

# Introduction à la Commande Stochastique

Jean-Pierre Quadrat

Michel Viot

Author address:

J.-P. QUADRAT, INRIA DOMAINE DE VOLUCEAU-ROCQUENCOURT,  
B.P. 105, 78153 LE CHESNAY CEDEX, FRANCE.

*E-mail address:* Jean-Pierre.Quadrat@inria.fr

M. VIOT, 01300 IZIEU, FRANCE.

1991 *Mathematics Subject Classification*. Commande Stochastique,  
Programmation Dynamique, Chaînes de Markov, Chaînes de Bellman,  
Agrégation, Perturbation, LQ, LQG.

## Table des matières

Chapitre 1. CHAÎNES DE MARKOV	9
1. Définitions	9
2. Exemples	9
3. Equation de Kolmogorov	12
4. Etude analytique des matrices	15
5. Problèmes ergodiques	20
6. Propriétés combinatoires	22
7. Base de $\mathcal{N}(A')$	23
8. Base de $\mathcal{N}(A)$	26
Chapitre 2. CHAÎNES DE BELLMAN	29
1. Algèbre max-plus	29
2. Matrices dans l'algèbre max-plus	31
3. Décision Optimale	39
4. Chaînes de Bellman	43
Chapitre 3. COMMANDE OPTIMALE STOCHASTIQUE	45
1. Introduction et exemples	45
2. Formulation précise du problème	46
3. Programmation dynamique en horizon fini	48
4. Programmation dynamique coût actualisé	50
5. Programmation dynamique ergodique	53
6. Commande en information incomplète	55
7. Réalisation et identification	57
Chapitre 4. PERTURBATION ET AGRÉGATION	61
1. Introduction et exemples	61
2. Chaînes de Markov perturbées	61
3. Commande des chaînes de Markov perturbées	66
Chapitre 5. DÉCOMPOSITION	73
1. Un Réseau de files d'attente	73
2. Factorisation de la mesure invariante	74
3. Feedbacks locaux optimaux	80
Chapitre 6. LE RÉGULATEUR LQG	85
1. Le filtre de Kalman	85
2. Le problème LQG	90
3. Le régulateur LQG en observation complète	92

4. Le problème LQG en observation incomplète	95
5. Rappels et compléments sur l'espérance conditionnelle	96
6. Rappel sur les variables aléatoires gaussiennes	97
Chapitre 7. PROPRIÉTÉS DES RÉGULATEURS LQ	99
1. Rappel sur le régulateur LQ	99
2. Placement de pôles par le régulateur LQ	100
3. Approche fréquentielle du régulateur LQ	102
4. Systèmes positifs et robustesse du régulateur LQ	106
5. Rappel sur la stabilité	110
Chapitre 8. PROBLÈMES	117
1. Une gestion de stock	117
2. Maintenance d'une automobile	121
3. Transport	126
4. Gestion de réservoir	129
5. Jeux de pile ou face	133
6. Processus de calcul	137
7. Filtrage et commande en observation incomplète	141
8. Stratégies	146
9. Clown équilibriste	149
Notations	165
Bibliographie	169
Index	173

Ce cours est une introduction à la commande optimale stochastique. On traite la commande des chaînes de Markov, à états en nombre fini, assez en détail. La commande des systèmes linéaires à temps discret avec bruits gaussiens et critère quadratique (LQG) est également étudié mais plus brièvement. L'étude de la robustesse des régulateurs obtenus est seulement esquissée.

On veut montrer dans ce cours comment formaliser les prises de décision en avenir incertain. Le point de vue adopté est celui de la recherche opérationnelle dans les parties concernant les chaînes de Markov et celui de l'automatique dans la partie concernant le LQG. Dans les deux cas il faut choisir une commande en "feedback" sur les événements effectivement réalisés.

Dans la première situation on veut minimiser un critère ayant un sens économique, par exemple, le coût de gestion d'un système de production, ou le coût de maintenance d'un atelier, ou le coût d'investissement échelonné dans le temps d'une entreprise — les variables de décision étant respectivement quand et combien produire, quand entretenir, quand et combien investir.

Dans la deuxième situation le but du feedback est d'améliorer la réaction, à un ordre, du système commandé. Le critère de qualité est la forme, dans le temps, de la réponse à l'ordre donné pour une famille plus ou moins bien spécifiée de systèmes, dans le voisinage d'un système nominal. Les critères ne sont plus alors que des intermédiaires de calcul sans grand sens physique. Par contre, d'autres préoccupations apparaissent : -la stabilité en temps, -la robustesse à des perturbations.

La différence entre les deux cas n'est sûrement pas aussi essentielle qu'on pourrait le croire a priori. Elle provient plus du point de vue que de la nature du phénomène. La saturation d'un amplificateur nous importe peu. Le débordement d'un grand barrage ou l'arrêt d'une centrale nucléaire nous paraît un beau gâchis.

Pour écrire ce cours nous avons fait des choix, nous avons exprimé un point de vue même si l'exposé semble très objectif — définition, théorème. Il n'est pas inutile, je crois, d'avertir le lecteur des choix conscients qui ont été faits et de leurs motivations.

On se place dans la situation mathématiquement la plus simple : -les chaînes de Markov à états en nombre fini, -les systèmes linéaires gaussiens. D'autres approches sont possibles comme les diffusions stochastiques. Si ce dernier cas est très formateur sur le plan mathématique, il noie, à notre avis, le point essentiel du débat dans un formalisme difficile, sans jamais apporter au bout du compte, jusqu'à présent, un résultat réellement nouveau. L'étude des diffusions doit venir, à notre avis, dans un second temps — pour ceux plus intéressés par les mathématiques soulevées que par la résolution pratique — lorsque ce qui est faisable et central dans les problèmes de commande a été bien identifié.

On parle des chaînes de Markov d'un point de vue ancien, c.a.d plus analytique que probabiliste. Les probabilités apparaissent surtout au niveau de la modélisation des motivations. Les raisonnements sont analytiques sur la loi de probabilité définie. C'est la façon de faire de Pallu De La Barrière dans son livre d'automatique, ou de Gantmacher dans son livre sur les matrices. On favorise souvent les manipulations des matrices au détriment de l'espérance conditionnelle. La raison est claire : la manipulation des matrices est mieux connue que l'espérance conditionnelle — bien que cette dernière s'applique à des situations beaucoup plus générales. De plus, le point de vue probabiliste est beaucoup plus adapté à la démonstration de théorèmes limites qu'aux méthodes numériques nécessaires aux calculs effectifs — dès que l'on sort des quelques cas très particuliers faisables à la main. Enfin la manipulation matricielle est à la base de toute l'ingénierie actuelle. Nous manipulerons, donc, surtout, les matrices stochastiques.

L'économie de temps faite par ces deux premiers choix permet d'aller plus loin dans la résolution numérique des problèmes. Privilégier cette résolution numérique est le troisième choix fait dans cet exposé. Il nous semble, en effet, sûrement à tort, que ces problèmes sont, en fait, relativement bien compris au fond. Par contre, il y a de grosses difficultés à les appliquer, dans le cas général, y compris en utilisant de gros ordinateurs, du fait de la complexité exponentielle du calcul avec la dimension du système. Il faut structurer, simplifier, agréger pour diminuer cette complexité. Nous avons indiqué quelques pistes dans cette direction : -méthodes de perturbations, -séparation des variables.

Le quatrième choix privilégie l'algèbre à l'analyse. Depuis quelques années, des structures algébriques sous-jacentes aux problèmes d'optimisation sont apparues. Elles font apparaître une dualité entre l'optimisation et les probabilités. Chaque fois que cela a été possible, dans ce cadre introductif, cette dualité a été explicitée.

Enfin la commande des chaînes de Markov est plus développée dans ce texte que la partie LQG. Ce choix a été fait pour deux raisons. Les structures discrètes sont plus faciles à utiliser que les espaces gaussiens. La théorie des systèmes est peu enseignée en dehors de certaines écoles d'ingénieur, malgré son grand intérêt pratique et théorique. Nous avons voulu donner, dans ce cours, seulement une ouverture à ce domaine.

Malgré son apparence spécialisée, cet exposé veut avoir une portée très générale. Les chaînes de Markov sont la meilleure illustration que nous connaissons de la notion d'*état*. La nécessité de *rétroactions* pour faire fonctionner un système s'est imposée dès l'antiquité et la *programmation dynamique* est le seul moyen général de les optimiser. La version déterministe, temps continue, de l'équation de la programmation dynamique n'est autre que l'équation d'*Hamilton-Jacobi* qui est centrale en mécanique analytique, elle même centrale en physique. Les *méthodes de perturbations* ont acquis depuis longtemps leurs lettres de noblesse en physique. Quant aux méthodes

de *séparation des variables* très liées à la théorie des groupes montrent à la fois la puissance et les limites de la recherche de solutions explicites.

Par ce cours nous avons voulu à la fois : -expliquer des outils utiles à des gestionnaires, -illustrer, dans une situation mathématiquement différente, des idées rencontrées en physique, -développer des outils fondamentaux en automatique.



## CHAPITRE 1

# CHAÎNES DE MARKOV

### 1. DÉFINITIONS

Une chaîne de Markov est définie à partir du quadruplet  $(\mathcal{T}, \mathcal{E}, M^n, p^0)$  où :

- $\mathcal{E} = \{1, 2, \dots, E\}$  désigne l'espace des états qui sera fini;
- $n \in \mathbb{N} = \mathcal{T}$  le temps;
- $M^n, n \in \mathcal{T}$  une famille de matrices  $E \times E$  appelées *matrices de transition* vérifiant :

$$M_{xy}^n \geq 0, \forall n \in \mathcal{T}, \forall x, y \in \mathcal{E},$$

$$\sum_{y \in \mathcal{E}} M_{xy}^n = 1, \forall n \in \mathcal{T}, \forall x \in \mathcal{E};$$

- $p^0$  une probabilité sur  $\mathcal{E}$  appelée *loi initiale*.

On appelle alors :

- $\Omega = \mathcal{E}^{\mathcal{T}}$  l'espace des trajectoires,  $\omega \in \Omega$ ;  $X^n(\omega) = \omega^n$  les *applications coordonnées*;
- $F^n = \{\mathcal{P}(\mathcal{E}^n) \times \{\mathcal{E}, \emptyset\}^{\mathbb{N}}\}$ ,  $n \in \mathcal{T}$  la famille de tribus croissantes sur  $\Omega$  rendant les applications coordonnées  $X^j, j \leq n$  mesurables;
- $F = \sigma \left\{ \bigcup_{n=1}^{\infty} F^n \right\}$  la tribu engendrée par les  $F^n$ .

Sur  $(\Omega, F)$  une loi de probabilité sera dite markovienne si :

$$\mathbb{P}(X^{n+1} = x^{n+1} | X^n = x^n, X^{n-1} = x^{n-1}, \dots, X^0 = x^0) = M_{x^n x^{n+1}}^n.$$

Si  $M^n = M, \forall n \in \mathcal{T}$ , on dira que la chaîne est *homogène* en temps. La loi de probabilité sur l'espace des trajectoires est définie de façon unique (grâce à un théorème de Kolmogorov) par ses lois marginales de dimension finie :

$$\mathbb{P}(X^{n+1} = x^{n+1}, \dots, X^0 = x^0) = p_{x^0}^0 \prod_{k=0}^n M_{x^k x^{k+1}}^k.$$

Pour illustrer ces définitions donnons quelques exemples.

### 2. EXEMPLES

## 2.1. JEU DE PILE OU FACE

Les variables aléatoires  $X^n$  représentent la fortune d'un joueur à pile ou face à l'instant  $n$  dont l'évolution est régie par  $X^{n+1} = X^n + Y^n$  où les  $Y^n$  forment une suite de variables aléatoires, dépendantes de  $X^n$ , à valeurs dans  $\{-1, 0, 1\}$  et de lois :

	$X^n = 0$	$0 < X^n < E$	$X^n = E$
$\mathbb{P}(Y^n = -1 X^n)$	0	1/2	0
$\mathbb{P}(Y^n = 0 X^n)$	1	0	1
$\mathbb{P}(Y^n = 1 X^n)$	0	1/2	0

Les  $X^n$  ainsi définies sont à valeurs dans  $\{0, 1, \dots, E\}$ . Lorsque  $X^n = 0$  le joueur est ruiné, il s'arrête. Lorsque  $X^n = E$  il a gagné, il s'arrête également. La chaîne de Markov décrivant l'évolution de la fortune du

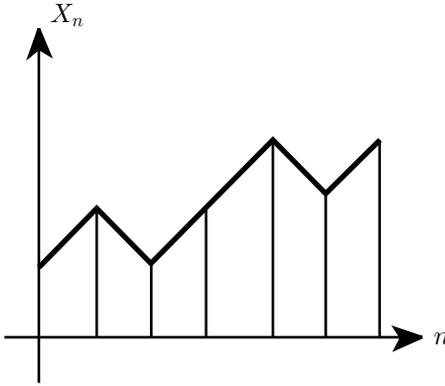


FIGURE 1. Trajectoire de la fortune du joueur

joueur est homogène en temps et de matrice de transition :

$$M = \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \cdots & \\ 0 & \frac{1}{2} & \ddots & & \\ \vdots & & & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & & & & 0 & 1 \end{bmatrix}$$

## 2.2. GESTION DE BARRAGE

On décrit ici l'évolution du niveau d'eau dans un barrage dont les apports sont aléatoires et sur lequel on peut agir en turbinant de l'eau pour produire de l'électricité.  $Y^n$  représente les apports,  $X^n$  la quantité d'eau en stock,  $Z^n$  la quantité turbinée. On suppose que les  $Y^n$  sont des variables aléatoires à valeurs dans  $\{0, 1\}$  indépendantes, de loi de probabilité  $(p, 1 - p)$ .

Les commandes

$$Z^n = \begin{cases} 1 & \text{si } X^n > E/2, \\ 0 & \text{si } X^n \leq E/2, \end{cases}$$

définissent une politique de gestion du réservoir en boucle fermée. Cela signifie que les décisions dépendent de l'état du système — ici la quantité en stock  $X^n$ .

On suppose la capacité de stockage égale à  $E$ . La matrice de transition décrivant l'évolution du stock d'eau s'écrit alors :

$$M^n = M = \begin{bmatrix} p & 1-p & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \\ \vdots & & p & 1-p & \\ \vdots & & p & 1-p & \\ 0 & & & \ddots & \ddots & p & 1-p \end{bmatrix} .$$

On désigne par  $c_{X^n Z^n}^n$  l'énergie produite à l'instant  $n$ , lorsque le stock est

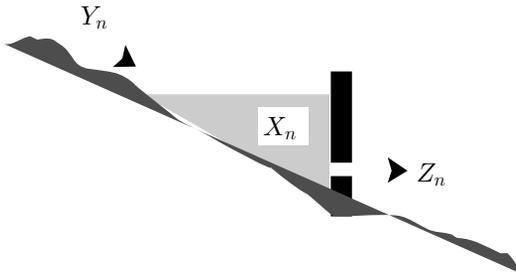


FIGURE 2. Evolution d'un stock d'eau

dans l'état  $X^n$  et la quantité turbinée vaut  $Z^n$ . On aimerait calculer :

$$\mathbb{E} \left\{ \sum_{n=0}^{N-1} c_{X^n Z^n}^n \mid X^0 = x^0 \right\} ,$$

c.a.d. l'espérance mathématique de l'énergie produite sur la période  $0, N$  sachant que le stock à l'instant 0 était  $x^0$ .

### 2.3. EVOLUTION DES POPULATIONS D'UN COUPLE DE PROIES ET DE PRÉDATEURS

On considère une population de proies et de prédateurs dont on veut prédire l'évolution et ses conséquences économiques. On note :

- $X_1^n$  la population de proies disons de souris;
- $X_2^n$  la population de prédateurs disons de chats;
- $Y_1^n$  une variable aléatoire à valeurs dans  $\{0, 1\}$  indiquant la naissance d'une souris entre  $i$  et  $i + 1$  :

$$\mathbb{P}(Y_1^n = 1 \mid X_1^n) = \begin{cases} \rho_1 X_1^n & \text{sur } X_1^n < E_1 , \\ 0 & \text{ailleurs ,} \end{cases}$$

- $Y_2^n$  une v.a. à valeurs dans  $\{0, 1\}$  indiquant la naissance d'un chat entre  $i$  et  $i + 1$  :

$$\mathbb{P}(Y_2^n = 1 | X_2^n) = \begin{cases} \rho_3 X_2^n & \text{sur } X_2^n < E_2, \\ 0 & \text{ailleurs,} \end{cases}$$

- $Z_1^n$  une v.a. à valeurs dans  $\{0, 1\}$  indiquant la mort d'une souris entre  $i$  et  $i + 1$ ,

$$\mathbb{P}(Z_1^n = 1 | X_1^n, X_2^n) = \rho_2 X_1^n X_2^n ;$$

- $Z_2^n$  une v.a. à valeurs dans  $\{0, 1\}$  représentant la mort d'un chat entre  $i$  et  $i + 1$ ,

$$\mathbb{P}(Z_2^n = 1 | X_1^n, X_2^n) = \rho_4 \frac{X_2^n}{1 + X_1^n} .$$

La population de souris est alors comprise entre 0 et  $E_1$  et évolue selon :

$$X_1^{n+1} = X_1^n + Y_1^n - Z_1^n .$$

La population de chats est comprise entre 0 et  $E_2$  et évolue selon :

$$X_2^{n+1} = X_2^n + Y_2^n - Z_2^n .$$

- EXERCICE 2.1. 1. Quelles sont les variables aléatoires à mémoriser pour rendre le processus markovien ?
2. Quelle est la matrice de transition correspondante ?
3. Si l'on désigne par  $c_{X_1^n}^n$  la consommation de grains des souris entre les instants  $n$  et  $n + 1$ , quelle est l'espérance mathématique de la perte totale de grains c.a.d. quelle est la valeur de :

$$\mathbb{E} \left\{ \sum_{n=0}^{N-1} c_{X_1^n}^n | X_1^0 = x_1^0, X_2^0 = x_2^0 \right\} ?$$

### 3. EQUATION DE KOLMOGOROV

#### 3.1. EQUATION DE KOLMOGOROV EN AVANT

PROPOSITION 3.1. Si  $\mathbb{P}$  désigne la mesure de probabilité d'une chaîne de Markov de matrice de transition  $M$  et de loi d'entrée  $p^0$  alors la loi marginale  $p_x^n$  définie par  $\mathbb{P}(X^n = x)$  est solution de :

$$\begin{cases} p^n = p^{n-1} M^{n-1} , \\ p^0 \text{ donné .} \end{cases} \quad (3.1)$$

PREUVE. Par définition d'une loi marginale on a :

$$p_x^n = (p^0 M^0 M^1 \dots M^{n-1})_x ,$$

d'où le résultat. □

## 3.2. EQUATION DE KOLMOGOROV ARRIÈRE

On considère l'évaluation de l'espérance mathématique d'une fonctionnelle additive de la trajectoire sur un horizon fini c.a.d. sur un nombre fini de périodes de temps. On évalue cette fonctionnelle par une équation récurrente rétrograde.

**PROPOSITION 3.2.** Etant donnée une chaîne de Markov de matrice de transition  $M^n$ , la fonctionnelle :

$$v_x^n = \mathbb{E} \left\{ \sum_{j=n}^{N-1} c_{X^j}^j \mid X^n = x \right\}$$

est solution de l'équation récurrente de Kolmogorov arrière :

$$\begin{cases} v^{n-1} &= c^{n-1} + M^{n-1}v^n, \\ v^N &= 0. \end{cases} \quad (3.2)$$

**PREUVE.** Démontrons ce résultat pour  $n = 0$ . On a :

$$v^0 = \sum_{n=0}^{N-1} \left( \prod_{j=0}^{n-1} M^j \right) c^n$$

grâce à l'équation de Kolmogorov avant et donc

$$v^0 = c^0 + M^0(c^1 + M^1(c^2 + M^2(c^3 + \dots + M^{N-1}c^{N-1}))),$$

d'où le résultat. □

## 3.3. EQUATION DE KOLMOGOROV ACTUALISÉE

On considère le même problème que précédemment mais cette fois sur un horizon infini et un coût par période actualisé c.a.d. décroissant de façon exponentielle avec le temps.

**PROPOSITION 3.3.** Etant donnée une chaîne de Markov homogène de matrice de transition  $M$ , la fonctionnelle

$$v_x = \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1 + \lambda)^{n+1}} c_{X^n} \mid X^0 = x \right\}$$

est solution de l'équation :

$$(A - \lambda)v + c = 0, \quad (3.3)$$

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

PREUVE. On a :

$$\begin{aligned}
 v_x &= \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} c_{X^n} \mid X^0 = x \right\}, \\
 &= \frac{1}{(1+\lambda)} c_x + \frac{1}{(1+\lambda)} \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} c_{X^{n+1}} \mid X^0 = x \right\}, \\
 &= \frac{1}{(1+\lambda)} c_x + \frac{1}{(1+\lambda)} \mathbb{E} \{ v_{X^1} \mid X^0 = x \}, \\
 &= \frac{1}{(1+\lambda)} c_x + \frac{1}{(1+\lambda)} (Mv)_x,
 \end{aligned}$$

d'où le résultat.  $\square$

### 3.4. EQUATION DE KOLMOGOROV ARRÊTÉ

On considère un problème du même type que celui du paragraphe précédent mais cette fois arrêté avant l'infini selon une procédure pouvant dépendre de l'état présent.

**DÉFINITION 3.4.** Une v.a.  $\nu : \Omega \rightarrow \mathbb{R}^+$  est un temps d'arrêt si  $\{\nu = n\}$  est  $F^n$  mesurable  $\forall n \in \mathbb{N}$ .

**EXEMPLE 3.5.** Les deux variables aléatoires suivantes sont des temps d'arrêt.

- $\nu(\omega) = n, \forall \omega \in \Omega$ .
- Etant donné un ensemble,  $B \subset \mathcal{E}$ ,  $\nu_B = \inf\{n : X^n \notin B\}$  est un temps d'arrêt appelé *temps de sortie* de l'ensemble  $B$ .

**EXERCICE 3.6.** Montrez que :

- $\nu_B$  est effectivement un temps d'arrêt;
- par contre  $\nu_B - 1$  (l'instant avant la première sortie) n'est pas un temps d'arrêt.

**PROPOSITION 3.7.** Etant donnée la chaîne de Markov homogène de générateur  $A$ ,  $B \subset E$ ,  $\nu_B$  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 suivant

$$\begin{cases} (A - \lambda)v + c = 0 & \text{sur } B, \\ v = \Phi/(1 + \lambda) & \text{sur } \mathbb{C}_B. \end{cases} \quad (3.4)$$

PREUVE. Il est immédiat d'adapter la preuve de la proposition précédente. Donnons une autre preuve plus technique mais montrant des idées utiles. Cette preuve consiste à trouver une interprétation stochastique à la solution du problème de Dirichlet.

Soit  $v$  solution de (3.4) on a alors :

$$\mathbb{E}^{F^n} \{v_{X^{n+1}} - (Mv)_{X^n}\} = 0, \text{ sur } X^n \in B,$$

à cause de la propriété de Markov de  $X_n$ .

Sur  $X^n \in B$  on a donc :

$$\begin{aligned} & \mathbb{E}^{F^n} \left\{ \frac{1}{(1+\lambda)^{n+1}} v_{X^{n+1}} - \frac{1}{(1+\lambda)^n} v_{X^n} \right\} \\ &= \mathbb{E}^{F^n} \left\{ \frac{1}{(1+\lambda)^{n+1}} (Mv)_{X^n} - \frac{1}{(1+\lambda)^n} v_{X^n} \right\}, \\ &= \mathbb{E}^{F^n} \left\{ \frac{1}{(1+\lambda)^{n+1}} ((A - \lambda I)v)_{X^n} \right\}, \\ &= -\mathbb{E}^{F^n} \left\{ \frac{1}{(1+\lambda)^{n+1}} c_{X^n} \right\}. \end{aligned}$$

Sur  $\mathcal{C}_B$  on définit  $c$  par  $c_x = -[(A - \lambda I)v]_x$  de façon à prolonger l'égalité précédente à tout  $\mathcal{E}$ . En notant :

$$R^{n+1} = \frac{1}{(1+\lambda)^{n+1}} v_{X^{n+1}} - v_{X^0} + \sum_{k=0}^n \frac{1}{(1+\lambda)^{k+1}} c_{X^k},$$

$R^n$  est martingale partant de 0.

Rappelons qu'une martingale  $R^n$  est un processus vérifiant

$$\mathbb{E}^{F^m} (R^n) = R^m, \quad \forall n, m, \quad n \geq m.$$

Alors  $R^{n \wedge v_B}$  est aussi une martingale partant de 0, en effet :

$$\mathbb{E}^{F^{n-1}} R^{n \wedge v_B} = \sum_{m=0}^{n-1} 1_{\{v_B=m\}} R^m + 1_{\{v_B \geq n\}} \mathbb{E}^{F^{n-1}} R^n = R^{(n-1) \wedge v_B},$$

puisque  $\{v_B \geq n\} = \mathcal{C}\{v_B \leq n-1\}$  est  $F^{n-1}$  mesurable.

On a donc :

$$\mathbb{E}^{F^0} \left\{ \frac{1}{(1+\lambda)^{n \wedge v_B}} v_{X_{n \wedge v_B}} + \sum_{k=0}^{n \wedge v_B - 1} \frac{1}{(1+\lambda)^{k+1}} c_{X^k} \right\} = v_{X^0}.$$

En passant alors à la limite en  $n$  on obtient :

$$\mathbb{E}^{F^0} \left\{ 1_{\{v_B < \infty\}} \frac{1}{(1+\lambda)^{v_B+1}} \Phi_{X^{v_B}} + \sum_{k=0}^{v_B-1} \frac{1}{(1+\lambda)^{k+1}} c_{X^k} \right\} = v_{X^0},$$

d'où le résultat, puisque  $1/(1+\lambda)^{v_B} = 0$  sur  $\{v_B = \infty\}$ .  $\square$

## 4. ETUDE ANALYTIQUE DES MATRICES

On étudie les matrices d'un point de vue analytique. On obtient ainsi des résultats qui peuvent être vus comme des corollaires de la mise sous forme de Jordan d'une matrice. L'approche donnée ici présente l'intérêt de se généraliser à la dimension infinie; elle est tirée du livre de Kato sur la perturbation des opérateurs. Comme corollaire de ces résultats on obtient la semi-simplicité des valeurs propres de module 1 des matrices stochastiques. On a alors les outils nécessaires pour comprendre le comportement asymptotique des chaînes de Markov.

## 4.1. DÉVELOPPEMENT DE LA RÉSOVANTE D'UN OPÉRATEUR AUTOUR D'UNE VALEUR PROPRE

On se donne un opérateur  $A$  (ici une matrice  $E \times E$ ).

**DÉFINITION 4.1.** On note  $R^\lambda$  la résolvante de l'opérateur  $A$ , définie par  $R^\lambda(A) = (A - \lambda I)^{-1}$  où  $\forall \lambda \in \mathbb{C} \setminus \mathcal{S}_A$  avec  $\mathcal{S}_A$  le spectre de  $A$  — c.a.d. l'ensemble des  $\lambda \in \mathbb{C}$  où  $A - \lambda I$  est singulière. On a alors l'égalité des résolvantes :

$$R^{\lambda_1} - R^{\lambda_2} = (\lambda_1 - \lambda_2) R^{\lambda_1} R^{\lambda_2}. \quad (4.1)$$

PREUVE.

$$\begin{aligned} R^{\lambda_1} - R^{\lambda_2} &= R^{\lambda_1} (A - \lambda_2) R^{\lambda_2} - R^{\lambda_1} (A - \lambda_1) R^{\lambda_2} \\ &= (\lambda_1 - \lambda_2) R^{\lambda_1} R^{\lambda_2}. \end{aligned}$$

□

**PROPOSITION 4.2.** L'application  $\lambda \rightarrow R^\lambda$  est analytique sur  $\mathbb{C} \setminus \mathcal{S}_A$ , et n'a pas de pôle à l'infini.

PREUVE. Grâce à l'égalité des résolvantes on a :

$$R^{\lambda_0} = R^\lambda + R^{\lambda_0} - R^\lambda = [1 - (\lambda - \lambda_0) R^{\lambda_0}] R^\lambda,$$

et donc :

$$\begin{aligned} R^\lambda &= [1 - (\lambda - \lambda_0) R^{\lambda_0}]^{-1} R^{\lambda_0}, \\ &= \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n (R^{\lambda_0})^{n+1}. \end{aligned}$$

La série est convergente dès que  $|\lambda - \lambda_0| < \frac{1}{|R^{\lambda_0}|}$  et donc  $R^\lambda$  est analytique dans  $\mathbb{C} \setminus \mathcal{S}_A$ .

D'autre part :

$$R^\lambda = (A - \lambda)^{-1} = -\lambda^{-1} (1 - \lambda^{-1} A)^{-1} = -\sum_{n=0}^{+\infty} \lambda^{-n-1} A^n$$

est convergente dès que  $\lambda > |A|$ .

□

**THÉORÈME 4.3.** Dans un voisinage de la valeur propre  $\lambda_0$  de  $A$ , on a le développement

$$R^\lambda = -(\lambda - \lambda_0)^{-1}P - \sum_{n=1}^{m-1} (\lambda - \lambda_0)^{-n-1} N^n + \sum_{n=0}^{+\infty} (\lambda - \lambda_0)^n S^{n+1}, \quad (4.2)$$

avec :

- $P$  projecteur spectral sur l'espace propre  $V$  associé à la valeur propre  $\lambda_0$ , supposé de dimension  $m$  ;
- $N$  le nilpotent associé à la valeur propre  $\lambda_0$  ;
- $S$  vérifiant  $S(A - \lambda_0) = (A - \lambda_0)S = I - P$  donc pseudo inverse de  $A$  sur le supplémentaire de l'espace propre  $V$ .

