\{TX - G, X - F\} = 0 \quad (6.41)$$

avec

$$\begin{cases} T &= I - \theta \Delta t A_h \\ X &= v_h^t \\ G &= I + (\Delta t(1 - \theta)A_h)v_h^{t+\Delta t} \\ F &= f_h \end{cases}$$

La matrice T est tridiagonale :

$$T = \begin{pmatrix} b_1 & c_1 & 0 & \cdots & 0 \\ a_2 & b_2 & c_2 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & a_{N-1} & b_{N-1} & c_{N-1} \\ 0 & \cdots & 0 & a_N & b_N \end{pmatrix}$$

La matrice T est dite uniformément cervice s'il existe $\alpha > 0$ tel que $TX.X \geq \alpha|X|^2$. Cette condition est vérifiée ici si $|r - \frac{\sigma^2}{2}| \leq \frac{\sigma^2}{h}$ et si $\frac{k}{2h}|r - \frac{\sigma^2}{2}| < 1$. Dans ce cas, on peut résoudre le

système 6.41 en utilisant l'algorithme suivant, appelé "algorithme de Brennan et Schwartz" ([3]) dans le cas $\theta = 1$ (schéma implicite). On pose

$$\begin{cases} b'_N = b_N, \\ g'_N = g_N, \\ \text{Pour } i = N - 1, \dots, 1, \\ b'_i = b_i - c_i \frac{a_{i+1}}{b'_{i+1}}, \\ g'_i = g_i - c_i \frac{g'_{i+1}}{b'_{i+1}}, \end{cases}$$

Il suffit de calculer X en partant du haut de la matrice (descente "américaine") :

$$\begin{cases} x_1 = \frac{g'_1}{b'_1}, \\ \text{Pour } i = 2, \dots, N, x_i = \max(g'_i - a_i \frac{x_{i-1}}{b'_i}, f_i). \end{cases} \quad (6.42)$$

On trouvera une implémentation de l'algorithme de Brennan et Schwartz dans [?], chapitre 8.

6.3 Résolution numérique de problème de Contrôle ergodique

6.4 Résolution numérique de problème de contrôle de diffusion avec sauts

6.5 Méthodes de Kushner

Preuve probabiliste de convergence d'approximation de la fonction valeur vers la fonction valeur.

6.6 Preuve de convergence par les solutions de viscosité

méthodes de Barles, Souganidis.

6.7 Remarques conclusives

- Des algorithmes rapides basés sur l'algorithme d'Howard et les méthodes multigrilles pour la résolution du système linéaire intervenant dans l'étape (6.19) de l'algorithme d'Howard ont été développés par M. Akian (voir [17, 18]) pour résoudre les équations d'HJB discrétisées : ce sont les algorithmes multigrille-Howard et Full-Multigrille-Howard. L'avantage de cette méthode est que chaque itération multigrille prend un temps de calcul de $\mathcal{O}(N_h)$ et contracte l'erreur par un facteur indépendant du pas de discretisation h (contrairement aux algorithmes de Jacobi, Gauss-Seidel, gradient conjugué).
- La méthode de la Programmation Dynamique, qui consiste donc à la résolution de l'équation de HJB associée au problème de contrôle stochastique considéré fournit la fonction valeur et le feedback optimal. C'est donc la meilleure méthode lorsqu'elle est applicable. Ses limites d'application sont essentiellement la dimension de l'état du système. On peut utiliser des algorithmes rapides du type multigrille mais on ne pourra pas

traiter par cette méthode des systèmes de très grande taille : par exemple gérer des portefeuilles contenant des centaines d'actions. Dans ce cas, il faut se tourner vers des méthodes d'optimisation directes en boucle ouverte, à l'inverse de la Programmation Dynamique qui conduit à des contrôles en feedback (boucle fermée) sur l'état.

– **Le cas des inéquations variationnelles.**

Le traitement numérique des inéquations variationnelles est identique à celui des équations d'HJB. En effet on a vu que les IV peuvent être formulées comme des équations d'HJB.

Le seul problème est que l'on perd le caractère uniformément elliptique de l'opérateur, et donc les preuves des théorèmes de convergence ne sont plus valides.