PREUVE. On développe  $R^\lambda$  en série de Laurent au voisinage de  $\lambda = \lambda_0$ , on obtient :

$$R^\lambda = \sum_{n=-\infty}^{+\infty} (\lambda - \lambda_0)^n B_n,$$

où :

$$B_n = \frac{1}{2\pi i} \int_{\Gamma} (\lambda - \lambda_0)^{-n-1} R^\lambda d\lambda,$$

et  $\Gamma$  est un cercle entourant la valeur propre  $\lambda_0$  à l'exclusion de toutes les autres, orienté dans le sens positif. En prenant  $\Gamma'$  un cercle entourant  $\Gamma$  ayant les mêmes

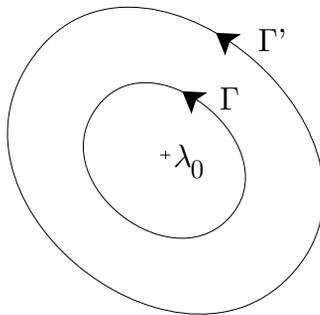


FIGURE 3. Contour d'intégration

propriétés que  $\Gamma$  on obtient en utilisant l'égalité des résolvantes :

$$\begin{aligned} B_n B_m &= \frac{1}{(2\pi i)^2} \int_{\Gamma'} \int_{\Gamma} (\lambda - \lambda_0)^{-n-1} (\lambda' - \lambda_0)^{-m-1} R^\lambda R^{\lambda'} d\lambda d\lambda' \\ &= \frac{1}{(2\pi i)^2} \int_{\Gamma'} \int_{\Gamma} \frac{R^{\lambda'} - R^\lambda}{(\lambda - \lambda_0)^{n+1} (\lambda' - \lambda_0)^{m+1} (\lambda' - \lambda)} d\lambda d\lambda'. \end{aligned}$$

D'autre part on a :

$$\frac{1}{(2\pi i)} \int_{\Gamma} (\lambda - \lambda_0)^{-n-1} (\lambda' - \lambda)^{-1} d\lambda = \eta_n (\lambda' - \lambda_0)^{-n-1} ,$$

avec

$$\begin{cases} \eta_n = 1 & \text{si } n \geq 0 , \\ \eta_n = 0 & \text{si } n < 0 . \end{cases}$$

En effet :

- si  $n < 0$ , donc si  $-n - 1 \geq 0$ , on a l'analyticit  de  $(\lambda - \lambda_0)^{-n-1}$ , et comme  $\Gamma'$  est   l'ext rieur de  $\Gamma$ , la formule de Cauchy montre que  $\eta_n = 0$ ;
- si  $n \geq 0$ , posant  $\mu = 1/(\lambda - \lambda_0)$ , par ce changement de variable  $\Gamma$  devient  $\Gamma_1$ ,  $\Gamma'$  devient  $\Gamma'_1$ , avec  $\Gamma'_1 \subset \Gamma_1$  et on a :

$$\begin{aligned} & \frac{1}{(2\pi i)} \int_{\Gamma} (\lambda - \lambda_0)^{-n-1} (\lambda' - \lambda)^{-1} d\lambda \\ &= -\frac{1}{(2\pi i)} \int_{\Gamma_1} \mu^{n+1} \left(\frac{1}{\mu'} - \frac{1}{\mu}\right)^{-1} \frac{d\mu}{\mu^2} \\ &= -\frac{1}{(2\pi i)} \int_{\Gamma_1} \mu^{n+1} \frac{\mu\mu'}{\mu - \mu'} \frac{d\mu}{\mu^2} \\ &= -\frac{1}{(2\pi i)} \int_{\Gamma_1} \mu' \frac{\mu^n}{\mu - \mu'} d\mu \\ &= (\mu')^{n+1} \\ &= \frac{1}{(\lambda' - \lambda_0)^{n+1}} . \end{aligned}$$

De m me on a :

$$\frac{1}{(2\pi i)} \int_{\Gamma'} (\lambda - \lambda_0)^{-m-1} (\lambda' - \lambda)^{-1} d\lambda' = (1 - \eta_m) (\lambda - \lambda_0)^{-m-1}$$

et donc on a :

$$B_n B_m = \frac{\eta_n + \eta_m - 1}{2\pi i} \int_{\Gamma} (\lambda - \lambda_0)^{-n-m-2} R^\lambda d\lambda , \quad (4.3)$$

$$= (\eta_n + \eta_m - 1) B_{n+m+1} . \quad (4.4)$$

Posons alors :

$$P = -B_{-1} ,$$

$$N = -B_{-2} ,$$

$$S = B_0 .$$

En particulierisant la formule fondamentale (4.4) on va obtenir le r sultat annonc  dans le th or me.

Nous rassemblons ces particularisations dans le tableau :

n	m	$\Rightarrow$	$\Rightarrow$	commentaire
-1	-1	$B_{-1}^2 = -B_{-1}$	$P^2 = P$	P est une projection
-2	-2	$B_{-2}^2 = -B_{-3}$	$B_{-3} = -N^2$	
-2	-3	$B_{-2}B_{-3} = -B_{-4}$	$B_{-4} = -N^3$	$B_{-n} = -N^{n-1} \quad n \geq 2$
0	0	$B_0B_0 = B_1$	$B_1 = S^2$	
1	0	$B_0B_1 = B_0^3 = B_2$	$B_2 = S^3$	$B_n = S^{n+1} \quad n \geq 0$
-1	-2	$B_{-1}B_{-2} = -B_{-2}$	$PN = NP = N$	
-1	0	$B_{-1}B_0 = 0$	$PS = SP = 0$	
-2	0	$B_{-2}B_0 = 0$	$NS = SN = 0$	

On a donc

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

Comme  $P$  est une projection, on a la décomposition de  $F = V \oplus W$  avec  $V = PF$  et  $W = (I - P)F$ . En posant :

$$Q^\lambda = -(\lambda - \lambda_0)^{-1}P - \sum_{n=1}^{+\infty}(\lambda - \lambda_0)^{-n-1}N^n,$$

$Q^\lambda$  est la partie principale d'une série de Laurent autour d'une singularité isolée.  $Q^\lambda$  est donc convergente pour  $\lambda \neq \lambda_0$ , et donc le rayon spectral de  $N$  est nul, et donc  $N$  est nilpotent, et donc, si  $m = \dim(PF)$ , on a  $N^m = 0$ .

Si l'on note par un indice  $i$  les opérateurs  $N, P, Q$  associés aux autres valeurs propres, on montre de même que :

$$P_i P_k = \delta_{ik} P_i,$$

$$\sum_k P_k = I,$$

$$R^\lambda = \sum_i Q_i^\lambda.$$

En remarquant alors que :

$$(A - \lambda_0)R^\lambda = (A - \lambda + \lambda - \lambda_0)(A - \lambda)^{-1} = 1 + (\lambda - \lambda_0)R^\lambda,$$

$$\begin{aligned} (A - \lambda_0)B_n &= \frac{1}{(2\pi i)} \int_{\Gamma} \frac{1}{(\lambda - \lambda_0)^{n+1}} R^\lambda (A - \lambda_0) d\lambda, \\ &= \frac{1}{(2\pi i)} \int_{\Gamma} \frac{1}{(\lambda - \lambda_0)^{n+1}} (1 + (\lambda - \lambda_0)R^\lambda) d\lambda, \\ &= \delta_{n0} + B_{n-1}, \end{aligned}$$

et donc pour  $n = 0$  :

$$(A - \lambda_0)S = S(A - \lambda_0) = 1 - P,$$

et pour  $n = -1$  :

$$P(A - \lambda_0) = (A - \lambda_0)P = N,$$

ce qui revient à dire que :

$$AP = \lambda_0 P + N,$$

on a obtenu la décomposition spectrale de  $A$ .

□

## 4.2. APPLICATION AUX MATRICES STOCHASTIQUES

Une matrice stochastique  $M$  est une probabilité de transition d'une chaîne de Markov. Elle vérifie donc :

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

$$\sum_{y \in \mathcal{E}} M_{xy} = 1, \quad \forall x \in \mathcal{E}.$$

PROPOSITION 4.4. Si l'on note :

- $|v|_\infty = \sup_{x \in \mathcal{E}} |v_x|$  pour  $v \in \mathbb{R}^E$ ,
- $|A|_\infty = \sup_{|v|_\infty \leq 1} |Av|_\infty$  pour  $A \in \mathbb{R}^{E \times E}$ ,

pour toute matrice stochastique  $M$ , on a  $|M|_\infty = 1$ , et donc toutes les valeurs propres de  $M$  ont des modules inférieurs à 1.

PREUVE.

$$|Mv|_\infty = \sup_x \left| \sum_y M_{xy} v_y \right| \leq (\sup_y |v_y|) (\sup_x \sum_y M_{xy}) = \sup_y |v_y|.$$

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

Du théorème donnant le développement de la résolvante, valide pour les matrices générales, on déduit le corollaire suivant sur les matrices stochastiques :

COROLLAIRE 4.5. Les nilpotents associés aux valeurs propres de module 1 d'une matrice stochastique sont nuls.

PREUVE. Soient,  $\lambda_0$  une valeur propre de module 1 de  $M$  — matrice stochastique — et  $\lambda = \lambda_0 \alpha \in \mathbb{C}$ ,  $\alpha > 1 \in \mathbb{R}$ . Alors

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

Et donc  $v_x^\lambda$  est borné  $\forall \lambda$ ,  $|\lambda| > 1$ . Or  $v^\lambda$  est solution de l'équation de Kolmogorov :

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

et donc

$$v^\lambda = -(\lambda - \lambda_0)R^\lambda(M)c \leq \sup_x |c_x|,$$

ce qui n'est possible, d'après le théorème précédent, que si le nilpotent de  $M$  associé à la valeur propre  $\lambda$  est nul.  $\square$

## 5. PROBLÈMES ERGODIQUES

Nous pouvons étudier maintenant l'équation de Kolmogorov dans le cas ergodique (c.a.d. lorsque le taux d'actualisation tend vers 0, ou plus classiquement le cas du coût moyen par unité de temps).

**COROLLAIRE 5.1.** Si l'on note

$$v_x^\lambda = \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{\lambda}{(1+\lambda)^{n+1}} c_{X^n} | X^0 = x \right\}, \quad \lambda \in \mathbb{R}^+,$$

on a  $\lim_{\lambda \rightarrow 0} v^\lambda = v = Pc$  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  $X^n$ .

**PREUVE.** D'après l'équation de Kolmogorov actualisée  $v^\lambda$ ,  $\lambda > 0$  est solution de

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

où  $A = M - I$ . La matrice  $A$  admet la valeur propre 0 et le nilpotent associé est nul grâce au corollaire (4.5). Comme  $v^\lambda = -\lambda R^\lambda(A)c$  le développement (4.2) donne le résultat.  $\square$

**PROPOSITION 5.2.** Le vecteur  $v = \lim_{\lambda \rightarrow 0} v^\lambda$  pour  $v^\lambda$  défini au corollaire (5.1) est l'unique solution d'un des systèmes d'équations équivalents suivant :

$$(v, q) = (c, q), \quad v \in \mathcal{N}(A), \quad \forall q \in \mathcal{N}(A'), \quad (5.1)$$

$$Aw + c = v, \quad v \in \mathcal{N}(A), \quad w \in \mathcal{R}(A). \quad (5.2)$$

où  $\mathcal{N}(A)$  désigne le nayau de  $A$  et  $\mathcal{R}(A)$  désigne l'image de  $A$ .

**PREUVE.** Dans le corollaire on a montré que  $v = Pc$ , prenons  $w = -Sc$ , alors (5.2) devient :

$$-ASc + c = Pc,$$

qui se déduit de  $AS = SA = I - P$ .

Montrons que (5.2)  $\Rightarrow$  (5.1). Pour cela il suffit de faire le produit scalaire de (5.2) par  $q$  et d'utiliser  $(q, Aw) = (A'q, w) = 0$  qui provient du fait que  $q \in \mathcal{N}(A')$ .

Montrons l'unicité de la solution de (5.1). Pour cela supposons qu'il y ait deux solutions. Notons  $z$  la différence, on a :

$$(z, q) = 0, \quad z \in \mathcal{N}(A), \quad \forall q \in \mathcal{N}(A'),$$

et donc  $z \in \mathcal{R}(A)$  et donc  $\mathcal{R}(A) \cap \mathcal{N}(A) \neq 0$  et donc il existe  $u$  tel que  $v = Au \in \mathcal{R}(A) \neq 0$  et  $Av = 0$  et donc  $A^2u = 0$ ,  $Au \neq 0$ , et donc le

nilpotent de  $A$  associé à la valeur propre 0 est non nul ce qui contredit le corollaire 4.5.  $\square$

EXERCICE 5.3. Démontrer que (5.1)  $\Rightarrow$  (5.2).

On a également un théorème ergodique qui peut être vu comme un corollaire du théorème 4.2.

COROLLAIRE 5.4 (Théorème ergodique).

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

où  $M$  désigne une matrice stochastique et  $P$  le projecteur spectral sur l'espace propre associé à la valeur propre 1 de  $M$ .

PREUVE.  $M = \sum_i (\lambda_i P_i + N_i)$  où  $\lambda_i$  désigne les valeur propres de  $M$ ,  $P_i$  les projecteurs spectraux et  $N_i$  les nilpotents associés.

On a  $\lim_{m \rightarrow \infty} (\lambda_i P_i + N_i)^m = 0$  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.  $\square$

REMARQUE 5.5. Puisque l'on a

$$\sum_{n=0}^{+\infty} \frac{\lambda}{(1 + \lambda)^{n+1}} = 1 ,$$

notant  $\nu$  une variable aléatoire sur  $\mathbb{N}$ , indépendante de  $X^n$ , de loi définie par :

$$\mathbb{P}\{\nu = n\} = \frac{\lambda}{(1 + \lambda)^{n+1}}$$

on a :

$$v_x^\lambda = \mathbb{E}\{c_{X^\nu} | X^0 = x\} .$$

D'autre part,

$$\mathbb{E}\nu = \sum_{n=0}^{+\infty} \frac{n\lambda}{(1 + \lambda)^{n+1}} = \frac{1}{\lambda} .$$

Et donc lorsque  $\lambda \rightarrow 0$ ,  $\mathbb{E}\nu \rightarrow \infty$ , ce qui montre que le corollaire (5.1) est une sorte de théorème ergodique dans lequel on a remplacé la loi

uniforme sur  $[0, m - 1]$  par une loi exponentielle et le passage à la limite sur  $m$  par un passage à la limite sur le taux de l'exponentielle.

EXERCICE 5.6. Vérifier la formule :  $\mathbb{E}v = \frac{1}{\lambda}$ .

## 6. PROPRIÉTÉS COMBINATOIRES

Etant donnée une chaîne de Markov de matrice de transition  $M$ , on lui associe l'application :

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

c.a.d l'application qui à un noeud du graphe (associé à la chaîne de Markov) fait correspondre l'ensemble des noeuds accessibles par une transition de probabilité non nulle.

Un chemin de longueur  $n$  allant de  $x^0$  à  $x^n \in \mathcal{E}$  sera la donnée de  $x^0 \cdots x^n$  avec  $x^k \in \Gamma(x^{k-1})$ . Un circuit sera un chemin revenant à son point de départ. Les noeuds sont vus comme des chemins de longueur zéro. On définit alors :

- 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  $\complement A$ ,
- une chaîne de Markov ayant une seule classe est dite *irréductible*.

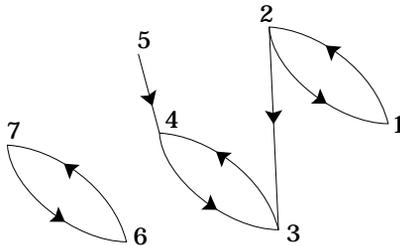


FIGURE 4. Exemple de chaîne de Markov

EXEMPLE 6.1. La chaîne de la figure 4 a 2 classes transitaires (1,2) (5) et 2 classes finales (3,4) (6,7). La sous chaîne (6,7) est irréductible.

7. BASE DE  $\mathcal{N}(A')$ 

Soit une chaîne de Markov de matrice de transition  $M$ , notons  $A = M - I$  son générateur. Un élément  $p$  de  $\mathcal{N}(A')$  vérifie  $A'p = 0$  donc  $pA = 0$  donc  $pM = p$  c'est donc une mesure invariante par  $M$ . On veut décrire, dans ce paragraphe, l'ensemble des mesures de probabilités invariantes par  $M$ .

**DÉFINITION 7.1.** Un opérateur  $A$  satisfait au :

- *principe du maximum [resp. principe du minimum] si :*

$$(Av)_x \leq 0, \forall x \in \arg \max_{x \in \mathcal{E}} v_x \text{ [resp. } (Av)_x \geq 0, \forall x \in \arg \min_{x \in \mathcal{E}} v_x \text{];}$$

- *principe du maximum positif [resp. principe du maximum strictement positif] si et seulement si :*

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

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

**PROPOSITION 7.2.** Si  $M$  est une matrice stochastique  $A = M - I$  satisfait au principe du maximum et du minimum et  $A - D$ ,  $D$  diagonale  $D_{ii} \geq 0$  [resp.  $> 0$ ], satisfait au principe du maximum [resp. strictement] positif — les matrices  $M - D$  sont dites sous-stochastiques.

**PREUVE.** Soit  $x \in \arg \max_{x \in \mathcal{E}} \{v_x\}$  on a alors :

$$(Av)_x = \sum_y M_{xy} v_y - v_x \leq \sum_y M_{xy} v_x - v_x \leq 0,$$

puisque  $M_{xy} \geq 0$ ,  $\sum_y M_{xy} = 1$  et  $v_y \leq v_x$ ,  $\forall y$ . □

**REMARQUE 7.3.** On peut définir de même le principe du minimum négatif [resp. strictement négatif] vérifié par les matrices stochastiques [resp. sous stochastiques].

Etant donné un vecteur  $p$ , il peut toujours s'écrire  $p = p^+ - p^-$  avec  $p_x^+ = \max(p_x, 0)$  et  $p_x^- = \max(-p_x, 0) \forall x \in \mathcal{E}$ . On note aussi  $p^+$  et  $p^-$  les vecteurs de dimension plus petite constitués des seuls éléments non nuls. Avec cet abus de notation et un changement de numérotation on notera également  $p = (p^+, p^-)$ .

**LEMME 7.4.**

$$p \in \mathcal{N}(A') \Rightarrow p^+ \in \mathcal{N}(A'), p^- \in \mathcal{N}(A').$$

**PREUVE.** On a la partition :

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

On a :

$$pA = \begin{bmatrix} p^+ A_{11} - p^- A_{21} & p^+ A_{12} - p^- A_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix},$$

mais  $-p^- A_{21} \leq 0$  car tous les coefficients de  $A_{21}$  sont positifs. On a alors  $p^- A_{21} = 0$  sinon  $(-p^- A_{21}, 1) < 0$  ce qui implique que  $(p^+ A_{11}, 1) > 0$  mais  $(p^+ A_{11}, 1) = (p^+, A_{11} 1) \leq 0$  d'après le principe du maximum positif appliqué à la sous matrice  $A_{11}$ . De même on a  $p^+ A_{12} = 0$  et donc le résultat.  $\square$

**THÉORÈME 7.5.** Il existe une base  $(p^1, p^2, \dots, p^m)$  de  $\mathcal{N}(A')$  telle que :

- les  $p^\alpha$  sont des mesures de probabilités de supports disjoints. Ces mesures de probabilités seront appelées dans la suite *mesure de probabilités invariantes extrêmes* (MPIE) de  $M$ ,
- leurs supports sont les classes finales de la chaîne de Markov.

PREUVE. Du lemme précédent découle l'existence d'une base de  $\mathcal{N}(A')$  : notée  $(q^1, \dots, q^m)$  telle que :

$$q_x^\alpha \geq 0, \quad \sum_x q_x^\alpha = 1. \quad (7.1)$$

En effet étant donnée une base quelconque  $(e_1 \cdots e_m)$ ,

$$(e_1^+, e_1^-, e_2^+, e_2^- \cdots e_m^+, e_m^-)$$

appartient à  $\mathcal{N}(A')$  et engendre  $\mathcal{N}(A')$  et on peut en extraire une base. Il suffit alors de normaliser cette base pour obtenir la propriété désirée.

Si les  $q^\alpha$  n'ont pas leurs supports disjoints on peut construire une nouvelle base ayant cette propriété. En effet supposons que  $q^\alpha$  et  $q^\beta$  soient dans ce cas alors  $\mu q^\alpha - \lambda q^\beta \in \mathcal{N}(A')$ ,  $\forall \lambda, \mu$  et donc  $(r^1, r^2, r^3)$  définis par :

$$r^1 = \begin{cases} q^\alpha & \text{sur } \mathbb{C}_{\text{supp } q^\alpha \cap \text{supp } q^\beta}^{\text{supp } q^\alpha}, \\ 0 & \text{ailleurs,} \end{cases}$$

$$r^2 = \begin{cases} q^\alpha & \text{sur } \text{supp } q^\alpha \cap \text{supp } q^\beta, \\ 0 & \text{ailleurs,} \end{cases}$$

$$r^3 = \begin{cases} q^\beta & \text{sur } \mathbb{C}_{\text{supp } q^\alpha \cap \text{supp } q^\beta}^{\text{supp } q^\beta}, \\ 0 & \text{ailleurs,} \end{cases}$$

génère l'espace vectoriel engendré par les  $q^\alpha$  et  $q^\beta$  et appartiennent à  $\mathcal{N}(A')$  grâce au lemme précédent.

On construit ainsi une base  $(r^1, r^2, \dots, r^m)$  de vecteurs vérifiant (7.1) et ayant leurs supports disjoints. Si l'on note alors  $f_1 \cdots f_m$  leurs supports et  $t = \mathbb{C}_{f_1 \cup f_2 \cup \dots \cup f_m}^\mathcal{E}$ ,  $A$  se réécrit :

$$A = \begin{array}{c|c|c|c|c} & f_1 & f_2 & f_m & t \\ \hline f_1 & A_1 & 0 & 0 & 0 \\ \hline f_2 & 0 & A_2 & 0 & 0 \\ \hline f_m & 0 & 0 & A_m & 0 \\ \hline t & A_{t1} & . & . & A_{tt} \end{array} .$$

En effet si l'on note :  $B_{ji} = \{M_{xy}, x \in f_j, y \in f_i\}$  alors  $B_{ji} = 0$  pour  $j \neq i$  sinon en prenant  $p = r^i - r^j$  on aurait :

$$(pA)_{f_i} = r^i A_i - r^j B_{ji} = 0 \Rightarrow B_{ji} = 0 .$$

En effet dans le cas contraire on aurait :  $r^j B_{ji} \geq 0$  avec  $(r^j B_{ji}, 1) > 0$ , mais alors  $(r^i A_i, 1) > 0$  or  $(r^i A_i, 1) = (r^i, A_i 1) \leq 0$  puisque  $A_i 1 \leq 0$  d'après le principe du maximum positif.

D'autre part les supports des MPIE sont des ensembles clos et donc des unions de classes closes minimales, et donc la dimension de  $\mathcal{N}(A')$  est plus petite que le nombre de classes finales. Inversement pour chaque classe finale  $f$ , notons  $A_f$  la restriction de  $A$  à l'ensemble  $f$ ,  $A_f$  est encore un générateur d'une chaîne de Markov, et a donc au moins une mesure de probabilité invariante  $q$ , alors :

$$r = \begin{cases} q & \text{sur } f , \\ 0 & \text{ailleurs ,} \end{cases}$$

vérifie  $rA = 0$  et la dimension de  $\mathcal{N}(A')$  est plus grande que le nombre de classes finales, d'où le résultat.  $\square$

REMARQUE 7.6.  $f_1, f_2, \dots, f_m$  sont donc les classes finales, et donc  $t$  est l'ensemble des états transitoires.

## 8. BASE DE $\mathcal{N}(A)$

Une base de  $\mathcal{N}(A)$  ayant une interprétation probabiliste peut être explicitée. Ce sont les probabilités d'atteindre une classe finale donnée partant d'un état transitoire. On a vu au paragraphe précédent que le générateur d'une chaîne de Markov a la structure suivante :

$$A = \begin{array}{c|cccc} & f_1 & f_2 & f_m & t \\ \hline f_1 & A_1 & 0 & 0 & 0 \\ \hline f_2 & 0 & A_2 & 0 & 0 \\ \hline f_m & 0 & 0 & A_m & 0 \\ \hline t & A_{t1} & . & . & A_{tt} \end{array} ,$$

où :

- $f_1, f_2, \dots, f_m$  sont les classes finales,
- $t$  est l'ensemble des états transitoires,
- $\dim(\mathcal{N}(A')) = \dim(\mathcal{N}(A))$  est égal au nombre de classes finales.

**PROPOSITION 8.1.** Toutes les valeurs propres de  $M_{tt} = A_{tt} + I$  sont de module strictement inférieur à 1.

PREUVE. Puisque les états appartenant à  $t$  sont transitoires, il existe un chemin de longueur  $m$  allant de tout élément  $x \in t$  à une classe finale  $f_i$ , mais alors tout chemin de longueur plus grande que  $m$ , va encore de  $t$  à  $f_i$

puisque le chemin précédent se prolonge par un chemin dans  $f_i$ , il existe alors  $n$  tel que

$$(M)^n = \begin{array}{c|cccc} & f_1 & f_2 & f_m & t \\ \hline f_1 & (M_1)^n & 0 & 0 & 0 \\ \hline f_2 & 0 & (M_2)^n & 0 & 0 \\ \hline f_m & 0 & 0 & (M_m)^n & 0 \\ \hline t & B_{t1} & . & . & (M_{tt})^n \end{array}$$

possède dans chaque ligne indexée dans  $t$ , un élément non nul en dehors du bloc diagonal  $M_{tt}$ . Et donc la somme des coefficients en ligne de  $(M_{tt})^n < 1$  et donc  $|(M_{tt})^n|_\infty < 1$  d'où le résultat.  $\square$

**THÉORÈME 8.2.** Une base de  $\mathcal{N}(A)$  est constituée par les vecteurs  $\chi^j$  solutions uniques des problèmes de Dirichlet :

$$\begin{cases} A\chi^j = 0 \text{ sur } t, \\ \chi^j = 1 \text{ sur } f_j, \\ \chi^j = 0 \text{ ailleurs.} \end{cases}$$

PREUVE. En utilisant la structure de la matrice  $A$  il suffit de montrer que :

$$(A_{tj}1^{f_j} + A_{tt}\chi^j)_x = 0, \quad \forall x \in t. \quad (8.1)$$

où  $1^{f_j}$  désigne le vecteur dont les composantes sont égales à 1 et définies pour les états appartenant à  $f_j$ . Mais  $A_{tt}$  est inversible puisque  $A_{tt}$  a ses valeurs propres strictement négatives et donc  $\chi_x^j = (-A_{tt}^{-1}A_{tj}1^{f_j})_x$ ,  $x \in t$  est la solution cherchée.  $\square$

**REMARQUE 8.3.** Potentiel d'une chaîne transitoire. La matrice :

$$R = \sum_{n=0}^{+\infty} (M_{tt})^n$$

est appelé *potentiel* de la chaîne transitoire de matrice de transition  $M_{tt}$ . On remarque que puisque  $|M_{tt}| < 1$ ,  $R < \infty$  on a  $\chi^j = RA_{tj}1^{f_j}$  en utilisant  $-A_{tt}^{-1} = (I - (I + A_{tt}))^{-1}$ .

**REMARQUE 8.4.** Interprétation de  $\chi^j$ .  $\chi_x^j$  s'interprète comme la probabilité d'aller dans  $f_j$  partant de l'état transitoire  $x$ .

$$\chi^j = \left[ \sum_{n=0}^{+\infty} (M_{tt})^n \right] A_{tj}1^{f_j}$$

est la probabilité d'y aller en une étape, plus la probabilité d'y aller en deux étapes, plus etc...



## CHAPITRE 2

# CHAÎNES DE BELLMAN

On construit dans ce chapitre un formalisme similaire au calcul des probabilités mais adapté à l'optimisation grâce à un changement d'algèbre. Aux chaînes de Markov correspondent alors des problèmes de programmation dynamique déterministe.

### 1. ALGÈBRE MAX-PLUS

#### 1.1. DÉFINITIONS

**DÉFINITION 1.1 (Demi-corps).** Un *demi-corps*  $\mathcal{K}$  est un ensemble muni de deux opérations  $\oplus$  and  $\otimes$  telles que :

- l'opération  $\oplus$  soit associative, commutative et possède un *zéro*  $\varepsilon$  ;
- l'opération  $\otimes$  définisse un groupe sur  $\mathcal{K}_* \stackrel{\text{def}}{=} \mathcal{K} \setminus \{\varepsilon\}$ , soit distributive par rapport à  $\oplus$  et son *unité*  $e$  vérifie  $\varepsilon \otimes e = e \otimes \varepsilon = \varepsilon$ .

On dit que le demi-corps est

- *idempotent* si la première opération est idempotente, c.a.d. si

$$a \oplus a = a, \forall a \in \mathcal{K} ;$$

- *commutatif* si le groupe est commutatif.

**THÉORÈME 1.2.** Le zéro  $\varepsilon$  d'un demi-corps idempotent est *absorbant* pour la seconde opération, c.a.d.  $\varepsilon \otimes a = a \otimes \varepsilon = \varepsilon, \forall a \in \mathcal{K}$ .

**PREUVE.** On a

$$\varepsilon = \varepsilon e = \varepsilon(\varepsilon \oplus e) = \varepsilon^2 \oplus \varepsilon = \varepsilon^2, ,$$

alors,

$$\forall a \in \mathcal{K}_*, \quad \varepsilon = \varepsilon e = \varepsilon a^{-1} a = \varepsilon(a^{-1} \oplus \varepsilon)a = \varepsilon a^{-1} a \oplus \varepsilon^2 a = \varepsilon^2 a = \varepsilon a .$$

□

**DÉFINITION 1.3.** Le symbole  $\mathbb{R}_{\max}$  [resp.  $\overline{\mathbb{R}}_{\max}$ ] désigne l'ensemble  $\mathbb{R} \cup \{-\infty\}$  [resp.  $\mathbb{R} \cup \{\infty, -\infty\}$ ] muni des deux opérations binaires  $\max$  et  $+$  (avec  $\infty - \infty = -\infty$ ) notées  $\oplus$  and  $\otimes$  respectivement. On a  $\varepsilon = -\infty$  et  $e = 0$ .

On appelle cette structure l'algèbre max-plus. On remarque que l'ordre naturel sur  $\mathbb{R}_{\max}$  peut être défini au moyen de l'opération  $\oplus$

$$a \leq b \quad \text{si} \quad a \oplus b = b .$$

**THÉORÈME 1.4.** La structure algébrique  $\mathbb{R}_{\max}$  est un demi-corps idempotent.

Si l'on compare les propriétés de  $\oplus$  et de  $\otimes$  avec celles de  $+$  et de  $\times$ , nous voyons que :

1. nous avons perdu la symétrie de l'addition (pour  $a$  donné, il n'existe pas d'éléments  $b$  tel que  $\max(b, a) = -\infty$  dès que  $a \neq -\infty$ );
2. nous avons gagné l'idempotence de l'addition.

Si l'on essaie de faire des calculs algébriques dans cette structure on s'aperçoit que l'idempotence est aussi utile que l'existence d'un élément symétrique pour la simplification des formules. Par exemple, l'analogue de la formule binomiale

$$(a + b)^n = \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{n-1} a b^{n-1} + \binom{n}{0} b^n$$

est

$$n \max(a, b) = \max(na, nb) .$$

D'un autre côté le  $\max$  n'est pas simplifiable  $\max(a, b) = b$  n'implique pas que  $a = \varepsilon$ .

## 1.2. NOTATION

Pour renforcer l'analogie avec le calcul conventionnel  $\max$  est noté  $\oplus$ , et  $+$  est noté  $\otimes$ . On introduit aussi le symbole  $\not\circlearrowleft$  pour l'inverse de l'opération de  $+$  qui joue le rôle de la multiplication (qui est donc le moins habituel). Par conséquent  $a \not\circlearrowleft b$  signifie  $a - b$ . On omet parfois  $\otimes$  si cela ne conduit pas à des confusions. La table suivante permet de vérifier que l'on a bien compris les notations.

$\mathbb{R}_{\max}$	Notations conventionnelles	=
$2 \oplus 3$	$\max(2, 3)$	3
$1 \oplus 2 \oplus 3 \oplus 4 \oplus 5$	$\max(1, 2, 3, 4, 5)$	5
$2 \otimes 3 = 5$	$2 + 3$	5
$2 \oplus \varepsilon$	$\max(2, -\infty)$	2
$\varepsilon = \varepsilon \otimes 2$	$-\infty + 2$	$-\infty$
$(-1) \otimes 3$	$-1 + 3$	2
$e \otimes 3$	$0 + 3$	3
$3^2 = 2^3$	$3 \times 2 = 2 \times 3 =$	6
$e = e^2 = 2^0$	$0 \times 2 = 2 \times 0$	0
$(2 \otimes 3) \not\circlearrowleft (2 \oplus 3)$	$(2 + 3) - \max(2, 3)$	2
$(2 \oplus 3)^3 = 2^3 \oplus 3^3$	$3 \times \max(2, 3) = \max(3 \times 2, 3 \times 3)$	9
$6 \not\circlearrowleft e$	$6 - 0$	6
$e \not\circlearrowleft 3$	$0 - 3$	-3
$\sqrt[2]{8}$	$8/2$	4
$\sqrt[5]{15}$	$15/5$	3

**EXERCICE 1.5.** Exprimer  $\min(a, b)$  en utilisant uniquement  $\oplus$  et  $\otimes$ .

## 2. MATRICES DANS L'ALGÈBRE MAX-PLUS

Dans cette section on étudie les systèmes linéaires. On est capable de résoudre explicitement les systèmes de la forme  $x = Ax \oplus b$  et  $Ax = b$  (le système général étant de la forme  $Ax \oplus b = Cx \oplus d$ ). On étudie également la théorie spectrale des matrices. Il y a une notion de vecteur propre et valeur propre satisfaisante mais il n'existe souvent qu'une seule valeur propre.

### 2.1. FONCTIONS SCALAIRES LINÉAIRES ET AFFINES

**DÉFINITION 2.1 (Fonction linéaire).** La fonction  $f : \mathbb{R}_{\max} \rightarrow \mathbb{R}_{\max}$  est *linéaire* si elle satisfait

$$f(c) = c \otimes f(e) , \quad \forall c \in \mathbb{R}_{\max} .$$

Par conséquent toute fonction linéaire est de la forme  $y = f(c) = a \otimes c$ , où  $a = f(e)$ . Les graphes de telles fonctions sont des droites de pente un et d'ordonnée  $a$  en 0.

**DÉFINITION 2.2 (Fonction affine).** La fonction  $f : \mathbb{R}_{\max} \rightarrow \mathbb{R}_{\max}$ ,

$$f(c) = ac \oplus b, \quad a, b \in \mathbb{R}_{\max}$$

est dite *affine*.

Son graphe est le sup de la fonction constante et d'une droite de pente 1. Le premier problème à résoudre est de trouver la solution de l'équation

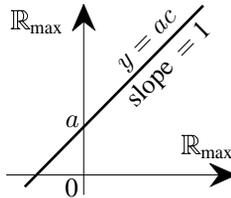


FIGURE 1. Une fonction linéaire

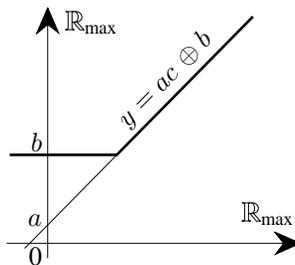


FIGURE 2. Une fonction affine

affine générale.

**DÉFINITION 2.3 (Equation affine).** L'équation *affine scalaire* générale est

$$ax \oplus b = a'x \oplus b' . \quad (2.1)$$

En effet puisque  $\oplus$  n'a pas d'inverse, l'équation (2.1) ne peut pas être réduite à  $ax \oplus b = \varepsilon$ .

**THÉORÈME 2.4.** La solution de l'équation scalaire affine générale est la suivante :

- Si

$$((a' < a) \text{ et } (b < b')) \text{ ou } ((a < a') \text{ et } (b' < b)) \tag{2.2}$$

sont vrais alors la solution est unique et donnée par

$$x = (b \oplus b') \not\wedge (a \oplus a') ;$$

- Si  $a \neq a', b \neq b'$ , et (2.2) n'est pas vérifiée, il n'existe pas de solution dans  $\mathbb{R}_{\max}$ ;
- Si  $a = a'$  and  $b \neq b'$ , la solution n'est pas unique et toutes les solutions sont données par  $x \geq (b \oplus b') \not\wedge a$ ;
- Si  $a \neq a'$  et  $b = b'$ , la solution n'est pas unique et toutes les solutions sont données par  $x \leq b' \not\wedge (a \oplus a')$ ;
- Si  $a = a'$  et  $b = b'$ , tout  $x \in \mathbb{R}$  est solution.

La preuve est immédiate à partir de l'interprétation géométrique voir figure 3. En pratique, il est préférable de simplifier (2.1) avant de la résoudre.

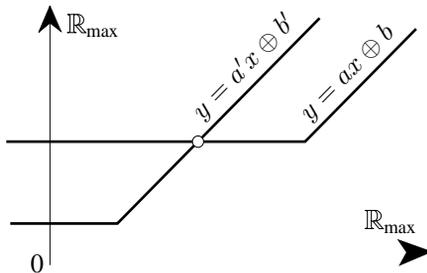


FIGURE 3. Une équation affine

Par exemple, si  $a > a'$  and  $b' > b$ , alors  $ax \oplus b = a'x \oplus b' \Leftrightarrow ax = b'$ .

### 2.2. STRUCTURES VECTORIELLES

La structure de moduloïde joue le rôle d'espace vectoriel lorsqu'on remplace le corps des scalaires par un demi-corps idempotent.

**DÉFINITION 2.5 (Moduloïde).** Un *moduloïde*  $\mathcal{M}$  sur un demi-corps idempotent  $\mathcal{K}$  (avec ses deux opérations  $\oplus$  and  $\otimes$ , son zéro  $\varepsilon$  et son unité  $e$ ) est un ensemble muni d'

- une opération interne notée aussi  $\oplus$  avec un zéro noté  $\varepsilon$ ;
- une opération externe définie sur  $\mathcal{K} \times \mathcal{M}$  à valeurs dans  $\mathcal{M}$ , notée par la juxtaposition des symboles du scalaire et du vecteur, qui satisfait les propriétés suivantes :
  1.  $\oplus$  est associatif, commutatif;
  2.  $\alpha(x \oplus y) = \alpha x \oplus \alpha y$ ;

$$3. (\alpha \oplus \beta)x = \alpha x \oplus \beta x;$$

$$4. \alpha(\beta x) = (\alpha\beta)x;$$

$$5. ex = x;$$

$$6. \varepsilon x = \varepsilon;$$

pour tout  $\alpha, \beta \in \mathcal{K}$  et tout  $x, y \in \mathcal{M}$ .

On utilise cette structure dans le cas particulier suivant.

EXEMPLE 2.6.  $(\mathbb{R}_{\max})^n$  est un moduloïde sur  $\mathbb{R}_{\max}$ . Son zéro, encore noté  $\varepsilon$ , vaut  $(\varepsilon, \dots, \varepsilon)$ .

DÉFINITION 2.7 (Algèbre Idempotente). Un moduloïde muni d'une opération interne notée  $\otimes$  est appelé *algèbre idempotente* si cette opération est associative, a une unité notée  $e$  et si elle est distributive par rapport à  $\oplus$ .

EXEMPLE 2.8. Soit  $(\mathbb{R}_{\max})^{n \times n}$  l'ensemble des matrices  $n \times n$  à coefficients dans  $\mathbb{R}_{\max}$  muni des deux opérations internes suivantes :

- l'addition terme à terme notée  $\oplus$ ;
- la multiplication matricielle notée  $\otimes$  définie par

$$(A \otimes B)_{ij} = \bigoplus_{k=1}^n A_{ik} \otimes B_{kj} ;$$

et l'opération externe

- $\forall \alpha \in \mathbb{R}_{\max}, \forall A \in (\mathbb{R}_{\max})^{n \times n}, \alpha A = (\alpha A_{ij})$ .

L'ensemble  $(\mathbb{R}_{\max})^{n \times n}$  est une algèbre idempotente dont le zéro est la matrice (encore notée  $\varepsilon$ ) ayant tous ses termes égaux à  $\varepsilon$ , dont l'identité est la matrice (encore notée  $e$ ) ayant sur la diagonale des  $e$  et partout ailleurs valant  $\varepsilon$ .

### 2.3. SYSTÈMES D'ÉQUATIONS LINÉAIRES DANS $(\mathbb{R}_{\max})^n$

Dans cette sous-section nous nous intéressons aux systèmes d'équations linéaires. Nous utilisons les notations matricielles pour les définir. Le système général s'écrit

$$Ax \oplus b = Cx \oplus d ,$$

où  $A$  et  $C$  sont des  $n \times n$ -matrices et  $b$  et  $d$  sont des  $n$ -vecteurs. Ce système peut être mis sous une forme simplifiée.

DÉFINITION 2.9 (Forme canonique d'un système). Le système

$$Ax \oplus b = Cx \oplus d$$

est dit être sous *forme canonique* si  $A, C, b$ , and  $d$  satisfont

- $C_{ij} = \varepsilon$  si  $A_{ij} > C_{ij}$ , et  $A_{ij} = \varepsilon$  if  $A_{ij} < C_{ij}$ ;
- $d_i = \varepsilon$  si  $b_i > d_i$ , et  $b_i = \varepsilon$  if  $b_i < d_i$ .

EXEMPLE 2.10. Considérons le système

$$\begin{pmatrix} 3 & 2 \\ \varepsilon & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} e \\ 3 \end{pmatrix} ,$$

qui peut être mis sous la forme simplifiée

$$\begin{pmatrix} \varepsilon & 2 \\ \varepsilon & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} = \begin{pmatrix} 4 & \varepsilon \\ 1 & \varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} \varepsilon \\ 3 \end{pmatrix} ,$$

ce qui implique

$$\left. \begin{array}{l} 2x_2 \oplus 1 = 4x_1 \\ 2x_2 = 1x_1 \oplus 3 \end{array} \right\} \Rightarrow 4x_1 = 1x_1 \oplus 3 \Rightarrow 4x_1 = 3 ,$$

$$\Rightarrow x_1 = -1 \Rightarrow x_2 = 1 .$$

Le système a une solution. En général, un système linéaire peut avoir ou ne pas avoir de solution. De plus même s'il existe une solution elle peut ne pas être unique.

Il y a deux classes de systèmes linéaires pour lesquels on a une théorie satisfaisante ce sont

- $x = Ax \oplus b$ ;
- $Ax = b$ .

Commençons par étudier le premier cas

### 2.3.1. SOLUTION DE $x = Ax \oplus b$ .

**DÉFINITION 2.11.** On appelle  $\mathcal{G}(A)$  le graphe de précedence associé à la matrice  $A$ . Ses noeuds sont  $\{1, \dots, n\}$  si la matrice  $A$  appartient à  $(\mathbb{R}_{\max})^{n \times n}$ . Il possède l'arc joignant  $i$  à  $j$  de poids  $A_{ij}$  si  $A_{ij} \neq \varepsilon$ .

**THÉORÈME 2.12.** Si, dans  $\mathcal{G}(A')$ , les poids des circuits sont négatifs ou nuls, il existe une solution à  $x = Ax \oplus b$  qui est donnée par  $x = A^*b$  avec

$$A^* \stackrel{\text{def}}{=} \bigoplus_0^{\infty} A^i = \bigoplus_0^{n-1} A^i .$$

De plus, si les poids des circuits sont strictement négatifs, la solution est unique.

**PREUVE.** Si  $A^*b$  existe, c'est une solution puisque

$$A(A^*b) \oplus b = (e \oplus AA^*)b = A^*b .$$

**EXISTENCE DE  $A^*b$ .** On peut interpréter  $[A^*]_{ji}$  comme le poids maximum des chemins de longueur quelconque joignant  $i$  à  $j$  dans  $\mathcal{G}(A')$ . Par conséquent, une CNS d'existence de  $[A^*]_{ij}$  est l'absence de circuit de poids positif dans  $\mathcal{G}(A')$ .

**UNICITÉ DE LA SOLUTION.** Supposons que  $x$  soit solution de  $x = Ax \oplus b$ . Alors  $x$  satisfait

$$\begin{aligned} x &= b \oplus Ab \oplus A^2x , \\ x &= b \oplus Ab \oplus \dots \oplus A^{k-1}b \oplus A^kx , \end{aligned} \quad (2.3)$$

et par conséquent  $x \geq A^*b$ . De plus, si tous les circuits ont des poids négatifs, alors  $A^k \rightarrow \varepsilon$  quand  $k \rightarrow \infty$ . En effet, les termes de  $A^k$  sont les poids des chemins de longueur  $k$  qui nécessairement contiennent des circuits de

$\mathcal{G}(A')$  dont le nombre tend vers l'  $\infty$  avec  $k$ . Mais les poids de ces circuits sont tous négatifs. Utilisant cette propriété dans l'équation (2.3), pour  $k$  suffisamment grand, on obtient que  $x = A^*b$ .

PREUVE DE  $A^* = e \oplus A \oplus \dots \oplus A^{n-1}$ . Tous les chemins de longueur supérieure ou égale à  $n$  sont composés nécessairement d'un circuit et d'un chemin de longueur strictement inférieure à  $n$ . Puisque les circuits ont des poids négatifs, par hypothèse, on a

$$\forall m \geq n, \quad A^m \leq e \oplus \dots \oplus A^{n-1} .$$

□

REMARQUE 2.13. Si le poids maximum des circuits est 0, une solution existe, mais il n'y a pas unicité. Par exemple, l'équation  $x = x \oplus b$  admet la solution  $x = e^*b = b$  mais tout  $x > b$  est aussi solution.

EXEMPLE 2.14. Considérons le système d'ordre 2

$$x = \begin{pmatrix} -1 & 2 \\ -3 & -1 \end{pmatrix} x \oplus \begin{pmatrix} e \\ 2 \end{pmatrix} .$$

Alors,

$$b = \begin{pmatrix} e \\ 2 \end{pmatrix}, \quad Ab = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad A^2b = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

$$A^3b = \begin{pmatrix} 3 \\ e \end{pmatrix}, \quad A^4b = \begin{pmatrix} 2 \\ e \end{pmatrix} \dots .$$

Par conséquent,

$$x = \begin{pmatrix} e \\ 2 \end{pmatrix} \oplus \begin{pmatrix} 4 \\ 1 \end{pmatrix} \oplus \begin{pmatrix} 3 \\ 1 \end{pmatrix} \oplus \begin{pmatrix} 3 \\ e \end{pmatrix} \oplus \begin{pmatrix} 2 \\ e \end{pmatrix} \oplus \dots = \begin{pmatrix} 4 \\ 2 \end{pmatrix},$$

est la solution unique.

On remarque de plus que  $A^2b, A^3b, \dots$  sont plus petits que  $b \oplus Ab$ .

2.3.2. SOLUTION DE  $Ax = b$ . La seconde classe de systèmes linéaires pour laquelle on dispose de résultats satisfaisants est  $Ax = b$ . Cependant, on doit se placer dans  $\overline{\mathbb{R}}_{\max}$  plutôt que  $\mathbb{R}_{\max}$ , et on doit affaiblir la notion de solution. Une *sous-solution* de  $Ax = b$  est un  $x$  qui satisfait  $Ax \leq b$ , où l'ordre sur les vecteurs est défini par  $x \leq y$  si  $x \oplus y = y$ .

THÉORÈME 2.15. Etant donnée une  $n \times n$ -matrice  $A$  et un  $n$ -vecteur  $b$  à coefficients dans  $\overline{\mathbb{R}}_{\max}$ , la plus grande sous-solution de  $Ax = b$  existe. Elle est donnée par :

$$-x_j = \max_i (-b_i + A_{ij}) .$$

PREUVE. On a

$$\begin{aligned} \{Ax \leq b\} &\Leftrightarrow \left\{ \bigoplus_j A_{ij} x_j \leq b_i, \forall i \right\} \\ &\Leftrightarrow \{x_j \leq b_i - A_{ij}, \forall i, j\} \\ &\Leftrightarrow \{x_j \leq \min_i [b_i - A_{ij}], \forall j\} \\ &\Leftrightarrow \left\{ -x_j \geq \max_i [-b_i + A_{ij}], \forall j \right\} . \end{aligned}$$

Inversement on vérifie que le vecteur  $x$  défini par

$$-x_j = \max_i [-b_i + A_{ij}], \forall j ,$$

est une sous-solution. Par conséquent c'est la plus grande.  $\square$

Pour résoudre  $Ax = b$ , on peut alors calculer la plus grande sous-solution et vérifier si c'est une solution ou non.

EXEMPLE 2.16. Calculons la plus grande sous-solution de l'égalité :

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 7 \end{pmatrix} .$$

Suivant les considérations précédentes on calcule matriciellement l'opposé de la plus grande sous-solution

$$\begin{pmatrix} -6 & -7 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & -2 \end{pmatrix} .$$

Alors la plus grande sous-solution vaut  $(x_1, x_2) = (3, 2)$ . En effet d'une part

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix} \leq \begin{pmatrix} 6 \\ 7 \end{pmatrix} ,$$

d'autre part il est facile de vérifier que la seconde inégalité n'est plus satisfaite si on augmente  $x_1$  et/ou  $x_2$ . Par conséquent la première inégalité ne peut pas être réduite à une égalité.

## 2.4. THÉORIE SPECTRALE

Etant donnée une matrice  $A$  à coefficients dans  $\mathbb{R}_{\max}$ , on étudie l'existence de valeurs propres et de vecteurs propres, c.a.d., de l'existence d'éléments non nuls  $\lambda$  et  $x$  tels que :

$$Ax = \lambda x . \quad (2.4)$$

Le résultat principal est le suivant

THÉORÈME 2.17. Si  $\mathcal{G}(A)$  est fortement connexe, il existe une et une seule valeur propre (il peut y avoir plusieurs vecteurs propres). La valeur propre est égal au poids moyen maximum des circuits du graphe :

$$\lambda = \max_{\zeta} \frac{|\zeta|_w}{|\zeta|_1} ,$$

où,  $\zeta$  parcourt l'ensemble des circuits de  $\mathcal{G}(A')$ , on note  $|\zeta|_w$  le poids du circuit  $\zeta$  et  $|\zeta|_l$  sa longueur.

PREUVE. EXISTENCE DE  $x$  ET DE  $\lambda$ . Considerons la matrice  $B = A/\lambda \stackrel{\text{def}}{=} (e/\lambda)A$ , où  $\lambda = \max_{\zeta} |\zeta|_w / |\zeta|_l$ . Le poids maximum des circuits de  $\mathcal{G}(B')$  est  $e$ . Par conséquent  $B^*$  and  $B^+ = BB^*$  existent. La matrice  $B^+$  a une colonne possédant un  $e$  sur la diagonale. Pour prouver cette affirmation choisissons un noeud  $k$  d'un circuit  $\xi$  tel que  $\xi \in \arg \max_{\zeta} |\zeta|_w / |\zeta|_l$ . Le poids maximum des chemins allant de  $k$  à  $k$  est  $e$ . Et donc nous avons  $e = B_{kk}^+$ . Soit  $B_{\cdot k}$  la  $k$ -ème colonne de  $B$ . Alors, puisque  $B^+ = BB^*$  et  $B^* = e \oplus B^+$  ( $e$  est la matrice unité), pour ce  $k$  on a

$$B_{\cdot k}^+ = B_{\cdot k}^* \Rightarrow BB_{\cdot k}^* = B_{\cdot k}^+ = B_{\cdot k}^* \Rightarrow AB_{\cdot k}^* = \lambda B_{\cdot k}^* .$$

Par conséquent  $x = B_{\cdot k}^* = B_{\cdot k}^+$  est un vecteur propre de  $A$  associé à la valeur propre  $\lambda$ .

L'ensemble des noeuds de  $\mathcal{G}(A')$  correspondant aux coefficients non nuls de  $x$  est appelé le *support* de  $x$ .

INTERPRÉTATION DE  $\lambda$  EN TERME DE GRAPHE. Si  $\lambda$  satisfait (2.4), il existe un coefficient non nul de  $x$ , disons  $x_1$ . Alors on a  $(Ax)_1 = \lambda x_1$  et il existe un coefficient  $x_{i_1} \neq \varepsilon$  tel que  $A_{1i_1} x_{i_1} = \lambda x_1$ . Alors  $(Ax)_{i_1} = \lambda x_{i_1}$  et il existe un coefficient  $x_{i_2}$  tel que  $A_{i_1 i_2} x_{i_2} = \lambda x_{i_1}$ , etc. jusqu'à ce que l'on obtienne une deuxième fois le même coefficient  $x_{i_l}$ . De cette façon on a défini un circuit  $\beta = (i_l, i_m, \dots, i_{l+1}, i_l)$ . Par multiplication le long du circuit, on obtient

$$A_{i_l i_{l+1}} A_{i_{l+1} i_{l+2}} \dots A_{i_m i_l} x_{i_{l+1}} x_{i_{l+2}} \dots x_{i_m} x_{i_l} = \lambda^{m-l+1} x_{i_l} x_{i_{l+1}} \dots x_{i_m} .$$

Puisque  $x_k \neq \varepsilon$  pour tout  $k$ , on peut simplifier l'équation ci dessus et on obtient que  $\lambda^{m-l+1}$  est le poids du circuit de longueur  $m - l + 1$ , ou, dit autrement,  $\lambda$  est le poids moyen du circuit  $\beta$ . Cet argument n'utilise pas la forte connexité du graphe.

SI  $\mathcal{G}(A')$  EST FORTEMENT CONNEXE LE SUPPORT DE  $x$  EST L'ENSEMBLE DE TOUS LES NOEUDS DE GRAPHE. Supposons que le support de  $x$  ne soit pas le graphe tout entier. Il existerait alors des arcs partant du support de  $x$  et joignant d'autres noeuds, en effet  $\mathcal{G}(A')$  a seulement une composante fortement connexe par hypothèse. Le support de  $Ax$  serait alors plus grand que le support de  $x$  ce qui serait en contradiction avec l'équation (2.4).

UNICITÉ DANS LE CAS FORTEMENT CONNEXE. Prenons un circuit  $\gamma = (i_1, \dots, i_p, i_1)$  tel que ses noeuds appartiennent au support de  $x$  (ici tous les noeuds de  $\mathcal{G}(A')$ ). On a

$$A_{i_2 i_1} x_{i_1} \leq \lambda x_{i_2} , \quad \dots , \quad A_{i_p i_{p-1}} x_{i_{p-1}} \leq \lambda x_{i_p} , \quad A_{i_1 i_p} x_{i_p} \leq \lambda x_{i_1} .$$

Par les mêmes arguments que ceux utilisés pour l'interprétation (dans cette preuve), on voit que  $\lambda$  est plus grand que le poids moyen de  $\gamma$ . Par conséquent  $\lambda$  est le poids moyen maximum des circuits et donc  $\lambda$  est unique.  $\square$

Il est important de comprendre le rôle du support de  $x$  dans cette preuve. Si  $\mathcal{G}(A')$  n'est pas fortement connexe, le support de  $x$  n'est pas nécessairement l'ensemble des noeuds tout entier et en général il n'y a plus unicité (voir l'exemple 2.20 ci dessous).

REMARQUE 2.18. La partie interprétation de la preuve montre que pour une matrice générale  $A$ , toute valeur propre est égale au poids moyen d'un circuit. Par conséquent le poids moyen maximum des circuits est égal à la valeur propre maximum de la matrice correspondante.

EXEMPLE 2.19. Sous les hypothèses du théorème 2.17, l'unicité du vecteur propre n'est pas garantie comme le montre l'exemple suivant

$$\begin{pmatrix} 1 & e \\ e & 1 \end{pmatrix} \begin{pmatrix} e \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ e \end{pmatrix} = 1 \begin{pmatrix} e \\ -1 \end{pmatrix},$$

et

$$\begin{pmatrix} 1 & e \\ e & 1 \end{pmatrix} \begin{pmatrix} -1 \\ e \end{pmatrix} = \begin{pmatrix} e \\ 1 \end{pmatrix} = 1 \begin{pmatrix} -1 \\ e \end{pmatrix}.$$

Les deux vecteurs propres ne sont pas "proportionnels".

EXEMPLE 2.20. • L'exemple suivant est un contre exemple trivial à l'unicité dans le cas non fortement connexe

$$\begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 2 \end{pmatrix} \begin{pmatrix} e \\ \varepsilon \end{pmatrix} = 1 \begin{pmatrix} e \\ \varepsilon \end{pmatrix}, \quad \begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 2 \end{pmatrix} \begin{pmatrix} \varepsilon \\ e \end{pmatrix} = 2 \begin{pmatrix} \varepsilon \\ e \end{pmatrix}.$$

- Dans l'exemple suivant le graphe est connexe mais pas fortement connexe néanmoins il n'y a qu'une seule valeur propre. On a

$$\begin{pmatrix} 1 & e \\ \varepsilon & e \end{pmatrix} \begin{pmatrix} e \\ \varepsilon \end{pmatrix} = 1 \begin{pmatrix} e \\ \varepsilon \end{pmatrix},$$

mais

$$\begin{pmatrix} 1 & e \\ \varepsilon & e \end{pmatrix} \begin{pmatrix} a \\ e \end{pmatrix} = \lambda \begin{pmatrix} a \\ e \end{pmatrix},$$

n'a pas de solution parce que la deuxième équation implique que  $\lambda = e$ , et par conséquent la première équation n'a pas de solution pour l'inconnue  $a$ .

- Dans l'exemple suivant le graphe est connexe mais pas fortement connexe mais il y a deux valeurs propres :

$$\begin{pmatrix} e & e \\ \varepsilon & 1 \end{pmatrix} \begin{pmatrix} e \\ \varepsilon \end{pmatrix} = e \begin{pmatrix} e \\ \varepsilon \end{pmatrix}, \quad \begin{pmatrix} e & e \\ \varepsilon & 1 \end{pmatrix} \begin{pmatrix} e \\ 1 \end{pmatrix} = 1 \begin{pmatrix} e \\ 1 \end{pmatrix}.$$

### 3. DÉCISION OPTIMALE

A partir de l'algèbre max-plus ou min-plus on peut construire un formalisme analogue au calcul des probabilités. On préfère utiliser ici l'algèbre min-plus qui est le dual de l'algèbre max-plus (le zéro étant cette fois  $+\infty$ ).

Nous présentons les principaux résultats de cette théorie sans démonstration. Dans beaucoup de cas l'adaptation de la démonstration faite en probabilité se fait facilement.

### 3.1. ESPACE DE DÉCISION, VARIABLES DE DÉCISION

Commençons par définir les mesures de coût qui sont les analogues des mesures de probabilité. L'idée est d'axiomatiser le fait que l'on peut associer à un ensemble de décisions un coût représentant l'infimum (le minimum n'existant pas nécessairement) des coûts des décisions appartenant à cet ensemble .

**DÉFINITION 3.1.** 1. On appelle *espace de décision* le triplet  $\{U, \mathcal{U}, \mathbb{C}\}$  où  $U$  est un espace topologique,  $\mathcal{U}$  est l'ensemble des ouverts de  $U$  et  $\mathbb{C}$  une application de  $\mathcal{U}$  dans  $\overline{\mathbb{R}}^+$  telle que

- (a)  $\mathbb{C}(U) = 0$  ,
- (b)  $\mathbb{C}(\emptyset) = +\infty$  ,
- (c)  $\mathbb{C}(\bigcup_n A_n) = \inf_n \mathbb{C}(A_n)$ ,  $\forall A_n \in \mathcal{U}$  .

2. L'application  $\mathbb{C}$  est appelée *mesure de coût*.

3. Une application  $c$  de  $U$  dans  $\overline{\mathbb{R}}^+$  telle que  $\mathbb{C}(A) = \inf_{u \in A} c(u)$ , pour tout  $A$  dans  $\mathcal{U}$  est appelée *densité de coût* de la mesure  $\mathbb{C}$ .

4. On peut définir aussi *le surcoût conditionnel* de prendre la décision dans  $A$  sachant qu'elle doit nécessairement être prise dans  $B$  par

$$\mathbb{C}(A|B) \stackrel{\text{def}}{=} \mathbb{C}(A \cap B) - \mathbb{C}(B) .$$

**THÉORÈME 3.2.** Soit  $c$  une fonction à valeurs réelles dont l'infimum est 0, l'expression  $\mathbb{C}(A) = \inf_{u \in A} c(u)$  pour tout  $A \in \mathcal{U}$  définit une mesure de coût.

Inversement on peut montrer que toute mesure définie sur les ouverts d'un ensemble polonais (métrisable, séparable et complet) admet une plus petite extension  $\mathbb{C}_*$  à  $\mathcal{P}(U)$  (l'ensemble des parties de  $U$ ). Cette extension est définie de façon unique et admet toujours une densité  $c$  qui est s.c.i. (semi continue inférieurement) et qui satisfait  $\inf_u c(u) = 0$ .

A cause de ce théorème dans la suite on ne s'intéressa qu'aux densités de coût qu'on appelle simplement coût.

On peut construire également l'analogue des variables aléatoires que l'on appelle variable de décision.

**DÉFINITION 3.3.** 1. Une *variable de décision*  $X$  sur  $\{U, \mathcal{U}, \mathbb{C}\}$  est une application de  $U$  dans  $E$  un espace topologique que l'on suppose polonais. Elle induit une mesure de coût  $\mathbb{C}_X$  sur  $(E, \mathcal{O})^1$  par  $\mathbb{C}_X(A) = \mathbb{C}_*(X^{-1}(A))$ , pour tout  $A$  dans  $\mathcal{O}$ . La mesure de coût  $\mathbb{C}_X$  admet une densité de coût notée  $c_X$ .

2. Quand  $E = \mathbb{R}$  [resp.  $\mathbb{R}^n$ , resp.  $\mathbb{R}_{\min}$ ] avec la topologie induite par la valeur absolue [resp. la distance euclidienne, resp.  $d(x, y) =$

<sup>1</sup> $\mathcal{O}$  représente l'ensemble des ouverts de  $E$ .

$|e^{-x} - e^{-y}|$  ] alors  $X$  est appelée *variable de décision réelle* [resp. *vectorielle*, resp. à *valeurs coûts*]

3. Deux variables de décisions  $X$  et  $Y$  sont dites *indépendantes* quand

$$c_{X,Y}(x, y) = c_X(x) + c_Y(y).$$

4. L' *optimum* d'une variable de décision réelle est défini par

$$\mathbb{O}(X) \stackrel{\text{def}}{=} \arg \min_x c_X(x)$$

quand le minimum existe. Quand une variable de décision  $X$  satisfait  $\mathbb{O}(X) = 0$ , on dit qu'elle est *centrée*.

5. On a souvent besoin de l'analogie formel de la moyenne défini par :

$$\mathbb{M}(f(X)) \stackrel{\text{def}}{=} \inf_x (f(x) + c_X(x)) .$$

6. Quand l'optimum d'une variable de décision réelle  $X$  est unique et qu'au voisinage de l'optimum on a

$$c_X(x) = \frac{1}{p} \left| \frac{x - \mathbb{O}(X)}{\sigma} \right|^p + o(|x - \mathbb{O}(X)|^p) ,$$

on définit la *sensibilité d'ordre  $p$*  de  $\mathbb{C}$  par  $\sigma^p(X) \stackrel{\text{def}}{=} \sigma$ . Quand une variable de décision vérifie  $\sigma^p(X) = 1$ , on dit qu'elle est *d'ordre  $p$  et normalisée*.