6.8 Exercices

Exercice : *Interprétation de l'algorithme d'Howard en terme d'algorithme de Newton.*

Montrer que l'algorithme d'Howard peut s'interpréter comme un algorithme de Newton. Que peut-on dire de la convergence dans le cas régulier? Comparer avec la convergence des itérations sur les valeurs.

Exercice : *Etude numérique des équations d'Hamilton-Jacobi-Bellman* On considère l'équation de la Programmation dynamique stationnaire discrète

$$\min_{P \in \mathcal{P}} (A^P v(x) + u^P(x)) = 0, \quad \forall x \in \mathcal{E} \quad (6.43)$$

où \mathcal{E} est un espace d'états finis, $A^P = M^P - (1 + \lambda)I$, M^P est une matrice de Markov, $\lambda > 0$ est le taux d'actualisation et I désigne l'identité. Pour résoudre l'équation (6.43), on considère l'algorithme d'Howard défini ainsi : partant d'une sur-solution v^0 de (6.43), c'est à dire satisfaisant

$$\min_{P \in \mathcal{P}} (A^P v^0(x) + u^P(x)) \leq 0, \quad \forall x \in \mathcal{E}$$

on construit 2 suites (P^n) et (v^n) définies $\forall n \geq 1$ par

$$\text{Etape 1 : } P^n \in \underset{P \in \mathcal{P}}{\text{Argmin}} (A^P v^{n-1} + u^P) \quad (6.44)$$

$$\text{Etape 2 : } v^n \text{ est solution de } A^{P^n} v + u^{P^n} = 0. \quad (6.45)$$

Dans la suite, on dira qu'une matrice R est positive et l'on notera $R \geq 0$ si tous ses coefficients sont positifs ou nuls.

On dira que v est une sur-solution de (6.19) si $A^{P^n} v + u^{P^n} \leq 0$.

1. Démontrer que l'équation linéaire (6.19) peut s'écrire sous la forme :

$$v = U^{-1} L v + U^{-1} u^{P^n}$$

avec U inversible, $U^{-1} \geq 0$ et $L \geq 0$.

2. Pour résoudre l'équation (6.19), on utilise la méthode itérative suivante : à n fixé

$$v_{k+1}^n = U^{-1} L v_k^n + U^{-1} u^{P^n}, \quad k \geq 0 \quad (6.46)$$

avec v_0^n une sur-solution de l'équation (6.19). Démontrer que $\forall k \geq 0, v_k^n$ est une sur-solution de (6.19).

3. Démontrer que la suite $(v_k^n)_{k \geq 0}$ est décroissante et minorée et qu'elle converge vers la solution v^n de (6.19).
4. On considère l'algorithme suivant consistant à remplacer dans l'étape 2 de l'algorithme d'Howard, la résolution exacte de l'équation linéaire (6.19) par k_n itérations de l'algorithme de point fixe (6.46). Plus précisément, on définit $v^n = v_{k_n}^n$ avec $v_0^n = v^{n-1}$. Démontrer que la suite $(v^n)_{n \geq 0}$ est décroissante.
5. Vérifier que si $k_n = 1 \forall n \geq 0$, l'algorithme défini en question 4 correspond à la méthode d'itération sur les valeurs, et que si $k_n > 1$, il converge plus vite que la méthode d'itération sur les valeurs.

Annexe

Rappels de Probabilité

Un *espace de Probabilité* (Ω, \mathcal{F}, P) est défini par un ensemble Ω , une σ -algèbre \mathcal{F} (ou tribu) et une mesure de probabilité P . Une σ -algèbre \mathcal{F} est un sous-ensemble de l'ensemble des parties de Ω tel que

- $\forall A_i, i \in I$ dénombrable, $A_i \in \mathcal{F}$, alors $\bigcap_i A_i$ et $\bigcup_i A_i \in \mathcal{F}$,
- si $A \in \mathcal{F}$, alors $C_A \in \mathcal{F}$,
- \emptyset et $\Omega \in \mathcal{F}$

Les éléments de \mathcal{F} sont appelés des *événements*.

Une *mesure de probabilité* P sur (Ω, \mathcal{F}) est une mesure positive sur \mathcal{F} de masse totale égale à 1, i.e. une fonction de \mathcal{F} dans $[0, 1]$ telle que

- $P(\Omega) = 1$
- $P(\bigcup_n A_n) = \sum_n p(A_n)$ si les A_n sont disjoints.

Une *variable aléatoire* X est une fonction mesurable de (Ω, \mathcal{F}) à valeurs dans \mathbb{R} muni de la tribu Borélienne. La *loi* ou *distribution* de X notée P_X est définie par :

$$\forall A \in \mathbb{R}, P_X(A) = P(X^{-1}(A)) = P\{\omega, X(\omega) \in A\}.$$

Plus généralement, on parlera de vecteur aléatoire lorsque X est à valeurs dans \mathbb{R}^n . La *densité* de X est une fonction réelle f positive ou nulle telle que

$$\forall A \in \mathbb{R}, P_X(A) = \int_A f(x) dx.$$

Si $(X_k), k \in K$ est une famille de variables aléatoires, on note $\sigma(X_k, k \in K)$ la plus petite σ -algèbre pour laquelle les X_k sont mesurables.

On dit qu'une variable aléatoire X a une *distribution normale* de paramètre μ et σ^2 si

$$P_X(A) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_A \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

L'*espérance* (ou *moyenne*) d'une variable aléatoire X est définie par

$$E X = \int_{\Omega} X(\omega) dP(\omega).$$

La *variance* de X , notée $\text{var } X$, est définie par :

$$\text{var } X = E((X - E(X))^2) \equiv \sigma^2.$$

σ est appelé l'écart type.

Soit X et Y deux variables aléatoires ; la *covariance* de X et Y est définie par

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

Elle vaut 0 si les variables aléatoires sont indépendantes.

Le *coefficient de corrélation* est

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Théorème 6.8.1. Soit \mathcal{G} une sous-tribu de \mathcal{F} . Pour toute variable aléatoire intégrable X , il existe une variable aléatoire intégrable \mathcal{G} -mesurable Y , unique aux ensembles négligeables près, telle que :

$$\forall G \in \mathcal{G}, E(X1_G) = E(Y1_G).$$

Y est appelée *espérance conditionnelle* de X sachant \mathcal{G} et notée $E(X|\mathcal{G})$. Elle représente l'espérance de X connaissant les informations contenues dans la sous-tribu \mathcal{G} .

Soit \mathcal{G} est une sous-tribu finie; on peut lui associer une partition finie (G_1, \dots, G_n) de Ω par des éléments de \mathcal{F} , qui engendrent \mathcal{G} : les G_i , sont des éléments non vides de \mathcal{G} qui ne contiennent pas d'autre élément de \mathcal{G} qu'eux-mêmes et la partie vide. On les appelle les atomes de \mathcal{G} . On a $E(X|\mathcal{G}) = \sum_i \frac{E(X1_{G_i})}{P(G_i)} 1_{G_i}$, la somme étant limitée aux atomes de probabilité non nulle. Ainsi sur chaque atome G_i , la valeur de $E(X|\mathcal{G})$ est la valeur moyenne de X sur G_i . Dans le cas de la tribu grossière $\mathcal{G} = \{\emptyset, \Omega\}$, on a $E(X|\mathcal{G}) = E(X)$.

Propriétés de l'espérance conditionnelle :

- Si X est \mathcal{G} -mesurable, $E(X|\mathcal{G}) = X$,
- $E(E(X|\mathcal{G})) = E(X)$
- pour toute variable aléatoire Z \mathcal{G} -mesurable et bornée,

$$E(ZE(X|\mathcal{G})) = E(ZX)$$

et

$$E(ZX|\mathcal{G}) = ZE(X|\mathcal{G})$$

- linéarité

$$E(\lambda X + \mu Y|\mathcal{G}) = \lambda E(X|\mathcal{G}) + \mu E(Y|\mathcal{G})$$

- si $X \geq Y$, alors $E(X|\mathcal{G}) \geq E(Y|\mathcal{G})$
- Si \mathcal{H} est une sous-tribu de \mathcal{G} , alors

$$E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$$

- Si X est indépendante de \mathcal{G} alors, $E(X|\mathcal{G}) = E(X)$. (La réciproque est fautive).