7. Pour  $1 \leq p < \infty$  les nombres

$$|X|_p \stackrel{\text{def}}{=} \inf \left\{ \sigma \mid c_X(x) \geq \frac{1}{p} |(x - \mathbb{O}(X))/\sigma|^p \right\} ,$$

et

$$\|X\|_p \stackrel{\text{def}}{=} |X|_p + |\mathbb{O}(X)| ,$$

définissent respectivement une semi-norme et une norme sur l'ensemble des variables de décision vérifiant  $\|X\|_p < \infty$ . L'ensemble des variables de décision correspondant est appelé  $\mathbb{D}^p$ . L'espace  $\mathbb{D}^p$  est un espace vectoriel au sens habituel et  $\mathbb{O}$  est un opérateur linéaire sur  $\mathbb{D}^p$ .

8. La densité d'une somme  $Z$  de deux variables de décision  $X$  et  $Y$  est donné par

$$c_Z(z) = \inf_{x,y,z=x+y} c_X(x) + c_Y(y) ,$$

appelé *inf-convolution* de  $c_X$  et de  $c_Y$ , noté  $c_X \star c_Y$ . On a donc  $c_{X+Y} = c_X \star c_Y$ .

## 3.2. TRANSFORMÉE DE FENCHEL

Le rôle de la transformée de Laplace ou de Fourier dans le calcul des probabilités est joué par la transformée de Fenchel dans cette théorie de la décision. Dans la suite l'ensemble des fonctions à valeurs réelles, convexes, s.c.i. et propres (ne valant pas partout  $-\infty$ ) est noté  $\mathcal{C}_x$ .

**DÉFINITION 3.4.** Soit  $c \in \mathcal{C}_x$ , sa transformée Fenchel est la fonction appartenant à  $\mathcal{C}_x$  définie par  $\hat{c}(\theta) = [\mathcal{F}(c)](\theta) \stackrel{\text{def}}{=} \sup_x [\theta x - c(x)]$ .

**EXEMPLE 3.5.** Si  $l_a(x) = ax$  on a  $[\mathcal{F}(l_a)](\theta) = \chi_a(\theta)$  avec

$$\chi_a(\theta) = \begin{cases} +\infty & \text{pour } \theta \neq a, \\ 0 & \text{pour } \theta = a. \end{cases}$$

**THÉORÈME 3.6.** Pour  $f, g \in \mathcal{C}_x$  on a :

1.  $\mathcal{F}(f) \in \mathcal{C}_x$ ,
2.  $\mathcal{F}$  est une involution c.a.d.  $\mathcal{F}(\mathcal{F}(f)) = f$ ,
3.  $\mathcal{F}(f \star g) = \mathcal{F}(f) + \mathcal{F}(g)$ ,
4.  $\mathcal{F}(f+g) = \mathcal{F}(f) \star \mathcal{F}(g)$  sous des hypothèses supplémentaires d'inf-compactité.

**DÉFINITION 3.7.** La fonction caractéristique d'une variable de décision est  $\mathbb{F}(X) \stackrel{\text{def}}{=} \mathcal{F}(c_X)$ .

Elle ne caractérise que les variables de décision ayant un coût appartenant à  $\mathcal{C}_x$ .

**THÉORÈME 3.8.** Si le coût d'une variable décision est d'ordre  $p$  on a :

$$\mathbb{F}(X)'(0) = \mathbb{O}(X),$$

$$\mathbb{F}[X - \mathbb{O}(X)]^{(p')}(0) = \Gamma(p')[\sigma^p(X)]^{p'}, \text{ avec } 1/p + 1/p' = 1,$$

où  $\Gamma$  désigne la fonction spéciale Gamma.

**THÉORÈME 3.9.** Pour deux variables de décision indépendantes  $X$  and  $Y$  d'ordre  $p$  et  $k \in \mathbb{R}$  on a :

$$\mathbb{F}(X + Y) = \mathbb{F}(X) + \mathbb{F}(Y), \quad [\mathbb{F}(kX)](\theta) = [\mathbb{F}(X)](k\theta),$$

$$\mathbb{O}(X + Y) = \mathbb{O}(X) + \mathbb{O}(Y), \quad \mathbb{O}(kX) = k\mathbb{O}(X),$$

$$\sigma^p(kX) = |k|\sigma^p(X), \quad [\sigma^p(X + Y)]^{p'} = [\sigma^p(X)]^{p'} + [\sigma^p(Y)]^{p'},$$

$$(|X + Y|_p)^{p'} \leq (|X|_p)^{p'} + (|Y|_p)^{p'}.$$

## 3.3. COÛTS STABLES

Les fonctions suivantes pour  $p \geq 1$ ,  $m \in \mathbb{R}$ ,  $\sigma \in \mathbb{R}^+$  sont stables par inf-convolution

$$\mathcal{M}_{m,\sigma}^p(x) = \frac{1}{p}(|x - m|/\sigma)^p,$$

plus précisément on a

$$\mathcal{M}_{m,\sigma}^p \star \mathcal{M}_{\bar{m},\bar{\sigma}}^p = \mathcal{M}_{m+\bar{m},[\sigma^{p'}+\bar{\sigma}^{p'}]^{1/p'}}^p \text{ avec } 1/p + 1/p' = 1 .$$

### 3.4. LES THÉOREMES LIMITES POUR LES VARIABLES DE DÉCISION

On étudie le comportement asymptotique de sommes de variables de décision indépendantes lorsque le nombre de termes tend vers l'infini. Pour cela on doit définir les topologies utilisées. Nous ne considérons ici que les deux types de convergence les plus importants.

**DÉFINITION 3.10.** Pour une suite de variable de décisions  $\{X^m, m \in \mathbb{N}\}$  on dit que

1.  $X^m$  converge faiblement vers  $X$ , que l'on note par  $X^m \xrightarrow{w} X$ , si pour tout  $f$  dans  $\mathcal{C}_b(E)$  (où  $\mathcal{C}_b(E)$  désigne l'ensemble des fonctions continues bornées inférieurement de  $E$  dans  $\mathbb{R}_{\min}$ ),

$$\lim_m \mathbb{M}[f(X^m)] = \mathbb{M}[f(X)] .$$

2.  $X^m \in \mathbb{D}^p$  converge en  $p$ -sensibilité vers  $X \in \mathbb{D}^p$ , notée  $X^m \xrightarrow{\mathbb{D}^p} X$ , si  $\lim_m \|X^m - X\|_p = 0$  .

On peut montrer le théorème suivant

**THÉORÈME 3.11.** La convergence en sensibilité entraîne la convergence faible et la réciproque est fausse.

On peut alors énoncer l'analogie de la loi des grands nombres et du théorème de la limite centrale.

**THÉORÈME 3.12** (des grands nombres et de la limite centrale). Etant donnée la suite  $\{X^m, m \in \mathbb{N}\}$  de variables de décision indépendantes et de coûts identiques (i.c.i.) appartenant  $\mathbb{D}^p$ ,  $\infty > p \geq 1$ , on a

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=0}^{N-1} X^m = \mathbb{O}(X^0) ,$$

où la limite est prise au sens de la convergence en  $p$ -sensibilité.

De plus si les  $X^m$  sont centrées

$$w - \lim_N \frac{1}{N^{1/p'}} \sum_{m=0}^{N-1} X^m = X, \text{ avec } 1/p + 1/p' = 1 ,$$

où  $X$  est une variable de décision de coût égal à  $\mathcal{M}_{0,\sigma^p(X^0)}^p$ .

### 3.5. TRANSFORMÉE DE CRAMER

La transformée de Cramer définie par ( $\mathcal{C} \stackrel{\text{def}}{=} \mathcal{F} \circ \log \circ \mathcal{L}$ , où  $\mathcal{L}$  désigne la transformée de Laplace) transforme les mesures de probabilités en fonction convexes et les convolutions en inf-convolutions :

$$\mathcal{C}(f * g) = \mathcal{C}(f) \star \mathcal{C}(g) .$$

TABLEAU 1. Propriétés de la transformée de Cramer.

$\mathcal{M}$	$\mathcal{C}(\mathcal{M})$
$\mu \geq 0$	$c$ convexe l.s.c.
$m_0 \stackrel{\text{def}}{=} \int d\mu = 1$	$\inf_x c(x) = 0$
$m_0 = 1, m \stackrel{\text{def}}{=} \int x d\mu$	$c(m) = 0$
$m_0 = 1, m_2 \stackrel{\text{def}}{=} \int x^2 d\mu$ $\sigma^2 = m_2 - m^2$	$c''(m) = 1/\sigma^2$
$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-m)^2/\sigma^2}$	$\mathcal{M}_{m,\sigma}^2$
distrib. stables	$\mathcal{M}_{m,\sigma}^p$

Elle convertit donc l'addition de variables aléatoires indépendantes en l'addition de variables de décision indépendantes. Elle joue un rôle important en mécanique statistique. Dans la table 1 on donne quelques unes de ses propriétés.

#### 4. CHAÎNES DE BELLMAN

Les chaînes de Bellman correspondent aux chaînes de Markov. L'analogie des équations de Kolmogorov sont les équations de la programmation dynamique beaucoup étudiées par Bellman et qui sont des versions discrètes des équations d'Hamilton-Jacobi.

**DÉFINITION 4.1.** Une chaîne de Bellman homogène en temps, prenant un nombre fini de valeurs, est définie à partir du triplet  $(\mathcal{E}, C, \phi)$  où

1.  $\mathcal{E}$  un ensemble fini appelé espace d'état possédant  $E$  éléments,
2.  $C : \mathcal{E} \times \mathcal{E} \mapsto \overline{\mathbb{R}}$  satisfaisant  $\inf_y C_{xy} = 0$  appelé coût de transition,
3.  $\phi$  une mesure de coût sur  $\mathcal{E}$  appelé coût initial,

comme étant la suite de variables de décision  $\{X_n\}$  (sur  $\{U, \mathcal{U}, \mathbb{C}\}$  à valeurs dans  $\mathcal{E}$ ), telle que

$$c_X(x_0, x_1, \dots) = \phi_{x_0} + \sum_{i=0}^{\infty} C_{x_i x_{i+1}}, \quad \forall x = (x_0, x_1, \dots) \in \mathcal{E}^{\mathbb{N}}.$$

**THÉORÈME 4.2.** Une chaîne de Bellman vérifie la propriété de Markov suivante

$$\mathbb{M}\{f(X_n) \mid X_0, \dots, X_{n-1}\} = \mathbb{M}\{f(X_n) \mid X_{n-1}\},$$

pour toute fonction  $f$  de  $\mathcal{E}$  dans  $\overline{\mathbb{R}}$ .

L'analogie de l'équation de Kolmogorov en avant permet de calculer récursivement le coût marginal d'être dans un état à un instant donné.

THÉORÈME 4.3. Le coût marginal  $w_x^n = \mathbb{C}(X_n = x)$  de la chaîne de Bellman est donné par

$$w^{n+1} = w^n \otimes C \stackrel{\text{def}}{=} \min_{x \in \mathcal{E}} (w_x^n + C_x), \quad w^0 = \phi .$$

Le coût d'une chaîne de Bellman est normalisé ce qui signifie que son infimum sur l'ensemble des trajectoires est 0. Dans certaines applications on aimerait enlever cette restriction. On peut le faire en introduisant l'analogie des fonctionnelles multiplicatives des trajectoires d'un processus stochastique.

THÉORÈME 4.4. La valeur

$$v_x^n \stackrel{\text{def}}{=} \mathbb{M} \left\{ \sum_{k=n}^{N-1} f(X_k) + \Psi(X_N) \mid X_n = x \right\}$$

avec  $f, \Psi \in \overline{\mathbb{R}}^E$  peut être calculée récursivement par

$$v^n = F \otimes C \otimes v^{n+1} = f. + \min_y (C_{.y} + v_y^{n+1}), \quad v^N = \Psi ,$$

où  $F = \text{diag } f$  est définie par  $F_{xy} = f_x$  si  $x = y$  et  $+\infty$  sinon.

## CHAPITRE 3

# COMMANDE OPTIMALE STOCHASTIQUE

### 1. INTRODUCTION ET EXEMPLES

Etant donnée une chaîne de Markov de matrice de transition dépendant d'un paramètre appelé commande, une observation ne dépendant que de l'état de la chaîne de manière markovienne, un critère qui est une fonctionnelle de la trajectoire de cette chaîne, le problème consiste à rechercher la commande ne dépendant que des observations passées réalisant le minimum du critère. Pour illustrer ce type de problèmes prenons quelques exemples.

#### 1.1. GESTION D'UN STOCK D'EAU

Etant donné le barrage décrit au chapitre précédent, on considère la politique de gestion suivante :

$$Z^n = \begin{cases} 1 & \text{si } X^n > a , \\ 0 & \text{si } X^n \leq a , \end{cases}$$

où  $a$  est un paramètre ou une fonction du temps à optimiser. La matrice de transition de la chaîne de Markov a alors la même structure que précédemment. Si l'on désigne par  $c_{X^n}^{Z^n}$  l'énergie produite changée de signe, lorsque le stock est dans l'état  $X^n$  et que l'on a turbiné  $Z^n$ . Le critère à optimiser s'écrit :

$$\min_a \mathbb{E} \left\{ \sum_{n=0}^{N-1} c_{X^n}^{Z^n} \mid X^0 = x \right\} .$$

Dans cet exemple :

- l'état est la variable aléatoire  $X^n$ ,
- l'observation est égale à  $X^n$ ,
- la commande est le paramètre  $a$ .

#### 1.2. PROBLÈME DE MAINTENANCE

On considère un matériel susceptible de tomber en panne : L'état de fonctionnement est noté  $M$ , se panne est noté  $\overline{M}$ .

Ce matériel est caractérisé par son âge  $x$  — temps passé depuis son dernier remplacement. On supposera que tout matériel d'âge  $N$  est remplacé. A chaque instant un matériel d'âge  $x$  a une probabilité  $\lambda_x$  de tomber en panne croissante avec son âge. A chaque instant on a la possibilité de remplacer ou de conserver ce matériel.

L'état du système sera la variable aléatoire pouvant prendre une des valeurs  $\overline{M}, 1, 2, \dots, N$  indiquant que le matériel est en panne ( $\overline{M}$ ), ou donnant son âge pouvant prendre les valeurs  $\{1, 2, \dots, N\}$ .

Pour ce système on peut supposer observer l'état<sup>1</sup>.

La commande  $u$  est la décision de remplacement :  $u = 1$  indique le remplacement,  $u = 0$  le non remplacement.

La matrice de transition de la chaîne de Markov décrivant l'évolution de l'état s'écrit alors :

$$M^u = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot \\ p_1^u & q_1^u & r_1^u & 0 & \cdot \\ p_2^u & q_2^u & 0 & r_2^u & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot \end{pmatrix},$$

où  $p_x^0 = \lambda_x$ ,  $p_x^1 = 0$ ,  $q_x^0 = 0$ ,  $q_x^1 = 1$ ,  $r_x^u = 1 - p_x^u - q_x^u$ .

Et donc la matrice de transition dépend du paramètre de commande  $u$ .

On associe un coût au fonctionnement du système constitué de la somme du coût de remplacement du matériel et d'un coût de défaillance ayant lieu lorsque le système est à l'arrêt:

$$c_x^u = c^u + d_x,$$

avec :

$$c^u = \begin{cases} k_1 & \text{sur } u = 1, \\ 0 & \text{ailleurs,} \end{cases}$$

$$d_x = \begin{cases} 0 & \text{si } x = 1, 2, \dots, N - 1, \\ k_2 & \text{si } x = \overline{M}, \\ k_1 & \text{si } x = N. \end{cases}$$

Le problème de commande consiste alors à minimiser, par exemple, le coût actualisé du fonctionnement du système c.a.d. :

$$\min_u \mathbb{E} \sum_{n=0}^{+\infty} \frac{1}{(1 + \lambda)^{n+1}} (c^{U^n} + d_{X^n}),$$

$\lambda$  étant un nombre positif représentant un taux d'actualisation.

## 2. FORMULATION PRÉCISE DU PROBLÈME

On se donne la chaîne de Markov  $(\mathcal{T}, \mathcal{E}, \mathcal{F}, \mathcal{G}, M^{uy}, p^0, c^{uy})$  commandée et observée avec :

- un temps  $n \in \mathcal{T} = \mathbb{N}$ ;
- un espace d'état  $\mathcal{E} = \{1, 2, \dots, E\}$ ;
- un espace de sorties  $\mathcal{G} = \{1, 2, \dots, G\}$ ;
- un espace d'entrée  $\mathcal{F} = \{1, 2, \dots, F\}$ ;

<sup>1</sup>Ce n'est pas toujours le cas, par exemple, pour un matériel de sécurité qui en général ne marche pas sauf s'il y a un incident, on est incapable, en dehors des instants de test, de savoir avec certitude s'il est en état de fonctionnement.

- une famille de matrices de probabilité de transition de  $\mathcal{E} \rightarrow \mathcal{E}$  notée  $\{M^{uy}, u \in \mathcal{F}, y \in \mathcal{G}\}$  qui s'interprète comme la probabilité d'aller d'un état à un autre et d'observer  $y$  si l'on prend la décision  $u$  ;
- $p^0$  une loi de probabilité sur  $\mathcal{E} \times \mathcal{G}$  appelée loi initiale ;
- une famille de coûts instantanés  $c^{uy} : \mathcal{E} \rightarrow \mathbb{R}^+, u \in \mathcal{F}, y \in \mathcal{G}$  .

On appelle :

- $X^n \in \mathcal{E}$  l'état du système à l'instant  $n$  ;
- $Y^n \in \mathcal{G}$  l'observation du système à l'instant  $n$  ;
- $U^n \in \mathcal{F}$  la commande à l'instant  $n$  .

On note  $(X^n \in \mathcal{E}, U^n \in \mathcal{F}, Y^n \in \mathcal{G}, n \in \mathcal{T})$  le processus canonique correspondant où la seule contrainte imposée à  $(U^n)$  est de ne dépendre que du passé des observations. On a alors la mesure de probabilité  $\mathbb{P}$  du processus définie par :

$$\mathbb{P}((x, y, u)_0, (x, y, u)_1, (x, y, u)_2, \dots) =$$

$$p_{x^0}^{0y^0} \rho^{0y^0u^0} m_{x^0x^1}^{u^0y^1} \rho^{1y^0y^1u^1} m_{x^1x^2}^{u^1y^2} \rho^{2y^0y^1y^2u^2} m_{x^2x^3}^{u^2y^3} \dots$$

où les  $\rho^i$  définissent la loi de probabilité de la commande qu'il faut optimiser.

On appelle  $\mathcal{R}$  l'ensemble des stratégies c.a.d. des suites  $(\rho^n)_{n \in \mathcal{T}}$ . Si chaque mesure  $\rho^n$  est concentrée en un point de  $\mathcal{F}$  on dira que la stratégie est déterministe — l'ensemble de telles stratégies sera noté  $\mathcal{S}$  — dans le cas contraire la stratégie est qualifiée de relaxée.

Le problème de commande optimale consiste alors à minimiser le coût de fonctionnement du système sur une période finie ou infinie — le minimum étant à prendre dans la classe  $\mathcal{R}$  ou dans une sous classe de  $\mathcal{R}$  — c.a.d. résoudre des problèmes du type :

$$\min_{\rho \in \mathcal{R}} \mathbb{E} \sum_{n=0}^{N-1} c_{X^n}^{U^n Y^n},$$

où  $N$ , fini ou infini, est appelé l'horizon du problème.

REMARQUE 2.1. La recherche de l'optimum dans  $\mathcal{R}$  est l'objectif principal de la commande optimale. Cependant cet optimum est souvent difficile voir pratiquement impossible à atteindre. On est alors conduit à étudier des cas particuliers plus faciles, ou à restreindre la classe des stratégies dans laquelle on optimise — on parlera alors de commandes sous-optimales.

Dans la suite nous étudierons surtout le cas particulier important appelé problème en observation complète, cas où  $X^n = Y^n$ . On montrera alors que l'on ne perd rien à optimiser dans la classe des boucles fermées sur l'état — feedback en anglais ou parfois politique markovienne — c'est à dire dans l'ensemble des fonctions déterministes de l'espace d'état dans l'espace de commande :

$$\mathcal{S}_F = \{s = (s^n : \mathcal{E} \rightarrow \mathcal{F})_{n \in \mathcal{T}}\}.$$

Dans de nombreux problèmes pratiques même ce cas particulier est “impossible” à résoudre. En effet  $E$  est souvent du type  $m^k$  —  $k$  est la dimension du système — et donc peut atteindre des tailles astronomiques. La seule mémorisation de la commande optimale est alors hors d’atteinte. On est conduit, pour éviter cette difficulté, à optimiser dans une classe plus restreinte, par exemple la classe des boucles ouvertes  $\mathcal{S}_O$  — open loop. C’est le cas où l’observation est vide  $\mathcal{G} = \emptyset$ . La commande est alors une fonction du temps seul :

$$\mathcal{S}_O = \{u = (u^n \in \mathcal{F})_{n \in \mathcal{T}}\} .$$

On parle parfois de feedback a priori lorsque l’on s’est ramené à la situation de la boucle ouverte en faisant un changement de variable sur la commande. La classe de stratégies est alors de la forme :

$$\mathcal{S}_{FP} = \{a = (a^n \in A, u^n = s_x^{a^n}, s : A \times \mathcal{E} \rightarrow \mathcal{F} \text{ donné})_{n \in \mathcal{T}}\} .$$

Lorsqu’on connaît la forme du feedback optimal, cette façon de poser le problème peut conduire à de grosses économies en temps de calcul à condition, bien sûr, de disposer d’une méthode efficace de résolution du problème de commande en boucle ouverte correspondant.

Dans la suite de ce chapitre nous étudierons le problème en observation complète dans trois situations :

- problème en horizon fini :

$$\min_{s \in \mathcal{S}} \mathbb{E} \left\{ \sum_{n=0}^{N-1} c_{X^n}^{nU^n} + \phi_{X^N} \right\} ,$$

où  $\phi$  est un coût particulier appelé coût final;

- problème en horizon infini avec un coût actualisé :

$$\min_{s \in \mathcal{S}} \mathbb{E} \sum_{n=0}^{+\infty} \frac{1}{(1 + \lambda)^{n+1}} c_{X^n}^{U^n} ;$$

- problème en horizon infini avec un coût non actualisé :

$$\min_{s \in \mathcal{S}} \lim_{\lambda \rightarrow 0} \mathbb{E} \sum_{n=0}^{+\infty} \frac{\lambda}{(1 + \lambda)^{n+1}} c_{X^n}^{U^n} ;$$

problème qui a même solution que :

$$\min_{s \in \mathcal{S}} \lim_{N \rightarrow \infty} \mathbb{E} \frac{1}{N} \sum_{n=0}^{N-1} c_{X^n}^{U^n} .$$

### 3. PROGRAMMATION DYNAMIQUE EN HORIZON FINI

Nous nous intéressons dans ce paragraphe à l’optimisation, d’un système régi par une chaîne de Markov, sur une période de gestion finie de  $N$  étapes. Nous élargirons les hypothèses du paragraphe précédent au cas  $\mathcal{F}$  compact. Nous nous plaçons dans le cas d’observation complète — dans ce cas puisque l’observation est égale à l’état  $M_{xx'}^{ux'}$  est notée  $M_{xx'}^u$ . On

accepte dans ce paragraphe que la matrice de transition et le coût dépendent du temps — nous les notons alors respectivement  $M^{nu}$  et  $c^{nu}$ ,  $n \in \mathcal{T}$ . Le théorème suivant permet de calculer par récurrence arrière la stratégie optimale. L'équation correspondante est appelée équation de la programmation dynamique ou de Bellman ou d'Hamilton-Jacobi-Bellman dans le contexte de la commande d'équations différentielles. Elle joue le rôle de l'équation de Kolmogorov pour les problèmes de commande stochastique.

**THÉORÈME 3.1.** Sous les hypothèses :

- $\mathcal{F}$  compact,
- $u \rightarrow (M^{nu}, c^{nu})$  continue pour tout  $n \in \mathcal{T}$ ,

la solution de :

$$\begin{cases} v_x^n = \min_{u \in \mathcal{F}} \{ (M^{nu} v^{n+1})_x + c_x^{nu} \} \\ v_x^n = \phi_x \end{cases} \quad \forall x \in \mathcal{E}, \quad (3.1)$$

donne le coût optimal :

$$v_x^n = \min_{S \in \mathcal{S}} \mathbb{E} \left\{ \sum_{j=n}^{N-1} c_{X^j}^{jU^j} + \phi_{X^N} | X^n = x \right\},$$

et les applications :

$$s^n : x \rightarrow u^* \in \arg \min_{u \in \mathcal{F}} \{ (M^{nu} v^{n+1})_x + c_x^{nu} \} \quad (3.2)$$

définissent une stratégie en boucle fermée optimale dans la classe  $\mathcal{S}$ .

**PREUVE.** Par récurrence rétrograde il est clair que (3.1) admet une solution, montrons qu'elle est optimale. Soit  $(v^n)$  solution de (3.1), soit  $\nu = (v^n)$  une stratégie dans  $\mathcal{S}$  on a :

$$\mathbb{E}^{n\nu} (v_{X^{n+1}}^{n+1} - v_{X^n}^n) = (M^{n\nu} v^{n+1} - v^n)_{X^n},$$

où  $\mathbb{E}^{n\nu}$  désigne l'espérance conditionnelle connaissant les états réalisés jusqu'à l'instant  $n$ , pour la mesure de probabilité construite à partir de la stratégie  $\nu$ . En sommant les égalités précédentes et en prenant l'espérance conditionnelle par rapport à la connaissance de l'état à l'instant 0, on obtient :

$$\begin{aligned} \mathbb{E}^{0\nu} v_{X^N}^N - v_{x_0}^0 &= \mathbb{E}^{0\nu} \sum_{n=0}^{N-1} (v_{X^{n+1}}^{n+1} - v_{X^n}^n), \\ &= \mathbb{E}^{0\nu} \sum_{n=0}^{N-1} (M^{n\nu} v^{n+1} - v^n)_{X^n}, \\ &\geq -\mathbb{E}^{0\nu} \sum_{n=0}^{N-1} c_{X^n}^{n\nu} \quad \text{grâce à (3.1),} \end{aligned}$$

et donc

$$v_{x_0}^0 \leq \mathbb{E}^{0\nu} \left\{ \sum_{n=0}^{N-1} c_{X^n}^{n\nu} + \phi_{X^N} \right\}, \quad \forall \nu \in \mathcal{S}.$$

De plus, par un raisonnement analogue, en faisant jouer le rôle de  $v$  à la stratégie (3.2), on montre l'égalité pour la stratégie définie par (3.2).  $\square$

REMARQUE 3.2. On a montré que dans le cas où on observe l'état il existe une stratégie optimale pour la classe de stratégies  $\mathcal{S}$  ne dépendant que de l'état présent donc appartenant à  $\mathcal{S}_F$ . Si l'on avait formulé le problème initial d'optimisation dans  $\mathcal{R}$  le feedback défini par (3.2) serait encore optimal.

#### 4. PROGRAMMATION DYNAMIQUE COÛT ACTUALISÉ

On s'intéresse dans ce paragraphe au problème de la commande optimale d'une chaîne de Markov en horizon infini, pour un coût actualisé, dans le cas où l'observation est complète. 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. Des méthodes itératives permettent de calculer sa solution avec un coût de calcul bien moindre que celui nécessaire pour résoudre le problème en l'horizon fini correspondant.

Nous donnerons deux méthodes de résolution de l'équation de la programmation dynamique : l'itération sur les valeurs, l'itération sur les politiques. Pour cela nous commençons par caractériser la politique optimale comme solution d'équation de la programmation dynamique.

THÉORÈME 4.1. Sous les hypothèses :

- $\mathcal{F}$  compact,
- $u \rightarrow (M^u, c^u)$  continue,
- $c_x^u \geq 0$ ,  $\forall u \in \mathcal{F}$ ,  $\forall x \in \mathcal{E}$ ,
- $\lambda > 0$ ,

en notant  $A^u = M^u - I$ ,  $v$  solution unique de :

$$\min_{u \in \mathcal{F}} \{[(A^u - \lambda)v]_x + c_x^u\} = 0, \quad \forall x \in \mathcal{E}, \quad (4.1)$$

est le coût optimal

$$v_x = \min_{s \in \mathcal{S}} \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1 + \lambda)^{n+1}} c_{X^n}^{U^n} \mid X^0 = x \right\}.$$

De plus

$$s : x \rightarrow u^* \in \arg \min_{u \in \mathcal{F}} \{(A^u v)_x + c_x^u\} \quad (4.2)$$

définit une stratégie optimale dans la classe  $\mathcal{S}$ . Cette stratégie est markovienne c.a.d.  $\in \mathcal{S}_F$ .

PREUVE. EXISTENCE D'UNE SOLUTION DE (4.1). On construit la solution au moyen de l'une des deux itérations suivantes :

ITÉRATION SUR LES VALEURS

$$v_x^{n+1} = \frac{1}{1 + \lambda} \min_{u \in \mathcal{F}} \{(M^u v^n)_x + c_x^u\} \quad (4.3)$$

cette itération en  $n$  est contractante car les  $M^u$  sont des matrices stochastiques donc de norme inférieure ou égale à 1.

#### ITÉRATION SUR LES POLITIQUES OU ALGORITHME DE HOWARD

- ÉTAPE  $2n - 1$ . A une stratégie  $s^n$  on associe  $v^n$  solution de:

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

Cette solution existe et est unique puisque  $A^{s^n} - \lambda$  est inversible car  $A$  est le générateur d'une chaîne de Markov et  $\lambda > 0$  et donc toutes les valeurs propres de  $A^{s^n} - \lambda$  sont strictement négatives.

- ÉTAPE  $2n$ . A un coût  $v^n$  on associe une nouvelle stratégie  $s^{n+1} : x \in \mathcal{E} \rightarrow u^{n+1}$  où:

$$u^{n+1} \in \arg \min_{u \in \mathcal{F}} \{[(A^u - \lambda)v^n]_x + c_x^u\} ,$$

qui a bien un sens puisque  $\mathcal{F}$  est compact et  $u \rightarrow (A^u, c^u)$  est continue.

On s'intéresse seulement à l'itération sur les politiques. On a construit par cette méthode d'itération sur les politiques deux suites  $(v^n)_{n \in \mathbb{N}}$ ,  $(s^n)_{n \in \mathbb{N}}$ . On va montrer que  $v^n$  est une suite décroissante positive qui admet une limite solution de l'équation de la programmation dynamique. Dans la suite on notera  $A^n = A^{s^n} - \lambda$ ,  $c^n = c^{s^n}$ .