Bibliographie

- [1] A. Bensoussan and J.L. Lions. *Applications des inéquations variationnelles en contrôle stochastique*. Dunod, 1978.
- [2] D.P. BERTSEKAS. *Dynamic Programming : Deterministic and Stochastic Models*. Prentice-Hall, Englewood Cliffs, N.J., 1987.
- [3] M.J. Brennan and E.S. Schwartz. The valuation of the american put option. *J. of finance*, 32 :449–462, 1977.
- [4] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms*. The MIT Press, 1989.
- [5] D. Dacunha-Castelle D. and M. Duflo. *Probabilité et Statistiques. Tome 2 : Problèmes à temps mobile*. Masson, 1983.
- [6] McCormick S.F. ed. *Multigrid methods*, volume 5 of *SIAM frontiers in applied mathematics*,. SIAM, 1987.
- [7] Shyriaev A. Folmer H., Protter P. Quadratic covariation and an extension of itô's formula. *Bernouilli*, (1) :149–169, 1995.
- [8] Barles G. *Solutions de viscosité des équations d'Hamilton-Jacobi-Bellman*, volume 17 of *Mathématiques et Applications*. édition ellipses, 1994.
- [9] W.K. Hastings. *Monte Carlo sampling methods using Markov chains and their applications*. Number 57. 1970.
- [10] Kushner H.J. and Dupuis P. *Numerical Methods for stochastic Control Problems in Continuous Time*. Springer Verlag, 1992.
- [11] P. Jaillet, D. Lamberton, and B. Lapeyre. Variationnal inequalities and the pricing of american options. *Acta Applicandae Mathematicae*, 21 :263–289, 1990.
- [12] J.Dieudonné. *Calcul Infinitésimal*. Hermann, Paris, 1980.
- [13] Quadrat J.P. Décision markovienne. unpublished manuscript.
- [14] T. Kato. *Perturbation of linear operators*. Springer-Verlag, 1966.
- [15] H.J. Kushner. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.
- [16] H.J. Kushner. Domain decomposition methods for large markov chain control problems and nonlinear elliptic-type equations. *Siam Journal on Scientific Computing*, 18(5) :1494–1516, 1997.
- [17] Akian M. Analyse de l'algorithme multigrille fmgh de résolution d'équations d'hamilton-jacobi-bellman. In A. Bensoussan and J. L. Lions, editors, *Analysis and Optimization of Systems*, volume 144 of *Lecture Notes in Contr. and Inf. Sci.*, pages 113–122. Springer Verlag, 1990. Proc. of the conf. on Analysis and Optimization of Systems.

- [18] Akian M. *Méthodes multigrilles en contrôle stochastique*. PhD thesis, Université Paris IX-Dauphine, 1990.
- [19] Akian M., Menaldi J.L., and Sulem A. Investment-consumption model with transaction costs. *SIAM J. of Control and Optim.*, 34(1), January 1996.
- [20] Davis M. and Norman A. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15 :676–713, 1990.
- [21] Crandall M.G., Ishii H., and Lions P.L. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27 :1–67, 1992.
- [22] J. Neveu. *Bases mathématiques du calcul des probabilités*. Masson, 1970.
- [23] Lions P.L. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, part 2 : Viscosity solutions and uniqueness. *Comm. in Partial Diff. Equ.*, 8(11) :1229–1276, 1983.
- [24] M.L Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. Probability and Mathematical Statistics : applied probability and statistics section. Wiley, 1994.
- [25] Bellman R. *Dynamic programming*. Princeton University Press, 1957.
- [26] Bellman R. *Introduction to the mathematical theory of control processes*. Academic Press, 1971.
- [27] P.A. Raviart and J.M. Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Masson, Paris, 1983.
- [28] L. Sainsaulieu. *Calcul scientifique*. Enseignement des mathématiques. Masson, 1996.
- [29] Feller W. *An introduction to probability theory and its application*. Wiley, 1957.
- [30] Hackbusch W. *Multigrid methods and applications*. Springer Verlag, Berlin, Heidelberg, 1985.
- [31] Hackbusch W. and Trottenberg U. eds. *Multigrid methods*, volume 960 of *Lecture notes in mathematics*. Springer Verlag, 1981.
- [32] Fleming W.H. and Soner H.M. *Controlled Markov Processes and Viscosity Solutions*. Springer Verlag, New York, 1993.
- [33] P. Whittle. *Optimization over Time, Dynamic Programming and Stochastic Control*. Wiley, New York, 1983.