- LA SUITE  $v^n$  EST  $\geq 0$ . On le voit grâce à l'interprétation stochastique de  $v^n$ . Montrons le également d'un point de vue analytique. Supposons qu'il n'en soit pas ainsi. Alors  $\min_x v_x^n < 0$ . Pour un  $x$  réalisant le minimum on aurait à cause du principe du minimum strictement négatif  $(A^n v^n)_x > 0$  et comme  $c_x^n \geq 0$  on a une contradiction.
- LA SUITE  $v^n$  EST DÉCROISSANTE. En effet on a :

$$A^n v^n + c^n = 0 .$$

Par différence entre deux équations successives on obtient :

$$A^n v^n - A^{n+1} v^{n+1} + c^n - c^{n+1} = 0 ,$$

$$A^{n+1}(v^n - v^{n+1}) + (A^n - A^{n+1})v^n + c^n - c^{n+1} = 0 ,$$

et donc grâce aux étapes (2n) de l'algorithme, on a :

$$A^{n+1}(v^n - v^{n+1}) \leq 0 .$$

Enfin par le même raisonnement que celui qui nous a permis de montrer la positivité de  $v^n$ , on a :

$$v^n - v^{n+1} \geq 0 \Rightarrow v^n \geq v^{n+1} .$$

- LA LIMITE DE  $v^n$  EST SOLUTION. La suite  $v^n$  est décroissante minorée, elle est donc convergente, notons  $v^*$  sa limite. L'ensemble  $\mathcal{F}$  étant compact l'ensemble des stratégies est compact comme produit

d'un nombre fini de compacts. Il existe donc une sous suite convergente ( $s^{n'}$ ). Notons  $s^*$  sa limite. Les hypothèses de continuité entraîne que  $A^{n'} \rightarrow A^*$  et  $c^{n'} \rightarrow c^*$  avec  $A^* = A^{s^*} - \lambda$  et  $c^* = c^{s^*}$ . Puisque :

$$(A^{n'} v^{n'-1} + c^{n'})_x \leq ((A^u - \lambda)v^{n'-1} + c^u)_x, \quad \forall u \in \mathcal{F}, \quad \forall x \in \mathcal{E},$$

et

$$A^{n'} v^{n'} + c^{n'} = 0,$$

on a

$$0 = A^* v^* + c^* \leq (A^s - \lambda)v^* + c^s, \quad \forall s \in \mathcal{S}_F.$$

Ce qui démontre l'existence d'une solution à l'équation de la programmation dynamique.

UNICITÉ DE LA SOLUTION DE (4.1). S'il y avait deux solutions  $v^1$  et  $v^2$  avec  $s^1$  et  $s^2$  deux stratégies optimales correspondantes, on aurait  $A^1 v^1 + c^1 \leq A^2 v^1 + c^2$ . Puisque  $A^1 v^1 + c^1 - A^2 v^2 - c^2 = 0$  se réécrit  $A^2(v^1 - v^2) + (A^1 - A^2)v^1 + c^1 - c^2 = 0$  on a  $A^2(v^1 - v^2) \geq 0$  en effet  $(A^1 - A^2)v^1 + c^1 - c^2 \leq 0$ . On en déduit que  $v^1 - v^2 \leq 0$  par le principe du maximum discret vérifié par  $A^2$  ou par l'interprétation probabiliste de la solution d'une équation  $A^2 v + d = 0, d \leq 0$ . Par symétrie du raisonnement on montre  $v^2 - v^1 \leq 0$ . Et donc on a  $v^1 = v^2$ .

INTERPRÉTATION STOCHASTIQUE. Il nous reste à monter que  $v$  solution de (4.1) est le coût du problème de commande optimale stochastique. Soit  $\nu$  une stratégie qcq en notant  $\rho = 1 + \lambda$  on a :

$$\begin{aligned} \mathbb{E}^{n\nu} \left\{ \frac{1}{\rho^{n+1}} v_{X^{n+1}} - \frac{1}{\rho^n} v_{X^n} \right\} &= \frac{1}{\rho^{n+1}} (M^{\nu^n} v - \rho v)_{X^n}, \\ &= \frac{1}{\rho^{n+1}} ((A^{\nu^n} - \lambda)v)_{X^n}, \\ &\geq -\frac{1}{\rho^{n+1}} c_{X^n}^{\nu^n}, \end{aligned}$$

et donc

$$\mathbb{E}^{0\nu} \sum_{n=0}^{+\infty} \left\{ \frac{1}{\rho^{n+1}} v_{X^{n+1}} - \frac{1}{\rho^n} v_{X^n} \right\} = -v_{X^0}, \geq -\mathbb{E}^{0\nu} \sum_{n=0}^{\infty} \frac{1}{\rho^{n+1}} c_{X^n}^{\nu^n},$$

et donc

$$v_{X^0} \leq \mathbb{E}^{0\nu} \sum_{n=0}^{+\infty} \frac{1}{\rho^{n+1}} c_{X^n}^{\nu^n}, \quad \forall \nu \in \mathcal{S},$$

et comme on a l'égalité pour la stratégie définie par (4.1), on a montré le théorème.  $\square$

## 5. PROGRAMMATION DYNAMIQUE ERGODIQUE

On étudie le problème de commande dans le cas ergodique sous l'hypothèse d'irréductibilité de la chaîne de Markov quelque soit la stratégie adoptée dans  $\mathcal{S}_F$ . On montre l'existence d'une solution à équation de la programmation en adaptant l'itération sur les politiques à cette situation.

**THÉORÈME 5.1.** Sous les hypothèses:

- $\mathcal{F}$  compact,
- $u \rightarrow (M^u, c^u)$  continue,
- la chaîne de Markov est irréductible pour tout  $s \in \mathcal{S}_F$ ,
- $c \geq 0$ ,

il existe une solution  $(v, w)$  à :

$$\min_u \{(A^u v)_x + c_x^u\} - w = 0, \forall x \in \mathcal{E}. \quad (5.1)$$

De plus  $w \in \mathbb{R}$  est unique et  $v : \mathcal{E} \rightarrow \mathbb{R}$  est définie à une constante près. La constante  $w$  s'interprète comme:

$$w = \min_{S \in \mathcal{S}_F} \lim_{\lambda \rightarrow 0} \left\{ \sum_{n=0}^{+\infty} \frac{\lambda}{(1 + \lambda)^{n+1}} c_{X^n}^{U^n} \right\}. \quad (5.2)$$

**PREUVE. EXISTENCE D'UNE SOLUTION DE (5.1).** L'idée d'itération sur les valeurs peut être adaptée à cette situation. Nous ne le ferons pas ici. Nous discutons uniquement de l'itération sur les politiques. On construit donc les suites  $(s^n)_{n \in \mathbb{N}}$ ,  $(v^n)_{n \in \mathbb{N}}$ ,  $(w^n)_{n \in \mathbb{N}}$  de la façon suivante :

- ÉTAPE  $2n - 1$  :  $s^n \rightarrow (v^n, w^n)$  solution de  $A^{s^n} v^n + c^{s^n} - w^n = 0$  qui existe et qui est unique (si l'on impose  $v^n \in \mathcal{R}(A^{s^n})$ ) grâce aux résultats de l'étude des chaînes de Markov dans le cas ergodique;
- ÉTAPE  $2n$  :

$$(v^n, w^n) \rightarrow s^{n+1} : x \rightarrow u^{n+1} \in \arg \min (A^u v + c^u)_x.$$

Étudions maintenant cette suite en utilisant des notations analogues à celles du paragraphe précédent.

- LES  $w^n$  SONT POSITIFS. S'il en était autrement on aurait  $w^n < 0$  qui implique  $c - w^n > 0$  et donc  $A^n v^n < 0$ . Or on a

$$x \in \arg \min_{x \in \mathcal{E}} v_x^n \Rightarrow (A^n v^n)_x \geq 0$$

grâce au principe du minimum ce qui induit une contradiction.

- LES CONSTANTES  $w^n$  SONT DÉCROISSANTES. En effet on a :

$$A^n v^n + c^n - w^n = A^{n+1} v^{n+1} + c^{n+1} - w^{n+1} = 0,$$

et donc

$$\begin{aligned} & (A^n - A^{n+1})v^n + c^n - c^{n+1} \\ & + A^{n+1}(v^n - v^{n+1}) - w^n + w^{n+1} = 0, \end{aligned}$$

et donc

$$A^{n+1}(v^n - v^{n+1}) - w^n + w^{n+1} \leq 0 ,$$

et donc le même raisonnement que celui prouvant la positivité de  $w^n$  montre que :

$$w^n - w^{n+1} \geq 0 \Rightarrow w^n \geq w^{n+1} .$$

- **PASSAGE À LA LIMITE SUR L'INDICE D'ITÉRATION  $n$ .** On a  $u^n \in \mathcal{F}$  compact, il existe donc une sous suite  $(s^{n'})$  convergente. Notons  $s^*$  sa limite. Le lieu du spectre de  $A^s$ ,  $s \in \mathcal{S}_F$  est compact puisque  $u \rightarrow \lambda_{A^s}$ , où  $\lambda$  désigne une valeur propre, est continue et que l'ensemble des stratégies est compact. Et donc, puisque  $\forall s \in \mathcal{S}_F$  les chaînes de Markov sont irréductibles, il existe un voisinage de 0 dans le plan complexe ne contenant aucune autre valeur propre de  $A^s$  que 0,  $\forall s \in \mathcal{S}_F$ . Donc en utilisant la semi-simplicité de la valeur propre 0, il existe  $k$  réel positif :

$$\|A^n|_{\mathcal{R}(A^n)}^{-1}\| \leq k$$

où  $\mathcal{R}(A^n)$  désigne l'image de  $A^n$ . La suite  $(v^n)$  reste donc bornée.

Il existe donc une sous-suite notée  $n''$  telle que:

$$\begin{aligned} A^{n''+1} &\rightarrow A^* , \\ v^{n''} &\rightarrow v^* , \\ v^{n''+1} &\rightarrow v^{**} , \\ w^{n''+1} &\rightarrow w^* , \\ w^{n''} &\rightarrow w^* , \\ c^{n''+1} &\rightarrow c^* . \end{aligned}$$

On a alors

$$A^{n''+1}v^{n''+1} + c^{n''+1} - w^{n''+1} = 0 ,$$

$$A^{n''+1}v^{n''} + c^{n''+1} \leq A^s v^{n''} + c^s , \forall s \in \mathcal{S}_F .$$

En passant à la limite sur  $n''$  on obtient :

$$A^* v^{**} + c^* - w^* = 0 ,$$

$$A^* v^* + c^* \leq A^s v^* + c^s , \forall s \in \mathcal{S}_F ,$$

$$A^* v^* + c^* - w^* \leq 0 ,$$

et donc en notant  $\tilde{v} = v^* - v^{**}$ :

$$A^* \tilde{v} \leq 0 .$$

Montrons que  $A^* \tilde{v} = 0$ . Si  $x$  désigne un indice  $\in \mathcal{E}$  où  $\tilde{v}$  atteint son minimum  $(A^* \tilde{v})_x \geq 0$  et donc

$$(A^* \tilde{v})_x = 0 , \tag{5.3}$$

ce qui n'est possible que si  $\tilde{v}$  est constante car (5.3) signifie que  $v_x$  est une moyenne pondérée de ses points voisins, et comme  $x$  est le minimum, les points voisins ont la même valeur et de proche en proche en utilisant l'irréductibilité de la chaîne de Markov on obtient :

$$A^* \tilde{v} = 0 .$$

Et donc on a

$$\min_u \{(A^u v^*)_x + c_x^u\} = w^* ,$$

d'où l'existence.

UNICITÉ DE  $w$ . Supposons qu'il existe deux solutions  $w^1$  et  $w^2$  on en déduit, en réutilisant les raisonnements précédents, que

$$A^2(v^1 - v^2) - w^1 + w^2 \geq 0 \Rightarrow w^1 - w^2 \leq 0 ,$$

et par symétrie  $w^2 - w^1 \leq 0$  d'où l'unicité.

INTERPRÉTATION STOCHASTIQUE DE  $w$ . Soit  $\nu$  une commande quelconque  $\in \mathcal{S}_F$ ,  $p^\infty$  la mesure invariante associée, on a :

$$p^\infty M^\nu = p^\infty , \text{ et donc } p^\infty A^\nu = 0 .$$

Si  $v$  et  $w$  désigne la solution de (5.1) on a donc

$$(A^\nu v, p^\infty) = 0 .$$

Mais  $p^\infty \geq 0$  et donc grâce à (5.1) on a

$$0 = (A^\nu v, p^\infty) \geq (w - c^\nu, p^\infty) ,$$

et donc  $w \leq (c^\nu, p^\infty)$ .

On montre par un raisonnement analogue que l'égalité est atteinte pour une stratégie vérifiant:

$$s \in \arg \min \{A^u v + c^u\} .$$

Les résultats du paragraphe sur la théorie ergodique des chaînes de Markov donnent alors l'interprétation souhaitée.  $\square$

## 6. COMMANDE EN INFORMATION INCOMPLÈTE

Nous étudions ici le problème de la commande optimale de chaîne de Markov dans le cas où on n'observe pas directement l'état mais seulement une fonction statique de l'état à valeurs dans un ensemble fini. Nous commençons par donner l'équation de filtre optimal c.a.d. l'équation régissant l'évolution de la loi conditionnelle de l'état connaissant le passé des observations. Nous donnons ensuite l'équation de la programmation dynamique du problème en information incomplète. L'état à mémoriser pour résoudre le problème d'optimisation n'est plus l'état du système mais la loi conditionnelle de l'état connaissant le passé des observations.

## 6.1. FILTRAGE OPTIMAL

Nous recherchons le meilleur estimé de  $X^n$ , état d'une chaîne de Markov non commandé, connaissant le passé des observations  $Y^0, Y^1, \dots, Y^n$ . Pour cela, on se donne une chaîne de Markov observée, c.a.d. le 5-uple :

$$(\mathcal{T}, \mathcal{E}, \mathcal{G}, M^y, p^0),$$

où

- $\mathcal{T}, \mathcal{E}, \mathcal{G}$  ont la même signification que précédemment;
- $p^0$  est une loi initiale sur  $\mathcal{E} \times \mathcal{G}$ ;
- $M_{xx'}^y$  représente la probabilité de transiter de l'état  $x$  à l'état  $x'$  et d'observer  $y$ .

La probabilité d'une trajectoire est alors donnée par :

$$\mathbb{P}((x^0, y^0), (x^1, y^1), \dots, (x^n, y^n)) = p_{x^0}^{0y^0} \prod_{i=0}^{n-1} M_{x^i x^{i+1}}^{y^{i+1}}. \quad (6.1)$$

On calcule alors la loi conditionnelle :

$$q_x^n = \mathbb{P}(X^n = x | Y^0 = y^0, \dots, Y^n = y^n).$$

**THÉORÈME 6.1.** La loi conditionnelle de l'état de la chaîne de Markov vaut :

$$q_x^n = \frac{p^{0y^0} \prod_{i=1}^n M^{y^i}}{p^{0y^0} \prod_{i=1}^n M^{y^i} \mathbf{1}}. \quad (6.2)$$

PREUVE. On calcule la loi marginale de  $(y^0, \dots, y^n, x^n)$  :

$$\sum_{\substack{x^i \\ i=0, \dots, n-1}} \left[ p_{x^0}^{0y^0} \prod_{i=0}^{n-1} M_{x^i x^{i+1}}^{y^{i+1}} \right] = \left[ p^0 \prod_{i=1}^n M^{y^i} \right]_{x^n}. \quad (6.3)$$

La loi  $q^n$  est obtenue en normalisant la formule précédente de façon à en faire une loi de probabilité en  $x$  donc en la divisant par :

$$p^0 \prod_{i=1}^n M^{y^i} \mathbf{1}.$$

□

**REMARQUE 6.2.** Le numérateur appelé équation du filtre non normalisé satisfait l'équation récurrente linéaire :

$$r^{n+1} = r^n M^{y_{n+1}},$$

$$r^0 = p^0.$$

L'équation du filtre est alors donnée par :

$$q^n = \frac{r^n}{(r^n, \mathbf{1})}.$$

Si la chaîne de Markov a  $E$  états,  $q^n$  est un processus aléatoire de sauts vivant dans le simplexe :

$$\mathcal{Z}^E = \left\{ x \mid \sum_{i=1}^n x_i = 1, x_i \geq 0, i = 1, \dots, n \right\} .$$

Ses sauts sont définis par :

$$q \rightarrow \frac{qM^y}{qM^y 1} \text{ avec la probabilité } qM^y 1 .$$

## 6.2. EQUATION DE LA PROGRAMMATION DYNAMIQUE

Nous nous plaçons maintenant dans le cadre général d'une chaîne de Markov commandée en observation incomplète. Nous nous donnons donc le 7-uple:

$$(T, \mathcal{E}, \mathcal{F}, \mathcal{G}, M^{uy}, p^0, c^{uy})$$

L'état à mémoriser pour pouvoir appliquer la programmation dynamique est la loi conditionnelle de l'état connaissant le passé des observations. La récurrence en temps devient alors  $\forall q \in \mathcal{Z}^E$ :

$$v^n(q) = \min_{u \in \mathcal{F}} \sum_{y \in \mathcal{G}} \left\{ v^{n+1} \left[ \frac{qM^{uy}}{qM^{uy} 1} \right] qM^{uy} 1 + (q, c^{uy}) \right\} .$$

L'argument du minimum donne une fonction de  $q$  qui est une fonction des observations par l'intermédiaire du filtre optimal.

On remarque que la résolution de cette équation est d'un ordre de grandeur plus difficile à résoudre que l'équation de la programmation dynamique dans le cas de l'observation complète. En effet si dans ce dernier cas l'état est fini de cardinalité  $E$ , dans le cas de l'observation incomplète il est de dimension  $E - 1$ . La discrétisation de cet espace en  $n$  points sur chaque dimension conduit à une autre chaîne de Markov dont l'état est de cardinalité  $n^{E-1}$ . Cette complexité est tellement grande que cette approche est, en général, sans intérêt pratique.

## 7. RÉALISATION ET IDENTIFICATION

Pour pouvoir faire de la commande il faut connaître la loi de la chaîne de Markov. Or bien souvent on ne la connaît pas. Par contre on observe souvent une trajectoire d'un processus que l'on aimerait bien modéliser par une chaîne de Markov. C'est le problème de l'identification. Nous donnons deux résultats : le premier est un résultat purement algébrique donnant la cardinalité de l'espace d'état dans le cadre le plus général, le second est intéressant pratiquement, bien que trivial, il donne un estimateur de la matrice de transition. Ce dernier n'est valable que dans le cas particulier de l'observation complète.

## 7.1. RÉALISATION D'UNE CHAÎNE DE MARKOV

On suppose avoir observé ou connaître la probabilité de séquences d'observations de longueurs qcq. On recherche une chaîne de Markov observée ayant un nombre fini d'états :

$$(\mathcal{T}, \mathcal{E}, \mathcal{G}, M^y, p^0)$$

capable d'expliquer cette loi de probabilité. Pour étudier ce problème il est commode d'utiliser les notations de la théorie des langages. On considère le monoïde des mots  $\mathcal{A}^*$  sur l'alphabet  $\mathcal{A}$ . Par exemple :

$$a = \alpha\beta\gamma\gamma \in \mathcal{A}^* \text{ si } \mathcal{A} = \{\alpha, \beta, \gamma\} .$$

On suppose connue la loi de probabilité  $\mathbb{P}$  de  $Y_n$  c.a.d. qu'à chaque  $a \in \mathcal{A}^*$  on sait associer sa probabilité que l'on écrira  $\langle P, a \rangle$ .

Se donner  $\mathbb{P}$  est équivalent à se donner la série génératrice notée  $P$  définie par :

$$P = \sum_{a \in \mathcal{A}^*} \langle P, a \rangle a .$$

Le problème posé revient alors au problème de la réalisation de cette série  $P$  qui peut être vue comme l'analogie dans ce contexte de la réponse impulsionnelle de la théorie des systèmes linéaires. On peut considérer alors la matrice de Hankel associée à  $P$  définie par

$$\mathcal{H}_{a,b} = \langle P, ab \rangle, \quad \forall a, b \in \mathcal{E}^* ,$$

$ab$  désigne la concaténation des deux mots  $a, b$ . Le problème de trouver la chaîne de Markov de taille minimale ( $E$ ) reste un problème ouvert. On dispose néanmoins du résultat suivant :

**THÉORÈME 7.1.** Le rang de la matrice de Hankel d'une chaîne de Markov observée vérifie :

$$Rg(\mathcal{H}) \leq E .$$

PREUVE. On supposera ici pouvoir écrire  $p^{0y^0}$  sous la forme  $p^0 M^{y^0}$  pour un certain  $p^0$ . Si  $P$  est la série génératrice d'une chaîne de Markov alors la matrice de Hankel dont les éléments sont :

$$\mathcal{H}_{a,b} = p^0 \prod_{\alpha \in a} M^\alpha \prod_{\beta \in b} M^\beta 1$$

se factorise donc sous la forme

$$\mathcal{O}\mathcal{C} ,$$

où  $\mathcal{O}$  est une matrice  $(\infty, E)$  dont la ligne d'index  $a$  vaut :

$$\mathcal{O}_a = p^0 \prod_{\alpha \in a} M^\alpha, \quad \forall a \in \mathcal{A}^* ,$$

et où  $\mathcal{C}$  est une matrice  $(E, \infty)$  dont la colonne d'index  $b$  vaut :

$$\mathcal{C}_b = \prod_{\beta \in b} M^\beta 1, \quad \forall b \in \mathcal{A}^* .$$

On a donc montré le rang de  $\mathcal{H}$  est plus petit que  $E$ .  $\square$

REMARQUE 7.2. La réciproque est un problème ouvert : si  $n = \text{rg}(\mathcal{H})$  on sait trouver des matrices  $M^y$  de rang  $n$  réalisant la série génératrice  $P$  mais on ne sait pas les réaliser avec des matrices qui sont des matrices de transition de chaînes de Markov.

## 7.2. IDENTIFICATION D'UNE CHAÎNE DE MARKOV DONT ON OBSERVE L'ÉTAT

Ce problème est très simple. On considère la chaîne de Markov :

$$(\mathcal{T}, \mathcal{E}, M, p^0)$$

On observe une trajectoire de l'état  $(X^n, n \in \mathcal{T})$  pendant un temps  $T$  fini, on veut estimer  $M$ .

Pour cela notons :

- $N_x^T$  le nombre de fois où  $X^n$  passe par  $x$ ,
- $N_{xy}^T$  le nombre de fois où  $X^n$  fait la transition  $xy$ .

THÉORÈME 7.3. Si la chaîne de Markov est irréductible l'estimateur :

$$\hat{M}_{xy} = \frac{N_{xy}^T}{N_x^T}, \quad \forall x, y \in \mathcal{E},$$

est un estimateur convergent, lorsque  $T \rightarrow \infty$ , asymptotiquement optimal.



## CHAPITRE 4

# PERTURBATION ET AGRÉGATION

## 1. INTRODUCTION ET EXEMPLES

### 1.1. INTRODUCTION

Dans ce chapitre on introduit de la structure sur la chaîne de Markov de façon à diminuer la complexité de la résolution des équations de Kolmogorov. On suppose que la matrice de transition  $M^\epsilon$  dépend d'un paramètre  $\epsilon$  et que la matrice  $M^0$  a une structure bloc-diagonale. Si on étudie la chaîne sur un horizon d'ordre 1 les termes en dehors des blocs diagonaux sont négligeables et le problème se décompose en sous problèmes indépendants de la taille de ces blocs. Par contre lorsque l'on s'intéresse à des horizons de l'ordre de  $1/\epsilon$ , même pour obtenir les termes d'ordre 1 dans un développement en  $\epsilon$ , on ne peut plus négliger les transitions d'ordre  $\epsilon$ . Le but de ce chapitre est d'étudier ce genre de phénomène y compris pour les problèmes de commande optimale. On retrouve ici — dans le cadre des chaînes de Markov — des phénomènes connus en mécanique céleste depuis fort longtemps — Gauss s'était déjà intéressé au sujet.

### 1.2. GESTION DE RÉSERVOIRS

Dans les problèmes de gestion de réservoirs ce phénomène apparaît dans le cas de barrages ayant plusieurs constantes de temps. Par exemple un gros stock et un petit stock en aval disposant des moyens de turbinage du même ordre de grandeur que le gros. Une autre situation de ce type apparaît si l'on considère une gestion annuelle d'un gros réservoir face à une demande fluctuante à l'échelle de la journée ou de la semaine.

### 1.3. FIABILITÉ

Dans les problèmes de fiabilité les taux de panne ne sont pas du même ordre de grandeur que les taux de marche. De plus on s'intéresse ici au système sur un temps long. Dans cet exemple la chaîne comporte d'états transitoires — situation plus compliquée que nous n'aborderons pas dans toute sa généralité.

## 2. CHAÎNES DE MARKOV PERTURBÉES

Etant donné un paramètre  $\epsilon$  destiné à être petit, on suppose que la matrice de transition de la chaîne de Markov a des probabilités de transition

de deux ordres de grandeur, les uns de l'ordre 1, les autres d'ordre  $\epsilon$ . Plus précisément on considère la chaîne de Markov :

$$(T, E, M^\epsilon, c, \mu)$$

où :

- $\epsilon \in \mathbb{R}^+$  un petit paramètre;
- $\mathcal{E}$  désigne l'espace d'état;
- $M^\epsilon$  la matrice de transition de la chaîne de Markov vérifie :

$$M^\epsilon - I = B + \epsilon A ,$$

avec  $B$  [resp.  $A$ ] ayant ses termes hors diagonaux tous positifs, la somme des termes d'une même ligne égale à zéro, la somme des termes hors diagonaux d'une même ligne inférieure à 1;

- $c : \mathcal{E} \rightarrow \mathbb{R}^+$  un coût;
- $\mu \in \mathbb{R}^+$  un taux d'actualisation.

On appelle : *chaîne rapide* la chaîne de Markov de matrice de transition  $M^r = I + B$ ; *chaîne lente* la chaîne de Markov de matrice de transition  $M^l = I + A$ . Considérons alors la fonctionnelle de la trajectoire suivante :

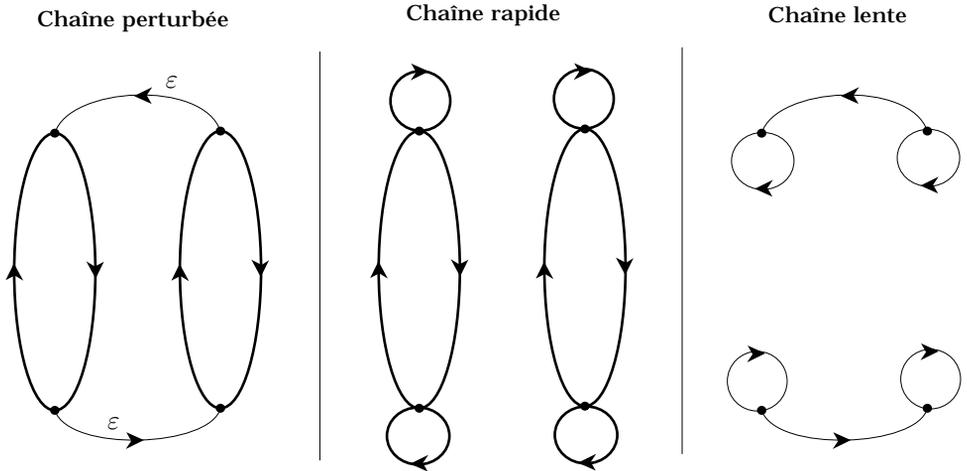


FIGURE 1. Exemple de chaîne perturbée

$$v_x^\epsilon = \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{\epsilon \mu}{(1 + \epsilon \mu)^{n+1}} c_{X^n} \mid X^0 = x \right\} , \quad (2.1)$$

$$= \mathbb{E}\{c_{X^\nu} \mid X^0 = x\} , \quad (2.2)$$

où  $\nu$  est un temps d'arrêt indépendant de  $X^n$  de loi définie par :

$$\mathbb{P}(\nu = n) = \frac{\epsilon \mu}{(1 + \epsilon \mu)^{n+1}} ,$$

et donc vérifie :

$$\mathbb{E}(\nu) = \frac{1}{\epsilon \mu} .$$

L'étude de ces fonctionnelles revient donc à étudier la chaîne de Markov sur un horizon d'ordre  $1/\epsilon$ , un temps suffisamment long pour que la chaîne lente ait une action non négligeable sur le système. On va donc étudier le comportement asymptotique de  $v^\epsilon$  lorsque  $\epsilon \rightarrow 0$ . On a  $v_x^\epsilon$  est borné indépendamment de  $\epsilon$  par  $\sup_{x \in \mathcal{E}} c_x$ .  $v^\epsilon$  satisfait l'équation de Kolmogorov

$$A_\mu v^\epsilon + \frac{1}{\epsilon} B v^\epsilon + c_\mu = 0, \quad (2.3)$$

où l'on a noté:  $A_\mu = A - \mu$  et  $c_\mu = \mu c$ . Mais  $A_\mu$  est inversible puisque  $A$  a toutes ses valeurs propres  $\leq 0$ , et donc  $v^\epsilon$  vaut:

$$\begin{aligned} v^\epsilon &= (\epsilon + A_\mu^{-1} B)^{-1} (-\epsilon A_\mu^{-1} c_\mu), \\ &= -\epsilon R(-\epsilon) A_\mu^{-1} c_\mu, \end{aligned}$$

où  $R$  désigne la résolvante de  $A_\mu^{-1} B$ . Le théorème de Kato (4.2) s'applique et nous donne donc le développement de  $v^\epsilon$  en  $\epsilon$ .

**THÉORÈME 2.1.** Le vecteur  $v^\epsilon$  solution de (2.3) admet le développement suivant :

$$v^\epsilon = -P A_\mu^{-1} c_\mu + \sum_{n=1}^{+\infty} (-\epsilon)^n S^n A_\mu^{-1} c_\mu,$$

où  $P$  est le projecteur spectral sur  $\mathcal{N}(A_\mu^{-1} B)$  et  $S$  satisfait :

$$S A_\mu^{-1} B = A_\mu^{-1} B S = I - P.$$

**PREUVE.** On a  $\mathcal{N}(A_\mu^{-1} B) = \mathcal{N}(B) \neq 0$ . Le théorème (4.2) donne le résultat après avoir remarqué que le nilpotent de  $A_\mu^{-1} B$  associé à la valeur propre 0 est nul car sinon  $v^\epsilon$  ne serait pas majoré indépendamment de  $\epsilon$ .  $\square$

En notant  $v^0 = -P A_\mu^{-1} c_\mu$  et  $v^1 = -S A_\mu^{-1} c_\mu$  la proposition suivante caractérise les deux premiers termes du développement de  $v^\epsilon$ .

**PROPOSITION 2.2.** On a le développement  $v^\epsilon = v^0 + \epsilon v^1 + o(\epsilon)$  où  $v^0$  est l'unique solution de l'un des deux systèmes équivalents suivants :

- (i)  $A_\mu v^0 + B v^1 + c_\mu = 0$ ,  $v^0 \in \mathcal{N}(B)$ ,  $v^1 \in A_\mu^{-1} R(B)$ ,
- (ii)  $(A_\mu v^0, p) + (c_\mu, p) = 0$ ,  $v^0 \in \mathcal{N}(B)$ ,  $\forall p \in \mathcal{N}(B')$ .

**PREUVE.** 1. *Existence d'une solution à (i).* Il suffit de vérifier que  $v^0 = -P A_\mu^{-1} c_\mu$  et  $v^1 = -S A_\mu^{-1} c_\mu$  sont solutions. Pour cela montrons que :

$$-A_\mu P A_\mu^{-1} c_\mu - B S A_\mu^{-1} c_\mu + c_\mu = 0. \quad (2.4)$$

En multipliant (2.4) par  $A_\mu^{-1}$  il suffit alors de prouver que :

$$-P A_\mu^{-1} c_\mu - A_\mu^{-1} B S A_\mu^{-1} c_\mu + A_\mu^{-1} c_\mu = 0.$$

Mais par définition de  $S$  on a  $A_\mu^{-1} B S = I - P$  d'où le résultat.

2. (i)  $\Rightarrow$  (ii). Faisons le produit scalaire de (i) avec  $p \in \mathcal{N}(B')$ , on obtient :

$$(A_\mu v^0, p) + (Bv^1, p) + (c_\mu, p) = 0 ,$$

mais  $(Bv^1, p) = (B'p, v^1) = 0$  puisque  $p \in \mathcal{N}(B')$ .

3. *Unicité de (ii)*

Supposons qu'il y ait deux solutions, notons  $w$  la différence, on a, puisque  $\mathcal{N}(B') = \mathcal{R}(B)^\perp$  :

$$(A_\mu w, p) = 0, \quad w \in \mathcal{N}(B), \quad \forall p \in \mathcal{N}(B') .$$

On a donc :

$$A_\mu w \in \mathcal{R}(B) \Rightarrow w \in A_\mu^{-1} \mathcal{R}(B) ,$$

et donc :

$$\exists z : w = A_\mu^{-1} Bz, \quad z \notin \mathcal{N}(B) .$$

Puisque

$$w \in \mathcal{N}(B) = \mathcal{N}(A_\mu^{-1} B) ,$$

il existe  $\exists z \neq 0$  tel que

$$A_\mu^{-1} Bw = (A_\mu^{-1} B)^2 z = 0 .$$

Le nilpotent de  $A_\mu^{-1} B$  associé à la valeur propre 0 est donc non nul, ce qui est en contradiction avec bornitude de  $v^\epsilon$ .

□

On peut donner, maintenant, l'interprétation, en terme de chaîne de Markov agrégée, du premier terme du développement de  $v^\epsilon$ .

**THÉORÈME 2.3.** Si l'on désigne par :

- $(f_1, f_2, \dots, f_m)$  les classes finales de la chaîne rapide,
- $t$  l'ensemble des états transitoires de la chaîne rapide,
- $A_{lk}$  les probabilités de transition de la chaîne lente,
- $p^i$  les mesures invariantes extrêmes (MPIE) de la chaîne rapide,
- $\chi_l^i$  les probabilités de tomber dans la classe finale  $i$  partant de l'état transitoire  $l$  de la chaîne rapide,

alors les  $q_{ij}$  définis par :

$$q_{ij} = \sum_{l \in f_i} \left( \sum_{k \in f_j} p_l^i A_{lk} + \sum_{k \in t} p_l^i A_{lk} \chi_k^j \right)$$

sont des probabilités de transition sur les classes finales de la chaîne rapide  $(f_1, f_2, \dots, f_m)$  et définissent donc une chaîne de Markov  $Z^n$  appelée chaîne agrégée dont les états sont donc ces classes finales.

Si l'on désigne par :

$$d_i = \sum_{l \in f_i} c_l p_l^i$$

moyenne au sens de la  $i^{\text{ème}}$  MPIE de  $c$ ,  $v^0 = \lim_{\epsilon \rightarrow 0} v^\epsilon$  ne dépend que des classes finales de la chaîne rapide et s'interprète comme le coût moyen — de la chaîne agrégée — suivant:

$$v_i^0 = \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{\mu}{(1+\mu)^{n+1}} d_{Z^n} | Z^0 = f_i \right\}. \quad (2.5)$$

PREUVE. On réécrit le (ii) de la proposition en utilisant les bases de  $\mathcal{N}(B)$  et de  $\mathcal{N}(B')$  que l'on a explicitées au premier chapitre :

$$(A_\mu v^0, p) + (c_\mu, p) = 0, \quad v^0 \in \mathcal{N}(B), \quad p \in \mathcal{N}(B'), \quad v^0 = \sum_j \alpha_j \chi^j, \quad (2.6)$$

où les  $\chi_j$  sont les solutions uniques de

$$\begin{cases} A\chi^j = 0 \text{ sur } t, \\ \chi^j = 1 \text{ sur } f_j, \\ \chi^j = 0 \text{ ailleurs.} \end{cases}$$

L'équation (2.6) se réécrit :

$$\sum_j -\mu(p^i, \chi^j) \alpha_j + \sum_j (p^i, A\chi^j) \alpha_j + (p^i, c_\mu) = 0, \quad i = 1 \cdots m. \quad (2.7)$$

On remarque alors que

$$(p^i, \chi^j) = \delta_i^j,$$

$$(p^i, A\chi^j) = \sum_{l \in f_i} \sum_{k \in f_j} p_l^i A_{lk} + \sum_{l \in f_i} \sum_{k \in t} p_l^i A_{lk} \chi_k^j,$$

$$(c, p^i) = d_i,$$

et donc (2.7) se réécrit

$$Q_\mu \alpha + d_\mu = 0, \quad (2.8)$$

avec avec  $Q_\mu = (q_{ij}) - \mu I$ . Il reste donc à vérifier que  $(I + Q)$  est une matrice stochastique. On a:

$$\sum_j q_{ij} = 0,$$

or

$$\sum_j q_{ij} = \sum_j (p^i, A\chi^j) = (p^i, A \sum_j \chi^j) = (p^i, A1) = 0.$$

D'autre part on a

$$\begin{aligned} \sum_{j \neq i} q_{ij} &= \sum_{l \in f_i} p_l^i \left( \sum_{\substack{k \notin f_j \\ k \notin t}} A_{lk} + \sum_{k \in t} A_{lk} \sum_{j \neq i} \chi_k^j \right), \\ &\leq \sum_{l \in f_i} p_l^i 1, \\ &\leq 1, \end{aligned}$$

et donc, en remarquant que  $q_{ij} \geq 0, i \neq j$ , on a bien que  $I + Q$  est une matrice stochastique. L'équation (2.8) devient alors l'équation de Kolmogorov de la chaîne de matrice de transition  $(I + Q)$  d'où le résultat.  $\square$

On note également  $Q$  par  $\bar{A}^B - A$  moyenné par  $B$ .

EXEMPLE 2.4. Pour illustrer ce résultat reprenons l'exemple considéré au début de ce chapitre. On a

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix},$$

et donc  $f_1 = \{1, 2\}$   $f_2 = \{3, 4\}$  et  $t = \emptyset$ . Les deux mesures de probabilité invariantes extrémales sont solutions de  $p_1 B_1 = 0$  et  $p_2 B_2 = 0$ . La chaîne agrégée a deux états  $f_1$  et  $f_2$  et ses probabilités de transition sont  $q_{12} = p_1^1 A_{13}$  et  $q_{21} = p_4^2 A_{42}$ .

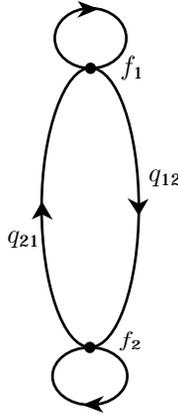


FIGURE 2. chaîne agrégée associée

### 3. COMMANDE DES CHAÎNES DE MARKOV PERTURBÉES

Etant donnée une chaîne de Markov perturbée et commandée:

$$(\mathcal{T}, E, F, M^{u^\varepsilon}, c^u, \mu)$$

où

- $\mathcal{T}$  désigne l'espace temps;
- $\mathcal{E}$  désigne l'espace d'état;
- $\mathcal{F}$  désigne l'espace de commande qui sera supposé compact,  $u \in \mathcal{F}$ ;

- $M^{u\epsilon}$  la famille de matrices de transition perturbées de paramètre  $\epsilon$  et commandées par la commande  $u$  vérifiant

$$M^{u\epsilon} - I = B^u + \epsilon A^u ,$$

avec  $B$  et  $A$  telles que  $\mathcal{N}(B^u)$  soit indépendant de  $u$ , l'application  $u \rightarrow \{A^u, B^u\}$  soit continue, la chaîne rapide n'ait pas de classe transitoire quelque soit  $s \in \mathcal{S}_F$  ;

- $c^u : \mathcal{E} \rightarrow \mathbb{R}^+$  un coût continu en  $u$ ;
- $\mu \in \mathbb{R}^+$  un taux d'actualisation.

on étudie le comportement asymptotique, lorsque  $\epsilon \rightarrow 0$ , du problème de commande optimale

$$v_x^\epsilon = \min_{s \in \mathcal{S}_F} \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{\mu \epsilon}{(1 + \mu \epsilon)^{n+1}} c_{X^n}^{U^n} | X^0 = x \right\} . \quad (3.1)$$

Nous savons que  $v^\epsilon$  est l'unique solution de :

$$-\mu v^\epsilon + \min_{u \in \mathcal{F}} \left\{ (A^u + \frac{1}{\epsilon} B^u) v^\epsilon + c_\mu^u \right\} = 0 , \quad (3.2)$$

équation de la programmation dynamique pour le problème (3.1). Le but de ce paragraphe est d'étudier un développement en  $\epsilon$  de  $v^\epsilon$  et d'en donner des interprétations en terme de commande optimale de chaînes de Markov. On commence par donner un théorème d'existence et d'unicité d'une équation caractérisant les deux premiers termes du développement. On donne ensuite l'interprétation du premier terme du développement. Dans certain cas —  $\mathcal{F}$  fini — on peut donner une caractérisation complète de la série. Nous ne le ferons pas ici. Le but de ce genre de travail est de réduire la complexité des algorithmes numériques de résolution de l'équation de la programmation dynamique. Ce sera ici un succès relatif puisque il y a réduction du nombre d'états mais malheureusement pas de la taille de l'espace de commande.

On note dans la suite :

$$H(v^0, v^1) = \min_{u \in \mathcal{F}} \{ A_\mu^u v^0 + B^u v^1 + c_\mu^u \} .$$

En utilisant les mêmes conventions qu'au paragraphe précédent on a le théorème caractérisant les deux premiers termes du développement suivant:

**THÉORÈME 3.1.** Il existe une solution a :

$$H(v^0, v^1) = 0 , \quad v^0 \in \mathcal{N}(B^{s^*}) , \quad v^1 \in (A_\mu^{s^*})^{-1} \mathcal{R}(B^{s^*}) ,$$

où :

$$s^* : x \rightarrow u^* \in \arg \min_{u \in \mathcal{F}} (A_\mu^u v^0 + B^u v^1 + c_\mu^u)_x ,$$

avec  $v^0$  défini de façon unique.

**PREUVE.** La démonstration combine les méthodes utilisées dans le paragraphe précédent et celles de la commande ergodique.

1. *Existence* On considère l'algorithme du type itération sur les politiques suivant :

- *Etape 1.* A  $s$  on associe  $(v^0, v^1)$  solution de

$$A_\mu^s v^0 + B^s v^1 + c_\mu^s = 0, \quad v^0 \in \mathcal{N}(B^s), \quad v^1 \in (A_\mu^s)^{-1} \mathcal{R}(B^s).$$

Cette solution existe et est unique grâce au paragraphe précédent. On sait de plus que  $v^0 \geq 0$ .

- *Etape 2.* A  $(v^0, v^1)$  donné on associe  $s$  définie par

$$s : x \rightarrow u \in \arg \min_{u \in \mathcal{F}} \{A_\mu^u v^0 + B^u v^1 + c_\mu^u\}.$$

Nous construisons ainsi trois suites  $\{v^{0n}\}, \{v^{1n}\}, \{s^n\}$ .

- *Montrons que la suite  $\{v^{0n}\}$  est décroissante.* Nous avons :

$$A^n v^{0n} + B^n v^{1n} + c_\mu^n = 0,$$

$$A^{n+1} v^{0,n+1} + B^{n+1} v^{1,n+1} + c_\mu^{n+1} = 0,$$

et donc :

$$A^{n+1} (v^{0n} - v^{0,n+1}) + B^{n+1} (v^{1n} - v^{1,n+1})$$

$$+ (A^n - A^{n+1}) v^{0n} + (B^n - B^{n+1}) v^{1n} + c_\mu^n - c_\mu^{n+1} = 0,$$

et donc :

$$A^{n+1} (v^{0n} - v^{0,n+1}) + B^{n+1} (v^{1n} - v^{1,n+1}) \leq 0, \quad (3.3)$$

et donc  $v^{0n} - v^{0,n+1} \geq 0$  grâce à l'interprétation probabiliste des equations du type  $Av + Bw + c = 0$  donné au paragraphe précédent.

- *Passage à la limite.* Puisque  $s^n(x) \in \mathcal{F}$  compact, il existe donc une sous suite  $\{s^{n'}\}$  convergente (nous notons  $s^*$  sa limite). Remarquons également que  $v^{1n} \in (A_\mu^{s^n})^{-1} \mathcal{R}(B^{s^n})$ ,  $u \rightarrow \text{spectre}((A^u)^{-1} \mathcal{R}(B^u))$  est continue et que  $\mathcal{N}(B^s)$  est fixe quelque soit  $s$ . Il existe alors un voisinage de 0 dans le plan complexe ne contenant aucune autre valeur propre de  $(A_\mu^s)^{-1} B^s$  que 0 et donc  $\|(A_\mu^s)^{-1} B^s|_{(A_\mu^s)^{-1} \mathcal{R}(B^s)}\|$  est borné indépendamment de  $s$  et donc  $v^{1n}$  est borné indépendamment de  $s$  et donc appartient à un compact. Il existe alors une sous suite de  $\{v^{1n'}\}$  notée  $\{v^{1n''}\}$  convergente vers  $v^{1*}$  et telle que  $\{v^{1,n''-1}\}$  soit convergente vers  $v^{1**}$ . On a alors :

$$A^{n''} v^{0n''} + B^{n''} v^{1n''} + c^{n''} = 0,$$

$$A^{n''} v^{0,n''-1} + B^{n''} v^{1,n''-1} + c^{n''} \leq A^s v^{0,n''-1} + B^s v^{1,n''-1} + c^s, \quad \forall s,$$

et donc en passant à la limite :

$$A^* v^{0*} + B^* v^{1*} + c^* = 0,$$

$$A^* v^{0*} + B^* v^{1**} + c^* \leq A^s v^{0*} + B^s v^{1**} + c^s, \quad \forall s,$$

$$A^* v^{0*} + B^* v^{1**} + c^* \leq 0,$$

et donc :

$$B^*(v^{1**} - v^{1*}) \leq 0, \quad (3.4)$$

et donc  $B^*(v^{1**} - v^{1*}) = 0$  par le même raisonnement que dans le cas du contrôle ergodique mais appliqué ici à chaque classe irréductible de la chaîne rapide. On conclut alors que  $(v^{0*}, v^{1**})$  est solution.

2. *Unicité de  $v^0$* . Supposons qu'il existe deux solutions  $(v^0, v^1, s)$  et  $(v'^0, v'^1, s')$  on aurait :

$$\begin{aligned} A^s v^0 + B^s v^1 + c^s &= 0, \\ A^{s'} v'^0 + B^{s'} v'^1 + c^{s'} &= 0, \\ A^s v^0 + B^s v^1 + c^s &\leq A^{s'} v'^0 + B^{s'} v'^1 + c^{s'}, \end{aligned}$$

et donc

$$A^{s'}(v^0 - v'^0) + B^{s'}(v^1 - v'^1) \geq 0 \Rightarrow v^0 - v'^0 \leq 0$$

grâce au paragraphe précédent. Le résultat en découle par symétrie du raisonnement. □

Montrons que les deux premiers termes du développement de  $v^\epsilon$  sont  $v^0$  et  $v^1$ .

**THÉORÈME 3.2.** Pour tout  $v^0$  tel que :

$$H(v^0, v^1) = 0, \quad v^0 \in \mathcal{N}(B^{s^*}), \quad v^1 \in (A_\mu^{s^*})^{-1} \mathcal{R}(B^{s^*}),$$

avec :

$$s^* : x \rightarrow u^* \in \arg \min (A_\mu^u v^0 + B^u v^1 + c^u)_x,$$

on a :

$$v^\epsilon = v^0 + o(1).$$

Si de plus :

$$(v^0, v^1) \rightarrow \arg \min \{A_\mu^u v^0 + B^u v^1 + c^u\}$$

est une fonction continue, on a  $v^\epsilon = v^0 + \epsilon v^1 + o(\epsilon)$ .

**PREUVE.** Prenons  $v = v^0 + \epsilon v^1$  nous avons  $H(v, v/\epsilon) = H(v^0 + \epsilon v^1, v^1)$  car  $v^0 \in \mathcal{N}(B^{s^*})$  qui est indépendant de la stratégie  $s^*$ . Comme

$$\lim_{\epsilon \rightarrow 0} H(v^0 + \epsilon v^1, v^1) = H(v^0, v^1) = 0,$$

où

$$H(v^0 + \epsilon v^1, v^1) = o(1),$$

on a donc :

$$H(v, \frac{v}{\epsilon}) = o(1).$$

Cette dernière équation peut être vue comme l'équation de la programmation dynamique de la chaîne de Markov de générateur  $B + \epsilon A$  pour la fonction coût  $c + o(1)$  et donc

$$|v - v^\epsilon| \leq \sup_{x \in \mathcal{E}} \sum_{n=0}^{+\infty} \mathbb{E} \left\{ \frac{\epsilon}{(1 + \epsilon \mu)^{n+1}} o(1) | X_0 = x \right\},$$

d'où la première partie du théorème.

Pour montrer la seconde partie. On prend  $v^2$  défini par  $(A_\mu^{s*})v^1 + B^{s*}v^2 = 0$  ce qui est possible puisque  $v^1 \in A_\mu^{-1}\mathcal{R}(B)$ . En prenant alors

$$v = v^0 + \epsilon v^1 + \epsilon^2 v^2,$$

on a

$$\begin{aligned} H\left(v, \frac{v}{\epsilon}\right) &= H(v^0 + \epsilon v^1 + \epsilon^2 v^2, v^1 + \epsilon v^2), \\ &= H(v^0, v^1) + \epsilon A_\mu^{s*} v^1 + \epsilon B^{s*} v^2 + o(\epsilon), \\ &= o(\epsilon), \end{aligned}$$

car dans ce cas  $H$  est différentiable en  $v^0$  et  $v^1$  et les dérivées sont  $D_{v^0}H = A^{s*}$ , et  $D_{v^1}H = B^{s*}$ . En faisant alors le même raisonnement que dans la première partie de cette démonstration on a montré que  $|v - v^\epsilon| = o(\epsilon)$ .  $\square$

Donnons maintenant une interprétation en terme de commande de chaîne de Markov de ce développement asymptotique.

**THÉORÈME 3.3** (Interprétation stochastique).  $v^0$  est solution du problème de contrôle stochastique suivant :

- les états sont les classes finales  $(f_1, f_2 \dots f_m)$  de la chaîne rapide qui ne dépendent pas du contrôle et qui forment une partition de l'ensemble des états — il n'y a pas de classes transitoires;
- les commandes sont les probabilités de transition d'une classe à l'autre  $q_{ij}$ ,  $i, j = 1 \dots m$  ;
- la fonction coût vaut  $v_i(q_{i1}, \dots, q_{im})$  où les  $v_i$  sont les solutions des problèmes de commande ergodiques suivants :
  - les états sont ceux de la  $i^{\text{ème}}$  classe rapide,
  - les probabilités de transition sont les probabilités de transition de la chaîne rapide  $M^r$ ,
  - les coûts sont :

$$v_i(q_i) = \min_s \sum_{k \in f_i} c_k^s p_k^{si},$$

sous les contraintes :

$$q_{ij} = \sum_{k \in f_i} \sum_{l \in f_j} p_k^{si} A_{kl},$$

où  $p^{si}$  désigne la  $i^{\text{ème}}$  MPIE de la chaîne rapide.

PREUVE. Notons  $\{\chi^i, i = 1 \cdots m\}$  les vecteurs supports des MPIE de la chaîne rapide — indépendants de  $s$  par hypothèse. Ils forment une base de  $\mathcal{N}(B^s)$ . On a donc :

$$\min_u \{A_\mu^u \sum_j \alpha_j \chi^j + B^u v^1 + c^u\} = 0 . \quad (3.5)$$

Cette équation peut être réécrite

$$\min_u \{B^u v^1 + (A_\mu^u \sum_j \alpha_j \chi^j + c^u)\} = \rho_i(\alpha) = 0, \quad i = 1 \cdots m ,$$

qui peut être interprétée comme la solution de  $m$  problèmes de commandes ergodiques associés à chaque classe irréductible de la chaîne rapide — le coût étant  $A_\mu^u \sum_j \alpha_j \chi^j + c^u$ . Mais nous pouvons également considérer les problèmes de commande ergodiques sous contraintes :

$$\begin{cases} \min_u \{B^u v^1 + (c^u + A_\mu^u \sum_j \alpha_j \chi_j)\} & = \xi(q, \alpha) , \\ (p^{si}, A_\mu^s \chi^j) & = q_{ij} . \end{cases}$$

On remarque que  $\xi_i(q, \alpha)$  est fonction uniquement de  $q_{i1} \cdots q_{im}$ . On a fait la convention que  $\xi = +\infty$  s'il n'y a aucun contrôle satisfaisant les contraintes. Le problème initial s'écrit alors :

$$\min_q \xi(q, \alpha) = 0 .$$

Mais on a

$$\xi_i(q, \alpha) = \min_s (p^{si}, c^s) + \sum_j q_{ij} \alpha_j - \mu \alpha_i$$

grâce à l'interprétation des problèmes ergodiques.

Ainsi si nous notons

$$v_i(q_{i1} \cdots q_{im}) = \min_u (p^{si}, c^s) ,$$

on a

$$\xi_i(q, \alpha) = v_i(q_i) + \sum_j q_{ij} \alpha_j - \mu \alpha_i ,$$

et donc (3.5) se réécrit

$$\min_{q_{ij}} \left\{ \sum_j q_{ij} \alpha_j + v_i(q_i) - \mu \alpha_i \right\} = 0 .$$

□

REMARQUE 3.4. Le résultat précédent est intéressant par son interprétation. Pour calculer la solution du problème on préférera "l'itération sur les politiques" car la résolution de problème ergodique sous contrainte est un problème algorithmiquement difficile à résoudre .



## DÉCOMPOSITION

La méthode de la programmation dynamique est utile lorsque le nombre d'états possibles n'est pas trop grand. Dans de nombreuses situations l'espace d'état  $\mathcal{E}$  est un produit cartésien d'ensembles finis,  $\mathcal{E} = \mathcal{E}^1 \times \mathcal{E}^2 \times \dots \times \mathcal{E}^I$ . Appelons "dimension" du système ce nombre  $I$ . Le nombre d'états possibles  $E$  croit de façon exponentielle avec  $I$ . Il devient rapidement impossible de stocker la solution même sur de gros ordinateurs. Le but de ce chapitre est de montrer un exemple, très utilisé dans la théorie des files d'attente, dans lequel la mesure invariante de la chaîne de Markov  $p_x$ ,  $x \in \mathcal{E}$  se factorise en  $p_{x^1} \dots p_{x^I}$  et ceci malgré la présence de couplages entre les sous-systèmes. Dans cet exemple il est possible d'optimiser dans la classe des feedbacks ne cassant pas cette structure (feedback locaux) avec des coûts de stockage et de calcul raisonnables.

### 1. UN RÉSEAU DE FILES D'ATTENTE

On considère une famille de chaînes de Markov. Cette famille a l'interprétation d'un réseau fermé de files d'attente. Chaque file a une capacité limitée et présente une dérivation des clients de l'entrée vers la sortie lorsqu'on sature la capacité de stockage. Précisons le modèle. Indexons par

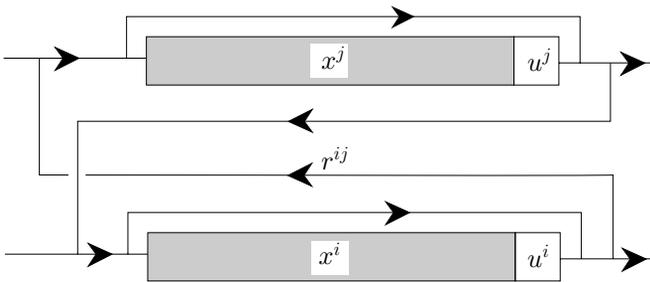


FIGURE 1. Le système de files d'attente

$i \in \mathcal{I} = \{1, 2, \dots, I\}$  les files d'attente. Chaque file a un état  $X^i$ : le nombre de clients en attente dans la file. On notera  $\mathcal{E}^i = \{1, 2, \dots, E^i\}$  l'espace d'état correspondant. L'espace d'état du système complet est alors

$$\mathcal{E} = \{(x^1, x^2, \dots, x^n) \in \prod_{i \in \mathcal{I}} \mathcal{E}^i : \sum_{i \in \mathcal{I}} x^i = E\},$$

où  $E$  est le nombre total de clients, constant, dans le système.

Afin de définir la matrice de transition de la chaîne de Markov on introduit :

- $u^i : \mathcal{E}^i \rightarrow \mathbb{R}^+$  le taux de service de la file  $i$  ;
- $r : \mathcal{I} \times \mathcal{I} \rightarrow [0, 1]$  une matrice stochastique dite de routage qui indique la loi des chemins suivis par les clients dans le réseau —  $r^{ij}$  représente la probabilité pour un client servi par la file  $i$  d'aller dans la file  $j$  si aucune file d'attente n'est pleine;
- $T^{ij} : \mathcal{E} \rightarrow \mathcal{E}$

$$T^{ij}(x^1, \dots, x^I) = (x^1, \dots, x^i - 1, x^{i+1}, \dots, x^j + 1, \dots, x^I) .$$

La matrice de transition  $M^u : \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}^+$  de la chaîne de Markov est alors définie par :

$$M_{xx'}^u = \begin{cases} u_{x^i}^i r_x^{+ij} & \text{s'il existe } i, j : x' = T^{ij}(x) \\ 0 & \text{sinon} \end{cases}$$

avec  $r_x^{+ij}$  la probabilité pour un client d'aller dans la file  $j$  sachant qu'il est sorti de la file  $i$  compte tenu de la saturation des files d'attente. La matrice  $r^+$  se calcule à partir de  $r$ . Pour cela notons  $\mathcal{J}(x^+) \subset \mathcal{I}$  l'ensemble des files non saturées lorsque l'état est  $x$  et que l'on a précisé de quelle file sort le client. On a alors :

$$r^+ = r^{\mathcal{J}} + r^{\mathcal{J}\bar{\mathcal{J}}} s^{\bar{\mathcal{J}}} r^{\bar{\mathcal{J}}\mathcal{J}} ,$$

avec la partition :

$$r = \begin{bmatrix} r^{\mathcal{J}} & r^{\mathcal{J}\bar{\mathcal{J}}} \\ r^{\bar{\mathcal{J}}\mathcal{J}} & r^{\bar{\mathcal{J}}} \end{bmatrix} ,$$

et

$$s^{\bar{\mathcal{J}}} = \sum_{k=0}^{\infty} (r^{\bar{\mathcal{J}}})^k .$$

Cette formule est obtenue en considérant tous les chemins possibles, pour atteindre une file non saturée, passant par les dérivations.

On s'intéresse alors à la mesure invariante  $p$  de la chaîne de Markov  $(\mathcal{T}, \mathcal{E}, M^u)$ .

## 2. FACTORISATION DE LA MESURE INVARIANTE

Nous allons montrer que la mesure invariante  $p$  de tels systèmes se calcule et se factorise  $p = p^1 \dots p^I$  où chaque  $p^i$  s'interprète comme une mesure invariante d'une file unique.

## 2.1. FILE D'ATTENTE MM1

Considérons d'abord la file d'attente à un serveur, à entrée et sortie exponentielle c.a.d. la chaîne de Markov à états dans  $\mathbb{N}$  de matrice de transition :

$$M_{xx'}^u = \begin{cases} u_x & \text{si } x' = x - 1, x' \geq 0, \\ e & \text{si } x' = x + 1, \\ 0 & \text{ailleurs.} \end{cases}$$

Sa mesure invariante doit vérifier, si elle existe, l'équation de Kolmogorov

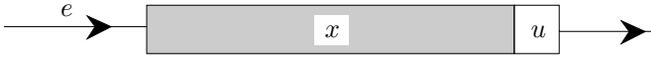


FIGURE 2. Le système ouvert à une file d'attente.

avant à l'équilibre :

$$\begin{cases} p_x(u_x + e) = p_{x+1}u_{x+1} + p_{x-1}e, x \geq 1, \\ p_0e = p_1u_1, \end{cases} \quad (2.1)$$

Le système (2.1) s'intègre facilement on obtient :

$$p_x = k \prod_{y=1}^x \frac{e}{u_y},$$

où  $k$  est une constante de normalisation

$$\frac{1}{k} = \sum_x \left( \prod_{y=1}^x \frac{e}{u_y} \right)$$

qui n'aura un sens que lorsque la file d'attente est stable par exemple lorsque  $e < u_x \forall x$ .

## 2.2. LE SYSTÈME À DEUX FILES D'ATTENTES

On considère cette fois un système fermé de deux files d'attente à capacités limitées. On supposera d'abord que chaque file a une capacité  $E^i$  supérieure au nombre total de clients du système  $E$ . L'équation définissant la mesure invariante vérifie cette fois :

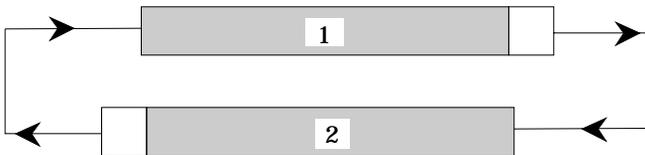


FIGURE 3. Le système fermé à 2 files d'attente.

$$\begin{cases} p_{x^1x^2}(u_{x^1}^1 + u_{x^2}^2) = (p_{x^1+1,x^2-1})u_{x^1+1}^1 + (p_{x^1-1,x^2+1})u_{x^2+1}^2, \\ (p_{0E})u_E^2 = (p_{1,E-1})u_1^1, \\ (p_{E0})u_E^1 = (p_{E-1,1})u_1^2. \end{cases} \quad (2.2)$$

On vérifie que

$$p_{x^1 x^2} = k \prod_{y=1}^{x^1} \frac{1}{u_y^1} \prod_{y=1}^{x^2} \frac{1}{u_y^2},$$

convient.

Si on limite la capacité  $E^i$  de chaque file à une valeur inférieure à la capacité totale du système  $E$ . Les équations (2.2) deviennent :

- (2.2.1) est inchangée,
- (2.2.2) devient (2.2.2') :

$$(p_{E-E^2, E^2}) u_{E^2}^2 = (p_{E-E^2+1, E^2-1}) u_{E-E^2+1}^1,$$

- une équation analogue pour (2.2.3)

et donc la mesure invariante solution de ce nouveau système est du type (2.5) avec une autre constante de normalisation  $k$ .

On constate à posteriori que l'équation (2.2.1) se découpe en deux équations d'équilibre donnant des informations plus précises :

$$\begin{cases} p_{x^1 x^2} u_{x^1}^1 & = p_{x^1-1, x^2+1} u_{x^2+1}^2 \\ p_{x^1 x^2} u_{x^2}^2 & = p_{x^1+1, x^2-1} u_{x^1+1}^1 \end{cases} \quad (2.3)$$

On remarque de plus dans ce cas particulier que (2.3.1) et (2.3.2) sont en fait la même équation mais appliquée à des endroits différents de l'espace d'état. Elles se réécrivent :

$$\frac{p_{x^1 x^2}}{p_{x^1-1, x^2+1}} = \frac{u_{x^2+1}^2}{u_{x^1}^1},$$

Ce sont des équations d'équilibre associées aux transitions joignant seule-

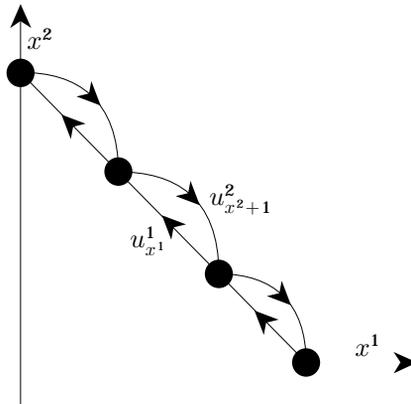


FIGURE 4. L'espace d'état pour le système à 2 files.

ment deux points de l'espace d'état.

Cet exemple est encore trop simple pour nous éclairer complètement sur le cas général puisque l'état est ici en fait de dimension 1.

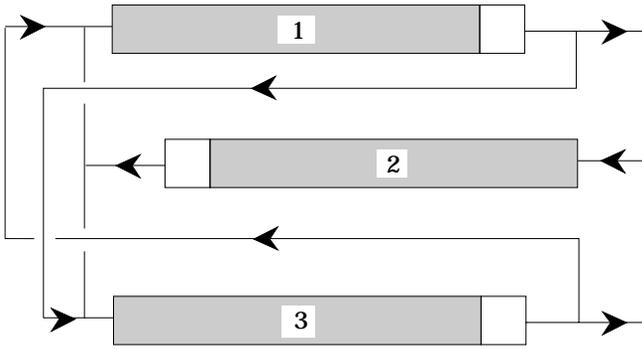


FIGURE 5. Le système fermé de trois files d'attente.

## 2.3. LE SYSTÈME FERMÉ DE TROIS FILES D'ATTENTE

On considère le système fermé à trois files d'attente ayant chacune une capacité supérieure au nombre total de clients du système. L'espace d'état  $\mathcal{E}$  est le simplexe :

$$\{(x^1, x^2, x^3) : x^1 + x^2 + x^3 = E, x^1 \geq 0, x^2 \geq 0, x^3 \geq 0\} .$$

Les équations d'équilibre définissant la mesure invariante s'écrivent :

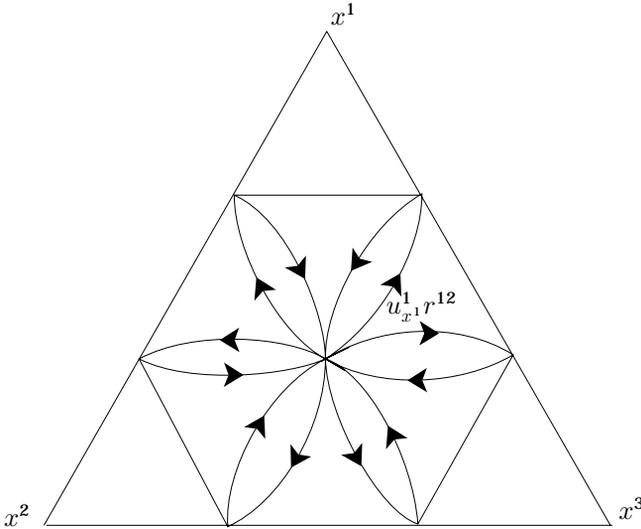


FIGURE 6. L'espace d'état du système à trois files.

$$\begin{aligned} & p_{x^1, x^2, x^3} [u_{x^1}^1 (r^{12} + r^{13}) + u_{x^2}^2 (r^{21} + r^{23}) + u_{x^3}^3 (r^{31} + r^{32})] \quad (2.4) \\ &= (p_{x^1+1, x^2-1, x^3}) u_{x^1+1}^1 r^{12} + (p_{x^1+1, x^2, x^3-1}) r^{13} u_{x^1+1}^1 \\ &+ (p_{x^1-1, x^2+1, x^3}) u_{x^2+1}^2 r^{21} + (p_{x^1, x^2+1, x^3-1}) r^{23} u_{x^2+1}^2 \\ &+ (p_{x^1-1, x^2, x^3+1}) u_{x^3+1}^3 r^{31} + (p_{x^1, x^2-1, x^3+1}) r^{32} u_{x^3+1}^2 . \end{aligned}$$

EXERCICE 2.1. • Ecrire les équations sur chacune des faces du simplexe.

- Ecrire les équations d'équilibre dans le cas où : — la capacité de chaque file est inférieure au nombre total de clients — on a la possibilité de dérivation indiquée au paragraphe précédent.

La solution s'obtient par séparation des variables. On la cherche sous la forme :

$$p_{x^1 x^2 x^3} = c \prod_{y=1}^{x^1} \frac{e^1}{u_y^1} \prod_{y=1}^{x^2} \frac{e^2}{u_y^2} \prod_{y=1}^{x^3} \frac{e^3}{u_y^3} .$$

En reportant dans (2.5) on obtient trois équations associées à des équilibres locaux de flux de probabilité :

$$\begin{cases} e^2 r^{21} + e^3 r^{31} = e^1 r^{12} + e^1 r^{13} , \\ e^1 r^{12} + e^3 r^{32} = e^2 r^{21} + e^2 r^{23} , \\ e^1 r^{23} + e^2 r^{23} = e^3 r^{31} + e^3 r^{32} . \end{cases} \quad (2.5)$$

L'équation (2.5.i) correspond à l'équilibre des flux de probabilité sur la file  $i$ . L'équation (2.5) se réécrit matriciellement

$$er = e . \quad (2.6)$$

Donc  $e$  doit être mesure invariante de la chaîne de Markov des routages  $(\mathcal{I}, r)$ . Supposons que la matrice des routages vérifie la condition d'équilibre local :

$$e^i r^{ij} = e^j r^{ji} . \quad (2.7)$$

On dit dans ce cas que la chaîne de Markov correspondante est "réversible". Ce qui signifie que : la chaîne de Markov obtenue en remontant le temps de matrice de transition  $r' = q_D^{-1} r^t q_D$  — où  $q_D$  désigne la mesure invariante  $r$  mise sur la diagonale d'une matrice (on vérifiera que cela définit bien une matrice de transition) — est la même que celle du sens direct ( $r' = r$ ). Sous cette condition d'équilibre la chaîne de Markov de la file d'attente est réversible et l'équation (6) se découpe cette fois en six équilibres locaux :

$$\begin{cases} e^2 r^{21} = e^1 r^{21} , \\ e^3 r^{31} = e^1 r^{13} , \\ \text{etc} \dots \end{cases} \quad (2.8)$$

On a alors le résultat plus fin :

$$\frac{p_{x'}}{p_x} = \prod_{yy' \subset C_{xx'}} \bar{M}_{yy'}^u , \quad (2.9)$$

où  $C_{xx'} = \{x y_1, y_1 y_2, \dots, y_n x'\}$  désigne un chemin quelconque de l'espace d'état joignant  $x$  à  $x'$

$$\bar{M}_{yy'}^u = \frac{M_{yy'}^u}{M_{y'y}^u} . \quad (2.10)$$

On a par exemple :

$$\bar{M}_{yy'}^u = \frac{u_{x^1}^1 r^{12}}{u_{x^2+1}^2 r^{21}}$$

pour  $y = (x^1, x^2, x^3)$  et  $y' = (x^1 - 1, x^2 + 1, x^3)$ . Cela revient à dire que :

$$\log \bar{M}_{xx'}^u = \log p_x - \log p_{x'} \quad \forall (x, x') \text{ connecté.}$$

Ce qui est l'analogie discret du fait qu'un champ de vecteur dérive d'un potentiel scalaire.

On peut maintenant étudier :

#### 2.4. LE CAS GÉNÉRAL

En remarquant que  $(T^{ji})^{-1} = T^{ij}$  la mesure invariante vérifie l'équation de Kolmogorov avant:

$$\sum_{i \neq j \in \mathcal{I}} p_{T^{ji}(x)} M_{T^{ji}(x), x}^u = p_x \sum_{i \neq j \in \mathcal{I}} M_{x, T^{ij}(x)}^u. \quad (2.11)$$

On a alors le:

**THÉORÈME 2.2.** Le système de files d'attente admet une unique mesure invariante dès que les  $u_y^i > 0, \forall y > 0$ , et que  $r$  est associée à une chaîne de Markov irréductible.

La mesure invariante du système de files d'attente se factorise en :

$$p_x = \begin{cases} k \prod_{i \in \mathcal{I}} \left( \prod_{y=1}^{x^i} \frac{e^i}{u_y^i} \right) & \text{sur } x : \sum_{i \in \mathcal{I}} x^i = E, \quad 0 \leq x^i \leq E^i, \\ 0 & \text{ailleurs,} \end{cases}, \quad (2.12)$$

où :

- $e \in \mathbb{R}^I$  est l'unique mesure invariante de  $r$  et donc satisfait :

$$er = e, \quad (e, 1) = 1; \quad (2.13)$$

- $k$  est une constante de normalisation.

**PREUVE.** L'unicité de la mesure invariante provient des résultats généraux sur les chaînes de Markov. Sous les hypothèses indiquées, la chaîne de Markov associée au système, a une seule classe finale et donc une unique mesure invariante. La formule (2.12) provient de la réalisation des équilibres locaux :

$$p_x \sum_{j \in \mathcal{I}, j \neq i} M_{x, T^{ij}(x)}^u = \sum_{j \in \mathcal{I}, j \neq i} p_{T^{ij}(x)} M_{T^{ij}(x), x}^u \quad \forall i \in \mathcal{I}. \quad (2.14)$$

On vérifie (2.14) en substituant  $p_x$  par sa valeur (2.12). On tombe alors sur la condition suffisante (2.13) de la même façon que dans le cas de trois files d'attente. Il faut ensuite vérifier que le résultat est encore valide sur chacun des bords de l'espace d'état.  $\square$

## 2.5. LE CAS D'UN SYSTÈME OUVERT SUR L'EXTÉRIEUR

Cette nouvelle situation peut être vue comme le cas particulier de la situation précédente dans lequel :

- une file d'attente d'indice 0 modélise l'extérieur;
- le nombre total de clients du système excède le nombre total de place du système interne;
- le nombre de places de la file extérieure est supérieur au nombre total de clients;
- le taux de sortie de la file extérieure est indépendant du nombre de clients en attente dans cette file;

La loi de probabilité du système peut se réécrire :

$$p_x = k p_{x^1}^1 \cdots p_{x^I}^I p_{E-x^1 \dots -x^I}^0$$

qui devient :

$$p_x = k' q_{x^1}^1 \cdots q_{x^I}^I$$

avec :

$$q_{x^i}^i = k^i \prod_{y=1}^{x^i} \frac{e^{i^y}}{u_y^i}$$

et :

$$e^{i^y} = \frac{e^i u^0}{e^0} .$$

Le support de la loi étant un produit cartésien — à cause des hypothèses faites sur le nombre de clients et de places dans le système — les lois d'équilibre sont bien indépendantes. Il est alors beaucoup plus facile de calculer les constantes de normalisation. En résumé, dans ce cas, le système à l'équilibre est équivalent à  $I$  files découplées avec un taux d'entrée  $e^{i^y}$  et un taux de service  $u_{x^i}^i$  .

## 3. FEEDBACKS LOCAUX OPTIMAUX

On étudie le problème de commande optimale  $(\mathcal{T}, \mathcal{E}, \mathcal{F}, M^u, c^u)$  dans le cas ergodique. Il faut donc minimiser :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} c_{X^n}^{U^n} = J(s) .$$

On cherche à optimiser dans la classe des feedbacks locaux  $\mathcal{S}_L$  c.a.d.

$$\mathcal{E} = \prod_{i=1}^I \mathcal{E}^i, \quad \mathcal{F} = \prod_{i=1}^I \mathcal{F}^i, \quad \mathcal{S}_L = \prod_{i=1}^I (\mathcal{F}^i)^{\mathcal{E}^i},$$

par opposition à la classe des feedbacks complets  $\mathcal{S} = \mathcal{F}^{\mathcal{E}}$ , de façon à conserver la forme produit. Entre d'autres termes on commande les taux de sortie  $u^i, i \in \mathcal{I}$  et on leur impose d'être des fonctions de l'état local seul c.a.d.  $u^i$  est fonction de  $x^i$  et non pas du vecteur d'état complet  $x$ . En

se restreignant à cette classe on perd un peu d'optimalité mais on gagne énormément en calcul et en place pour mémoriser la stratégie optimale. D'autre part on se placera, dans toute cette partie, dans le cas d'un système ouvert pour simplifier les résultats mais ils peuvent s'étendre au cas d'un système fermé. A cause de l'absence de convexité à priori du problème, il sera très difficile d'obtenir l'optimum global dans la classe  $\mathcal{S}_L$  mais nous pouvons obtenir une stratégie que l'on ne peut pas améliorer en optimisant file par file ce sont des points de Nash.

**DÉFINITION 3.1.** On appelle point de Nash pour le critère à optimiser  $J(u)$  un point  $u^*$  vérifiant :

$$J(u^{*1}, \dots, u^{*I}) \leq J(u^{*1}, \dots, u^{*i-1}, u^i, u^{*i+1}, \dots, u^{*I}), \quad \forall i \in \mathcal{I}.$$

**REMARQUE 3.2.** Dans le cas convexe différentiable le seul point de Nash coïncide avec le minimum global (dans la classe considérée ici classe des feedbacks locaux).

**EXERCICE 3.3.** On vérifiera sur des exemples en dimension 2 que l'hypothèse de différentiabilité est nécessaire même dans le cas convexe. Les lignes de non différentiabilité peuvent dans certains cas être des points de Nash.

### 3.1. CARACTÉRISATION DES FEEDBACK LOCAUX NASH OPTIMAUX

Nous caractérisons les feedbacks locaux Nash optimaux comme des solutions d'un système couplé d'équation de la programmation dynamique pour cela nous devons définir un coût local moyen

$$c^i : \begin{array}{cccc} \mathcal{S}_L & \times & \mathcal{E}^i & \times & \mathcal{F}^i & \longrightarrow & \mathbb{R}^+ \\ s & & y & & v & & \mathbb{E}_{p^{s\infty}}(c_x^s | x^i = y, u^i = v) \end{array}$$

qui peut s'écrire à partir de la formule explicite de

$$p_x^s = \prod_{j=1}^I p_{x^j}^{js},$$

$$c_y^{iv} = \sum_{x, x^i=y, s^i(x^i)=v} c_x^s \prod_{j \neq i} p_{x^j}^{js}. \quad (3.1)$$

On rappelle que :

$$p_{x^j}^{js} = k^j \prod_{z=1}^{x^j} \frac{e^{t^j}}{s^j(z)}, \quad (3.2)$$

où  $k^j$  est la constante de normalisation.

Si on voit ce problème comme un problème de jeux à  $n$  joueurs, chaque joueur étant identifié à une file d'attente,  $c^j$  représente le coût du  $j^{\text{ème}}$  joueur connaissant seulement l'information de sa file et connaissant la stratégie globale, mais sans avoir accès aux événements qui se sont effectivement produits sur les autres files. La caractérisation suivante devient naturelle.

THÉORÈME 3.4. Sous les hypothèses :

- $s(x) > 0, \forall x \in \mathcal{E}$ ,
- $r$  est irréductible,
- $\mathcal{F}^i$  compacts,
- $u \rightarrow c_x^u$  continue,

une CNS pour que  $w$  soit le coût associé à une stratégie Nash optimal est que le système suivant soit vérifié :

$$\forall i \in \mathcal{I} \left\{ \begin{array}{l} p^{is} A^{is} = 0, \quad \sum_{x \in \mathcal{E}^i} p_x^{is} = 1, \\ \min_u \{ (A^{iu} v^i)_x + c_x^{iu} \} = w, \quad \forall x \in \mathcal{E}^i, \end{array} \right. \quad (3.3)$$

où :

- $c^i$  est défini par (3.1),
- l'équation (3.3.2) définit la stratégie optimale  $s^i$  en associant à chaque  $x^i \in \mathcal{E}^i$  un élément réalisant le minimum de (3.3.2),
- $A^{iu}$  est le générateur de la file i c.a.d.

$$(A^{iu} v^i)_x = e^i v_{x+1}^i + u v_{x-1}^i - (e^i + u) v_x^i.$$

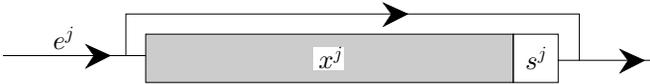


FIGURE 7. File i équivalente.

REMARQUE 3.5. L'équation (3.3.1) définit la mesure invariante de la file  $i$ . L'équation (3.3.2) n'est que l'équation de la programmation dynamique pour le joueur  $i$  optimisant le coût  $c^i$ , sa file étant indépendante des autres grâce à la propriété de la forme produit.

PREUVE. Montrons la condition nécessaire et laissons en exercice la condition suffisante. Si  $s^*$  est Nash optimale elle vérifie :

$$w = J(s^*) \leq J(s^{*1}, \dots, s^i, s^{*i+1}, \dots, s^{*I}),$$

or

$$J(s^{*1}, \dots, s^i, s^{*i+1}, \dots, s^{*I}) = \sum_{y \in \mathcal{E}^i} c_y^{is^i} p_y^{is^i}, \quad (3.4)$$

par définition de  $c^i$ . Les résultats de la théorie du contrôle dans le cas ergodique nous indiquent que le minimum de (3.4) par rapport  $s^i$  vérifie l'équation (3.3.2) d'où le résultat puisque cela doit être vrai  $\forall i \in \mathcal{I}$ . On remarquera que les hypothèses faites assurent l'irréductibilité de la chaîne de Markov du système et l'existence de tous les minima considérés.  $\square$

### 3.2. UN ALGORITHME DE RELAXATION POUR CALCULER UNE STRATÉGIE NASH OPTIMALE

Pour résoudre (3.3) on pourra utiliser l'algorithme suivant qui n'est rien d'autre qu'un point fixe dans l'espace des stratégies :

1. démarrer avec une stratégie  $s \in \mathcal{S}_L$ ,
2. choisir un joueur  $j$ ,
3. calculer  $c^{js}$ ,
4. calculer une nouvelle stratégie  $s^j$  en résolvant l'équation de la programmation dynamique pour le joueur  $j$ ,
5. remettre à jour la probabilité du joueur  $j$  en résolvant l'équation de Kolmogorov avant correspondante,
6. retourner en 2) jusqu'à ce qu'il ne soit plus possible d'améliorer la stratégie  $s$  quel que soit le joueur choisi à l'étape 2).

Par cet algorithme on obtient une suite décroissante de coûts. Les coûts sont minorés si on fait l'hypothèse que  $c \geq 0$ . Cette suite converge vers un coût Nash optimal — car inaméliorable par modification d'un feedback local à la fois.

REMARQUE 3.6. Dans cet algorithme la partie la plus difficile est le calcul des coûts  $c^{js}$ . En fait on ne pourra faire ce calcul que dans des cas particuliers : lorsque le coût est de la forme  $c_x^u = \psi(\sum_i c_{x^i}^{iu^i})$ . On fera alors des convolutions réitérées pour calculer les  $c^{js}$ .



## CHAPITRE 6

# LE RÉGULATEUR LQG

### 1. LE FILTRE DE KALMAN

On se donne le système récurrent stochastique :

$$\begin{cases} x_{k+1} = Ax_k + w_k, & x_0 = \xi, \\ y_k = Cx_k + v_k, \end{cases} \quad (1.1)$$

Ou on suppose que

1.  $x_k$  (resp.  $y_k$ ) sont de dimension  $n$  (resp.  $p$ ),
2. les v.a.  $\xi$ ,  $w_k$  et  $v_k$  sont indépendantes gaussiennes,
3.  $\xi$  est de moyenne  $m_0$  et de variance  $P_0$ ,
4.  $w_k$  (resp.  $v_k$ ) sont centrées, de variance  $M$  (resp.  $N$ ).

Les  $w$  (resp. les  $v$ ) correspondent au bruit sur l'état (resp. l'observation). On peut faire dépendre les matrices  $A$ ,  $C$ ,  $M$ ,  $N$  du temps  $k$  si c'est nécessaire.

On se pose le problème dit du "filtrage", consistant à trouver le meilleur estimé de l'état  $x_k$  connaissant les observations  $y_0, \dots, y_k$ .

On sait que ce meilleur estimé au sens  $L^2$  (moindres carrés), est donné par l'espérance conditionnelle  $\mathbb{E}(x_k | \mathcal{Y}^k)$ , où  $\mathcal{Y}^k$  est la tribu engendrée par  $(y_0, \dots, y_k)$ . Le filtre de Kalman donne une forme récursive, en temps, au calcul de ces espérances conditionnelles.

#### 1.1. CALCUL DU FILTRE

Notons :

1. le prédicteur de  $x_k$  à un pas en temps  $\hat{x}_k = \mathbb{E}(x_k | \mathcal{Y}^{k-1})$ ,
2. la variance de l'erreur de prévision  $\Sigma_k = \text{var}(x_k - \hat{x}_k)$ ,
3. l'état filtré  $\hat{x}_k^+ = \mathbb{E}(x_k | \mathcal{Y}^k)$ ,
4. la variance de l'erreur de filtrage  $\Sigma_k^+ = \text{var}(x_k - \hat{x}_k^+)$ .

**THÉORÈME 1.1.** Ces quantités évoluent selon les équations :

$$\hat{x}_k^+ = \hat{x}_k + K_k(y_k - C\hat{x}_k), \quad \hat{x}_{k+1} = A\hat{x}_k^+, \quad \hat{x}_0 = m_0,$$

le gain du filtre  $K_k$ , étant donné par :

$$K_k = \Sigma_k C' (C \Sigma_k C' + N)^{-1},$$

$$\Sigma_k^+ = (1 - K_k C) \Sigma_k, \quad \Sigma_{k+1} = A \Sigma_k^+ A' + M, \quad \Sigma_0 = P_0.$$

**REMARQUE 1.2.** L'état initial de (1.1) et les bruits, étant gaussiens, le système étant linéaire, les états  $x_k$  et les observations  $y_k$  sont eux-mêmes gaussiens. Le prédicteur à un pas et le meilleur estimé sont aussi gaussiens.

PREUVE. CALCUL DU FILTRE. On appelle innovation le processus :

$$\tilde{y}_k \stackrel{\text{def}}{=} y_k - \mathbb{E}(y_k | \mathcal{Y}^{k-1}) = y_k - C\hat{x}_k = C(x_k - \hat{x}_k) + v_k . \quad (1.2)$$

D'après (6.1) on a :

$$\begin{aligned} \hat{x}_k^+ &\stackrel{\text{def}}{=} \mathbb{E}(x_k | \mathcal{Y}^k) = \mathbb{E}(x_k | \mathcal{Y}^{k-1}) + \mathbb{E}(x_k | \tilde{y}_k) - \mathbb{E}(x_k) \\ &= \hat{x}_k + \mathbb{E}(x_k | \tilde{y}_k) - \mathbb{E}(x_k) = \hat{x}_k + K_k \tilde{y}_k = \hat{x}_k + K_k(y_k - C\hat{x}_k) , \end{aligned}$$

avec

$$K_k = \text{cov}(x_k, \tilde{y}_k) \text{var}(\tilde{y}_k)^{-1} .$$

CALCUL DU GAIN. Pour calculer récursivement  $K_k$  il faut calculer récursivement  $\text{cov}(x_k, \tilde{y}_k)$  et  $\text{var}(\tilde{y}_k)$ .

$$\begin{aligned} \text{cov}(x_k, \tilde{y}_k) &= \text{cov}(x_k, C(x_k - \hat{x}_k)) = \text{cov}(x_k, x_k - \hat{x}_k)C' \\ &= \text{cov}(x_k - \hat{x}_k, x_k - \hat{x}_k)C' = \Sigma_k C' . \end{aligned}$$

Par ailleurs l'indépendance de  $x_k - \hat{x}_k$  et de  $v_k$  entraîne :

$$\text{var}(\tilde{y}_k) = \text{var}(C(x_k - \hat{x}_k)) + \text{var}(v_k) = C\Sigma_k C' + N .$$

On en déduit la valeur du gain du filtre :

$$K_k = \Sigma_k C' (C\Sigma_k C' + N)^{-1} .$$

CALCUL DE LA VARIANCE. De l'orthogonalité  $x_k - \hat{x}_k^+ \perp \tilde{y}_k$  on déduit :

$$\text{var}(x_k - \hat{x}_k) = \text{var}(x_k - \hat{x}_k^+) + \text{var}(K_k \tilde{y}_k) ,$$

et donc

$$\Sigma_k^+ = \Sigma_k - K_k \text{var}(\tilde{y}_k) K_k' = \Sigma_k - K_k C \Sigma_k .$$

CALCUL DU PRÉDICTEUR ET DE SA VARIANCE. On a :

$$\hat{x}_{k+1} = A\mathbb{E}(x_k | \mathcal{Y}^k) + \mathbb{E}(w_k | \mathcal{Y}^k) ,$$

mais  $w_k$  est indépendant de  $\mathcal{Y}^k$  et centrée et donc

$$\hat{x}_{k+1} = A\hat{x}_k^+ .$$

Enfin de la relation :

$$x_{k+1} - \hat{x}_{k+1} = A(x_k - \hat{x}_k^+) + w_k$$

et de l'indépendance de  $x_k - \hat{x}_k^+$  et  $w_k$ , il vient :

$$\Sigma_{k+1} = A\Sigma_k^+ A' + M .$$

□

REMARQUE 1.3. 1. Si les bruits ne sont plus gaussiens, le théorème précédent donne malgré tout, l'évolution de la régression linéaire de  $x_k$  en  $\mathcal{Y}^k$ .

2. Si on suppose les bruits  $w_k, v_k$  corrélés  $\text{cov}(w_k, v_k) = O$ , on se ramène au cas précédent en introduisant

$$\tilde{w}_k = w_k - \mathbb{E}(w_k | v_k) = w_k - ON^{-1}v_k .$$

qui est non corrélé avec  $v_k$  et donc :

$$\begin{aligned} x_{k+1} &= Ax_k + ON^{-1}(y_k - Cx_k) + \tilde{w}_k \\ &= (A - ON^{-1}C)x_k + ON^{-1}y_k + \tilde{w}_k . \end{aligned}$$

L'équation de  $\hat{x}_k$  devient alors :

$$\hat{x}_{k+1} = (A - ON^{-1}C)\hat{x}_k + ON^{-1}y_k ,$$

les autres équations du filtre restent inchangées à la substitution de  $M$  (variance de  $w$ ) par  $M - ON^{-1}O$  près (variance  $\tilde{w}_k$ ).

EXERCICE 1.4. Donner les équations récurrentes satisfaites par

$$\mathbb{E}(x_{k+d} | \mathcal{Y}^k)$$

1. pour  $d \geq 1$  (problème de prévision),
2. pour  $d \leq -1$  (problème de lissage).

EXERCICE 1.5. Si toutes les valeurs propres de  $A$  sont de module strictement inférieur à 1, montrer que le système

$$x_k = Ax_k + w_k, \quad y_k = Cx_k + v_k ,$$

est stable au sens suivant :

1.  $x_k$  converge en loi pour  $k \uparrow \infty$  vers une loi gaussienne centrée de variance  $P$ .  $P$  étant la solution positive de l'équation

$$P = APA' + M .$$

2.  $y_k$  converge en loi vers une loi gaussienne centrée de variance

$$N + CPC' .$$

3. Si  $M$  est symétrique positive,  $P = \sum_{k=0}^{\infty} A^k M (A')^k$  est l'unique solution positive et symétrique de

$$P = APA' + M .$$

## 1.2. COMPORTEMENT ASYMPTOTIQUE DU FILTRE

D'après le théorème 1.1, l'équation des  $\hat{x}_k$  est donnée par,

$$\begin{cases} \hat{x}_{k+1} &= (A - \tilde{K}_k C)\hat{x}_k + \tilde{K}_k y_k , \\ \tilde{K}_k &= AK_k = A\Sigma_k C' (C\Sigma_k C' + N)^{-1} , \\ \Sigma_{k+1} &= A[\Sigma_k - \Sigma_k C' (C\Sigma_k C' + N)^{-1} C\Sigma_k]A' + M . \end{cases} \quad (1.3)$$

On peut réécrire cette dernière équation

$$\Sigma_{k+1} = (A - \tilde{K}_k C)\Sigma_k (A - \tilde{K}_k C)' + \tilde{K}_k N \tilde{K}_k' + M .$$

et remarquer que  $\tilde{K}_k$  réalise le minimum de la covariance à l'instant  $k + 1$  sachant celle à l'instant  $k$ .

$$\Sigma_{k+1} = \min_K [(A - KC)\Sigma_k(A - KC)' + KNK' + M]. \quad (1.4)$$

Le but de ce paragraphe va être d'étudier la convergence en loi des  $\hat{x}_k$ , c'est-à-dire le comportement asymptotique de la variance d'erreur  $\Sigma_k$ .

**THÉORÈME 1.6.** Si  $M = BB'$  avec le triplet  $(A, B, C)$  complètement accessible<sup>1</sup> et observable alors  $\Sigma_k$  converge pour  $k \uparrow \infty$ , vers l'unique solution  $\geq 0$  de :

$$\Sigma = (A - \tilde{K}C)\Sigma(A - \tilde{K}C)' + \tilde{K}N\tilde{K}' + M, \quad (1.5)$$

$$\tilde{K} = A\Sigma C'(C\Sigma C' + N)^{-1}. \quad (1.6)$$

Toutes les valeurs propres de  $A - \tilde{K}C$  sont de module  $< 1$ .

**PREUVE.** 1. LA SUITE  $\Sigma_k$  EST MAJORÉE PAR UNE SUITE CONVERGENTE. En effet la suite étant complètement observable il existe une matrice  $L$  telle que les valeurs propres de  $A - LC$  soient strictement plus petite que 1. Si l'on considère le filtre approché

$$\bar{x}_{k+1} = (A - LC)\bar{x}_k + Ly_k,$$

et la différence  $\tilde{x} \stackrel{\text{def}}{=} x - \bar{x}$  on a

$$\tilde{x}_{k+1} = (A - LC)\tilde{x}_k + w_k - Lv_k,$$

et  $\text{var}(\tilde{x}_k) \geq \Sigma_k$ . Le système gaussien régissant  $\tilde{x}$  étant stable  $\text{var}(\tilde{x}_k)$  converge vers la solution positive de

$$\Sigma = (A - LC)\Sigma(A - LC)' + LNL' + M.$$

2. CONVERGENCE DANS LE CAS  $\Sigma_0 = 0$ . Si  $\Sigma_0 = P_0 = 0$ , la suite  $\Sigma_k$  est croissante dans l'ensemble des matrices  $\geq 0$ . On a  $\Sigma_1 = M \geq \Sigma_0$ . Si par récurrence on suppose  $\Sigma_k \geq \Sigma_{k-1}$ ,

$$\begin{aligned} \Sigma_{k+1} &\geq (A - \tilde{K}_k C)\Sigma_{k-1}(A - \tilde{K}_k C)' + \tilde{K}_k N \tilde{K}_k' + M \\ &\geq \min_K \{(A - KC)\Sigma_{k-1}(A - KC)' + KNK' + M\} \\ &= \Sigma_k \text{ d'après (1.4).} \end{aligned}$$

D'après 1 et 2, pour  $\Sigma_0 = 0$ ,  $\Sigma = \lim_k \Sigma_k$  existe, est  $\geq 0$  et vérifie

$$\begin{cases} \Sigma &= (A - \tilde{K}C)\Sigma(A - \tilde{K}C)' + \tilde{K}N\tilde{K}' + M \\ \tilde{K} &= \lim_k \tilde{K}_k = A\Sigma C'(C\Sigma C' + N)^{-1}. \end{cases} \quad (1.7)$$

<sup>1</sup>Un système  $(A, B)$  est totalement accessible si pour tout  $x$  il est possible de trouver une séquence de commandes transférant en temps fini l'état de 0 à  $x$ .

3. LES VALEURS PROPRES DE  $A - \tilde{K}C$  SONT DE MODULE  $< 1$ . En effet supposons qu'il existe  $z$  tel que

$$(A - \tilde{K}C)'z = \lambda z, \quad |\lambda| \geq 1.$$

De (1.7), il vient<sup>2</sup>

$$(1 - |\lambda|^2)z^* \Sigma z = z^* \tilde{K} N \tilde{K}' z + z^* B B' z.$$

Le premier membre est  $\leq 0$ , les termes du second membre sont  $\geq 0$ ,  $N$  est une matrice inversible. Donc nécessairement,  $\tilde{K}' z = 0$  et  $B' z = 0$ .

Soit alors la matrice  $L$  telle que les valeurs propres de  $A - BL$  soient de module strictement inférieur à 1 (qui existe grâce à l'accessibilité du couple  $(A, B)$ ), on obtient

$$(A - \tilde{K}C)'z = A'z = \lambda z,$$

$$(A - BL)'z = A'z = \lambda z,$$

avec  $|\lambda| \geq 1$  d'où une contradiction.

4. CONVERGENCE DANS LE CAS  $\Sigma_0 = \rho I_d$ . Pour  $\Sigma_0 = \rho I_d$ ,  $\rho > 0$ , soit

$$\Psi_{k,l} = (A - \tilde{K}_{k-1}C)(A - \tilde{K}_{k-2}C) \cdots (A - \tilde{K}_lC).$$

Il est facile de voir que

$$\Sigma_{k+1} = \Psi_{k,0} \Sigma_0 \Psi_{k,0}' + \Phi_k \text{ avec } \Phi_k \geq 0.$$

Donc

$$\Sigma_{k+1} \geq \rho \Psi_{k,0} \Psi_{k,0}'.$$

Mais la suite des  $\Sigma_k$  étant bornée (point 1), il en est de même des  $\Psi_{k,0}$ .

On peut montrer (laissée en exercice) la relation

$$\Sigma_{k+1} - \Sigma = (A - \tilde{K}C)(\Sigma_k - \Sigma)(A - \tilde{K}C)'.$$

D'où

$$\Sigma_{k+1} - \Sigma = (A - \tilde{K}C)^k (\Sigma_0 - \Sigma) \Psi_{k,0}.$$

D'après le point 3,  $(A - \tilde{K}C)^k \downarrow 0$ . D'après le point 4 si  $\Sigma_0 = \rho I_d$   $\Psi_{k,0}$  reste borné. Donc dans ce cas on a encore,  $\lim_k \Sigma_k = \Sigma$ .

5. CONVERGENCE DANS LE CAS GÉNÉRAL. En utilisant (1.4) on montre que, si  $\Sigma_k(\Sigma_0)$  désigne la variance d'erreur pour une condition initiale  $\Sigma_0$ , cette variance d'erreur est une fonction croissante de  $\Sigma_0$ . Donc il existera toujours  $\rho > 0$  tel que :

$$0 \leq \Sigma_0 \leq \rho I_d \Rightarrow \Sigma_k(0) \leq \Sigma_k(\Sigma_0) \leq \Sigma_k(\rho I_d).$$

La convergence de  $\Sigma_k(\Sigma_0)$  résulte donc, des points 2 et 5.

6. UNICITÉ. L'unicité de la solution positive de 1.5 provient du point 5. En effet s'il y avait deux solutions  $\Sigma_1$  et  $\Sigma_2$  on pourrait prendre  $\Sigma_0 = \Sigma_1$  et  $\Sigma = \Sigma_2$  alors  $\Sigma_k(\Sigma_0) = \Sigma_1$  tout  $k$  et comme dans cette section on a montré que  $\lim_k \Sigma_k = \Sigma$  on concluerait que  $\Sigma_1 = \Sigma_2$ .

□

<sup>2</sup>\* signifie transposée conjuguée

COROLLAIRE 1.7. Le filtre approché :

$$\hat{x}_{k+1}^a = (A - \tilde{K}C)\hat{x}_k^a + \tilde{K}y_k, \quad (1.8)$$

où  $\tilde{K}$  est défini par (1.6), a comme variance d'erreur asymptotique  $\text{var}(x - \hat{x}^a)$  la solution positive de (1.5).

PREUVE. En effet notons  $\tilde{x} \stackrel{\text{def}}{=} x - \hat{x}^a$  on a alors

$$\tilde{x}_{k+1} = (A - \tilde{K}C)\tilde{x}_k + w_k - \tilde{K}v_k,$$

et puisque les valeurs propres de  $A - \tilde{K}C$  sont plus petites que 1 ce système linéaire gaussien est stable. Donc la loi de  $\tilde{x}_k$  converge vers une loi gaussienne centrée de variance donnée par

$$\Sigma = (A - \tilde{K}C)\Sigma(A - \tilde{K}C)' + \tilde{K}N\tilde{K}' + M,$$

c.a.d. (1.5). □

## 2. LE PROBLÈME LQG

### 2.1. LA DYNAMIQUE DU SYSTÈME

Les systèmes dynamiques considérés ont pour état  $x_k$ , observation  $y_k$  et commande  $u_k$ . Ils sont donnés par les équations récurrentes stochastiques suivantes :

$$\begin{cases} x_{k+1} &= Ax_k + Bu_k + w_k, \quad x_0 = \xi, \\ y_k &= Cx_k + v_k. \end{cases} \quad (2.1)$$

On suppose que :

1.  $x_k$  (resp.  $y_k$ ) (resp.  $u_k$ ) de dimension  $n$  (resp.  $p$ ) (resp.  $m$ ),
2. les v.a.  $\xi$ ,  $w_k$  et  $v_k$  sont indépendantes gaussiennes,
3.  $\xi$  est de moyenne  $m_0$  et de variance  $P_0$ ,
4.  $w_k$  (resp.  $v_k$ ) sont centrées, de variance  $M$  (resp.  $N$ ).

Les  $w$  (resp. les  $v$ ) correspondent au bruit sur l'état (resp. l'observation). On peut faire dépendre les matrices  $A, B, C, M, N$  du temps  $k$  si c'est nécessaire.

### 2.2. LA STRUCTURE DES COMMANDES

A un instant  $k$  donné, on dispose des observations passées et présentes ( $y_i, i = 0, \dots, k$ ). Il est naturel d'imposer au contrôle  $u_k$  de n'être fonction que de ces informations, donc d'être de la forme :

$$u_k = s_k(y_0, \dots, y_k), \quad k \geq 0. \quad (2.2)$$

Précisons un peu plus cette condition (2.2). Désignons par

1.  $(\Omega, \mathcal{A}, P)$ , un espace de probabilité sur lequel sont définies les v.a.  $\xi, (w_k, v_k), k \geq 0$ ,
2.  $H = L^2(\Omega, \mathcal{A}, P)$  l'espace de Hilbert des v.a. de carré intégrable,
3.  $H^{m \times T}$  l'espace produit, des  $u = (u_k, k = 0, \dots, T - 1)$  avec  $u_k \in H^m$ ,

4.  $\mathcal{Y}^k(u)$ , pour  $u \in H^{m \times T}$ , la tribu engendrée par les observations

$$(y_i, i = 0, \dots, k),$$

solution de (2.1).

La condition (2.2) est équivalente à  $u_k$  est  $\mathcal{Y}^k(u)$  mesurable pour tout  $k$ .

**DÉFINITION 2.1.** Nous désignons par  $\mathcal{U}_{ad}$  l'ensemble des  $u \in H^{m \times T}$  adaptés c.a.d. vérifiant (2.2).

**PROPOSITION 2.2.**  $\mathcal{U}_{ad}$  est un sous-espace de Hilbert de  $H^{m \times T}$ .

Ce résultat est une conséquence du

**LEMME 2.3.** Pour  $u \in \mathcal{U}_{ad}$ , la famille de sous-tribus engendrée par les observations,  $\mathcal{Y}^k(u)$  est fixe (indépendante de  $u$ ).

**PREUVE.** Soit  $(x_k^o, y_k^o)$  la solution de (2.1) associée au contrôle  $u_k = 0$  pour tout  $k$  :

$$\begin{cases} x_{k+1}^o &= Ax_k^o + w_k, \quad x_0^o = \xi, \\ y_k^o &= Cx_k^o + v_k. \end{cases}$$

Et soit  $\mathcal{Y}^k(0)$  la tribu engendrée par  $y_0^o, \dots, y_k^o$ . Par ailleurs pour  $u \in \mathcal{U}_{ad}$ , posons

$$\begin{cases} x_{k+1}^u &= Ax_k^u + Bu_k, \quad x_0^u = 0, \\ y_k^u &= Cx_k^u. \end{cases}$$

Il est clair que, pour tout  $k$  il existe une fonction  $h_k$  de  $\mathbb{R}^{mk}$  dans  $\mathbb{R}^p$ , telle que

$$y_k^u = h_k(u_0, u_1, \dots, u_{k-1}). \quad (2.3)$$

De plus la linéarité des équations implique que la solution de (2.1) associée au contrôle  $u$  s'écrit encore :

$$x_k = x_k^o + x_k^u; \quad y_k = y_k^o + y_k^u. \quad (2.4)$$

Remarquons que l'on a toujours  $y_0^o = y_0$  et donc que  $\mathcal{Y}^0(u) = \mathcal{Y}^0(0)$ . Supposons par récurrence que

$$\mathcal{Y}^i(u) = \mathcal{Y}^i(0), \quad \forall i = 0, \dots, k-1.$$

De (2.3) on déduit que  $y_k^u$  est  $\mathcal{Y}^{k-1}(0)$  mesurable et donc d'après (2.4) que  $y_k$  est  $\mathcal{Y}^k(0)$  mesurable. D'où l'inclusion  $\mathcal{Y}^k(u) \subset \mathcal{Y}^k(0)$ . Mais comme le raisonnement précédent est symétrique en  $y_k$  et  $y_k^o$ , on obtient également l'inclusion inverse.

La proposition est alors conséquence du lemme 5.2.  $\square$

**REMARQUE 2.4.** Si l'observation  $y_k$  n'est pas disponible à l'instant  $k$  mais seulement à l'instant  $k+1$  les contrôles deviennent des feedbacks de la forme

$$u_k = s_k(y_0, y_1, \dots, y_{k-1})$$

ou encore  $u_k$  est  $\mathcal{Y}^{k-1}(u)$  mesurable pour tout  $k$ . Cette différence avec (2.2) ne modifie pas la structure du problème.

### 2.3. LE CRITÈRE

On cherche à minimiser pour  $u \in \mathcal{U}_{ad}$ , le critère quadratique :

$$J(u) = \frac{1}{2} \mathbb{E} \left\{ \sum_{k=0}^{T-1} x'_k Q x_k + u'_k R u_k + x'_T Q_T x_T \right\} \quad (2.5)$$

où la matrice  $Q$  est semi-définie positive et la matrice  $R$  est définie positive. On peut faire dépendre les matrices  $Q$  et  $R$  du temps. L'horizon  $T$  pourra être infini auquel cas la matrice  $Q_T$  sera nulle.

Rappelons que le produit scalaire d'un espace  $H^k$  est défini par

$$\langle x, y \rangle = \mathbb{E}(x'y) ,$$

de telle sorte que le critère  $J(u)$  a la forme,

$$J(u) = \frac{1}{2} \left\{ \sum_{k=0}^{T-1} \langle x_k, Q x_k \rangle + \langle u_k, R u_k \rangle + \langle x_T, Q_T x_T \rangle \right\} . \quad (2.6)$$

La situation est donc classique : “étant donné un système linéaire sur un espace de Hilbert, minimiser un critère quadratique sur un ensemble de contrôle formant lui-même un espace de Hilbert”. On va donc écrire les conditions d'optimalité (principe du minimum de Pontryagin). Et il s'agira d'interpréter de façon probabiliste ce principe du minimum.

On résoudra ce problème en deux étapes. Dans la première étape on se placera en observation complète c.a.d. que l'on supposera que  $y_k = x_k$  pour tout  $k \geq 0$ . Dans la deuxième étape on prouvera le théorème de séparation montrant que la commande optimale, pour le problème en observation incomplète, est égal au feedback optimal obtenu dans le cas de l'observation complète appliqué au meilleur estimé de l'état (au lieu d'être appliqué à l'état lui même qui n'est pas accessible dans cette situation).

## 3. LE RÉGULATEUR LQG EN OBSERVATION COMPLÈTE

On suppose dans ce paragraphe que  $y_k = x_k$  pour tout  $k \geq 0$ .

**LEMME 3.1.** Le critère  $J(u)$  est  $\alpha$ -convexe sur  $\mathcal{U}_{ad}$ , de dérivée :

$$\langle J'(u), v - u \rangle = \sum_{k=0}^{T-1} \langle B' p_{k+1}(u) + R u_k, v_k - u_k \rangle \quad (3.1)$$

où  $p_k$  est solution du système adjoint,

$$p_k(u) = A' p_{k+1}(u) + Q x_k(u), \quad p_T(u) = Q_T x_T(u) . \quad (3.2)$$

PREUVE. Notons  $\delta u = v - u$  et  $\delta x = x(v) - x(u)$ . L'application  $u \rightarrow x(u)$  est affine continue de  $\mathcal{U}_{ad}$  dans  $H^{n \times (T+1)}$  et donc  $J$  est  $\alpha$ -convexe sur  $\mathcal{U}_{ad}$ , de dérivée

$$\begin{aligned} \langle J'(u), \delta u \rangle &= \sum_{k=0}^{T-1} \langle Qx_k(u), \delta x_k \rangle + \langle Ru_k, \delta u_k \rangle \\ &\quad + \langle Q_T x_T(u), \delta x_T \rangle . \end{aligned} \quad (3.3)$$

Mais on a

$$\delta x_{k+1} = A\delta x_k + B\delta u_k, \quad \delta x_0 = 0, \quad (3.4)$$

puisque la dynamique du système est linéaire.

Alors en faisant le produit scalaire de (3.4) par  $p_{k+1}$  et (3.2) par  $\delta x_k$  et en sommant sur  $k$  la différence des résultats on obtient

$$\langle p_T, \delta x_T \rangle = \sum_{k=0}^{T-1} [-\langle Qx_k, \delta x_k \rangle + \langle p_{k+1}, B\delta u_k \rangle],$$

et donc

$$\langle Qx_T, \delta x_T \rangle + \sum_{k=0}^{T-1} \langle Qx_k, \delta x_k \rangle = \sum_{k=0}^{T-1} \langle p_{k+1}, Bu_k \rangle .$$

Cette dernière formule et (3.3) donne alors le résultat par substitution.  $\square$

**LEMME 3.2 (Critère d'optimalité).** Le contrôle  $u \in \mathcal{U}_{ad}$  est optimal si et seulement si pour tout  $k = 0, 1, \dots, T-1$ ,

$$B' \hat{p}_{k+1}^k(u) + Ru_k = 0 \quad (3.5)$$

où  $\hat{p}_{k+1}^k$  est l'espérance conditionnelle de l'état adjoint. Nous avons utilisé la notation :

$$\hat{p}_k^l(u) = \mathbb{E}(p_k(u) \mid \mathcal{X}^l) \quad (3.6)$$

où  $\mathcal{X}^k$  désigne la tribu engendrée par  $(x_0, x_1(u), \dots, x_k(u))$ .

PREUVE.  $\mathcal{U}_{ad}$  étant un espace de Hilbert (proposition 2.2) le contrôle  $u$  est optimal si et seulement si :

$$\langle J'(u), v - u \rangle = 0, \quad \forall v \in \mathcal{U}_{ad}. \quad (3.7)$$

Prenons pour  $v$  :

$$v_i = u_i, \quad \forall i \neq k, \quad v_k = u_k + z_k,$$

avec  $z_k$  de carré intégrable et  $\mathcal{X}^k$ -mesurable. Les tribus  $\mathcal{X}^k$  étant fixes, on a  $v \in \mathcal{U}_{ad}$ . La condition (3.7) devient,

$$\langle B' p_{k+1} + Ru_k, z_k \rangle = 0,$$

ou encore,

$$\mathbb{E}(z_k'(B' p_{k+1} + Ru_k)) = 0.$$

Comme cette relation est vraie pour tout  $z_k$ ,  $\mathcal{X}^k$ -mesurable, les propriétés des espérances conditionnelles (lemme 5.4), impliquent bien (3.5).  $\square$

**THÉORÈME 3.3.** La solution du problème LQG en observation complète est donnée par le feedback linéaire  $u_k = K_k x_k$  avec

$$K_k = -(R + B' P_{k+1} B)^{-1} B' P_{k+1} A, \quad (3.8)$$

où la matrice semi-définie positive  $P_k$  est définie par

$$P_k = A' P_{k+1} A + Q - A' P_{k+1} B (R + B' P_{k+1} B)^{-1} B' P_{k+1} A, \quad (3.9)$$

$$P_T = Q_T.$$

**PREUVE.** D'après la condition d'optimalité (3.5), il s'agit de trouver une solution au système d'équations

$$x_{k+1} = Ax_k + Bu_k + w_k, \quad x_0 = \xi,$$

$$p_k = A' p_{k+1} + Qx_k, \quad p_T = Q_T x_T,$$

$$u_k = -R^{-1} B' \hat{p}_{k+1}^k.$$

Supposons par récurrence que  $\hat{p}_l^l = P_l x_l$  (où  $P_l$  est une matrice semi définie positive) pour  $l = k + 1$  montrons le pour  $k$ . La condition est vraie pour  $l = T$  puisque

$$\hat{p}_T^T = Q_T x_T.$$

Par ailleurs,  $w_k$  étant indépendante de  $\mathcal{X}^k$  et de moyenne nulle

$$\begin{aligned} \hat{p}_{k+1}^k &= P_{k+1} \mathbb{E}(x_{k+1} \mid \mathcal{X}^k) = P_{k+1} \mathbb{E}(Ax_k + Bu_k + w_k \mid \mathcal{X}^k), \\ &= P_{k+1} (Ax_k + Bu_k). \end{aligned}$$

D'après (3.5) on a

$$B' P_{k+1} (Ax_k + Bu_k) + Ru_k = 0.$$

d'où l'on déduit le feedback :

$$u_k = -(R + B' P_{k+1} B)^{-1} B' P_{k+1} Ax_k.$$

Enfin, les relations

$$\hat{p}_k^k = A' \hat{p}_{k+1}^k + Qx_k = A' P_{k+1} (Ax_k + Bu_k) + Qx_k,$$

donnent, en remplaçant  $u_k$  par (3.8), la condition  $\hat{p}_k^k = P_k x_k$ , où  $P_k$  est donnée par (3.9).  $\square$

**REMARQUE 3.4.** Le feedback ne dépend pas de la covariance des bruits. Il est donc le même que dans le cas déterministe ( $w_k = 0, \forall k$ ) ("principe d'équivalence au certain"). Ce résultat est très lié au cadre choisi ici. Il n'est pas vrai en général.

**EXERCICE 3.5.** On peut montrer les résultats de ce paragraphe par la méthode de la programmation dynamique. Soit

$$V(x, k) = \mathbb{E} \left\{ \sum_{i=k}^{T-1} (x_i' Q x_i + u_i' R u_i) + x_T' Q_T x \mid x_k = x \right\}$$

le coût optimal associé au système (2.1) partant de l'état  $x$  à l'instant  $k$ .

1. Montrer que  $V(x, k)$  vérifie l'équation de la programmation dynamique

$$V(x, k) = \min_{u \in \mathbb{R}^m} \{x' Q x + u' R u + \mathbb{E}[V(x_{k+1}, k+1) \mid x_k = x]\} .$$

2. Montrer que  $V(x, k)$  est de la forme :

$$V(x, k) = x' P_k x + s_k$$

où  $P_k$  est une matrice semi-définie positive et  $s_k$  un scalaire et donner les équations d'évolution de  $P_k$  et  $s_k$ .

3. Montrer que  $P_k$  vérifie (3.9).
4. Observer la relation entre  $s_k$  et la covariance des bruits  $Q$ .

**EXERCICE 3.6.** On suppose le système contraint par un retard de  $d$  unités de temps sur l'observation, c'est-à-dire qu'à l'instant  $k \geq d$ , on ne dispose que des observations  $(x_0, x_1, \dots, x_{k-d})$ . Montrer que la solution du problème LQG correspondant est encore linéaire avec le même gain (3.8) où on a remplacé  $x_k$  par la meilleure prédiction de  $x_k$  à l'instant  $k-d$ , à savoir  $\hat{x}_k^{k-d} = \mathbb{E}(x_k \mid \mathcal{X}^{k-d})$ .

#### 4. LE PROBLÈME LQG EN OBSERVATION INCOMPLÈTE

Au problème de commande

$$\left\{ \begin{array}{l} x_{k+1} = Ax_k + Bu_k + w_k, \quad x_0 = \xi, \\ y_k = Cx_k + v_k, \\ \min_{u \in \mathcal{U}_{ad}} J(u) = \mathbb{E} \left\{ \sum_{k=0}^{T-1} (x'_k Q x_k + u'_k R u_k) + x'_T Q x_T \right\} \end{array} \right. , \quad (4.1)$$

on associe le problème de commande déterministe en observation complète

$$(\mathcal{C}) \left\{ \begin{array}{l} x_{k+1} = Ax_k + Bu_k, \quad x_0 = 0, \\ \min_{u \in \mathcal{U}_{ad}} J(u) = \sum_{k=0}^{T-1} (x'_k Q x_k + u'_k R u_k) + x'_T Q x_T \end{array} \right. . \quad (4.2)$$

et le problème de filtrage

$$(\mathcal{O}) \left\{ \begin{array}{l} x_k = Ax_{k-1} + w_{k-1}, \quad x_0 = \xi, \\ y_k = Cx_k + v_k, \\ \mathbb{E}(x_k \mid \mathcal{Y}^k) \end{array} \right. . \quad (4.3)$$

**THÉORÈME 4.1.** Le solution du problème LQG en observation incomplète est donnée par le feedback linéaire

$$u_k = K_k^c \hat{x}_k^+ \quad (4.4)$$

où  $K^c$  est le gain optimal du problème de commande déterministe  $\mathcal{C}$  et  $\hat{x}_k^+$  est donné

$$\hat{x}_k = A \hat{x}_{k-1}^+ + Bu_{k-1}, \quad (4.5)$$

$$\hat{x}_k^+ = \hat{x}_k + K_k^o(y_k - C\hat{x}_k), \quad (4.6)$$

où  $K^o$  est le gain du filtre de Kalman associé au problème  $\mathcal{O}$ .

PREUVE. Il suffit de reprendre la démarche utilisé dans le cas de l'observation complète, en remplaçant les tribus  $\mathcal{X}^k$  par les tribus d'observation  $\mathcal{Y}^k$ , pour aboutir au critère d'optimalité :

$$B' \hat{p}_{k+1}^k + Ru_k = 0, \quad \hat{p}_{k+1}^k = \mathbb{E}(p_{k+1} | \mathcal{Y}^k). \quad (4.7)$$

Supposons par récurrence que :

$$\hat{p}_{k+1}^{k+1} = \mathbb{E}(p_{k+1} | \mathcal{Y}^{k+1}) = P_{k+1} \hat{x}_{k+1}^+,$$

où  $P_k$  est la solution de l'équation de Riccati donnant le gain  $K^c$ . Grâce à (4.7) et de (4.6) et à l'orthogonalité de  $y_{k+1} - C\hat{x}_{k+1}$  avec  $\mathcal{Y}^k$  on obtient :

$$\begin{aligned} \hat{p}_{k+1}^k &= \mathbb{E}(\hat{p}_{k+1}^{k+1} | \mathcal{Y}^k) = \mathbb{E}[P_{k+1}(\hat{x}_{k+1} + K_k^o(y_{k+1} - C\hat{x}_{k+1})) | \mathcal{Y}^k], \\ &= P_{k+1} \hat{x}_{k+1} = P_{k+1}(A\hat{x}_k^+ + Bu_k). \end{aligned}$$

On en déduit :

$$u_k = -(R_k + B'P_{k+1}B)^{-1}B'P_{k+1}A\hat{x}_k^+,$$

qui est (4.4) si l'on se rappelle de la valeur de  $K^c$  (3.8).

Par ailleurs on a :

$$\hat{p}_k^k = A' \hat{p}_{k+1}^k + Q\hat{x}_k^+ = A'[P_{k+1}(A\hat{x}_k^+ + Bu_k) + Q\hat{x}_k^+] = P_k \hat{x}_k^+$$

en remplaçant  $u_k$  par (4.4). □

EXERCICE 4.2. Montrer que le contrôleur peut être vu comme la sortie du système dynamique suivant dont l'entrée est  $y$  et la sortie  $u$  :

$$\begin{aligned} \hat{x}_{k+1}^+ &= (I - K_{k+1}^o)(A + BK_k^c)\hat{x}_k^+ + K_{k+1}^o y_{k+1}, \\ u_{k+1} &= K_{k+1}^c \hat{x}_{k+1}^+. \end{aligned}$$

EXERCICE 4.3. Montrer le théorème de séparation lorsqu'il a un retard sur l'observation en plus d'une observation incomplète c.a.d qu'à l'instant  $k$  on ne dispose que des informations  $y_l$ ,  $l \leq k - d$ , où  $d$  est un entier positif.

## 5. RAPPELS ET COMPLÉMENTS SUR L'ESPÉRANCE CONDITIONNELLE

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité et  $(X_i)_{i \in I}$ , une famille quelconque de v.a. définies sur cet espace. On appelle *tribu engendrée* par cette famille et on note  $\mathcal{B}(X_i, i \in I)$ , la plus petite sous-tribu de  $\mathcal{A}$  rendant mesurable les  $X_i$  c'est-à-dire la tribu engendrée par les ensembles mesurables de  $\mathcal{A}$ , de la forme  $\{X_i \in A\}$ ,  $i \in I$ ,  $A$  borélien de  $\mathbb{R}$ . Si  $\mathcal{B}$  est une sous-tribu de  $\mathcal{A}$ , une v.a.  $X$  est dite  $\mathcal{B}$ -mesurable, si  $\{X \in A\} \in \mathcal{B}$ , pour tout borélien  $A$ .

Soit  $(X_i)_{i \in I}$  une famille finie de v.a., une v.a.  $Y$  est  $\mathcal{B}(X_i, i \in I)$ -mesurable si et seulement si il existe une fonction borélienne  $f$  sur  $\mathbb{R}^{|I|}$  telle que  $Y = f(X_1, X_2, \dots, X_{|I|})$ .

DÉFINITION 5.1. Deux sous-tribus  $\mathcal{B}_1, \mathcal{B}_2$  sont *indépendantes* si

$$\forall B_i \in \mathcal{B}_i, i = 1, 2, P(B_1 \cap B_2) = P(B_1)P(B_2),$$

ou ce qui revient au même :

$$\forall Y_i \geq 0, \mathcal{B}_i - \text{mesurable}, i = 1, 2, \mathbb{E}(Y_1 Y_2) = \mathbb{E}(Y_1)\mathbb{E}(Y_2).$$

En particulier, une sous-tribu  $\mathcal{B}$  et une v.a.  $X$  sont dite indépendantes si les sous-tribus  $\mathcal{B}$  et  $\mathcal{B}(X)$  le sont.

LEMME 5.2. Si  $\mathcal{B}$  est une sous-tribu de  $\mathcal{A}$ , l'espace  $L^2(\Omega, \mathcal{B}, P)$  des v.a.  $Y$ ,  $\mathcal{B}$ -mesurable et de carré intégrable, est un sous-espace de Hilbert de l'espace  $L^2(\Omega, \mathcal{A}, P)$  de toutes les variables aléatoires de carré intégrable définies sur  $(\Omega, \mathcal{A}, P)$ .

On pourra considérer une suite de Cauchy dans  $L^2(\Omega, \mathcal{B}, P)$  (elle converge dans  $L^2(\Omega, \mathcal{A}, P)$ ) et utiliser le fait que d'une suite convergente dans  $L^2$  on peut extraire une sous-suite convergente avec probabilité 1 vers la même limite.

DÉFINITION 5.3. L'espérance conditionnelle d'une variable aléatoire  $Y \in L^2(\Omega, \mathcal{A}, P)$ , par rapport à une sous-tribu  $\mathcal{B}$  de  $\mathcal{A}$ , est définie comme la projection orthogonale de  $Y$  sur le sous-espace de Hilbert  $L^2(\Omega, \mathcal{B}, P)$ . On la note  $\mathbb{E}(Y | \mathcal{B})$ . Si  $\mathcal{B} = \mathcal{B}(X)$ , on écrira souvent  $\mathbb{E}(Y | X)$ .

LEMME 5.4. L'espérance conditionnelle  $\mathbb{E}(Y | \mathcal{B})$  est l'unique variable aléatoire de  $L^2(\Omega, \mathcal{B}, P)$  telle que pour tout  $Z \in L^2(\Omega, \mathcal{B}, P)$  <sup>3</sup>

$$\mathbb{E}(ZY) = \mathbb{E}(Z\mathbb{E}(Y | \mathcal{B})).$$

$\mathbb{E}(\cdot | \mathcal{B})$  a les propriétés suivantes :

1.  $\mathcal{B}_1 \subset \mathcal{B}_2 \Rightarrow \mathbb{E}(Y | \mathcal{B}_1) = \mathbb{E}(\mathbb{E}(Y | \mathcal{B}_2) | \mathcal{B}_1)$ ; <sup>4</sup>
2. Si  $Z$  est  $\mathcal{B}$ -mesurable on a  $\mathbb{E}(Z | \mathcal{B}) = Z$ ; <sup>5</sup>
3.  $Y$  indépendante de  $\mathcal{B} \Rightarrow \mathbb{E}(Y | \mathcal{B}) = \mathbb{E}(Y)$ . <sup>6</sup>

## 6. RAPPEL SUR LES VARIABLES ALÉATOIRES GAUSSIENNES

PROPOSITION 6.1. Deux variables aléatoires gaussiennes sont indépendantes si et seulement si elles sont orthogonales (covariances nulles).

PROPOSITION 6.2. Les espérances conditionnelles de vecteurs gaussiens, coïncident avec leurs régressions affines.

Si  $(X, Y)$  est un vecteur gaussien,  $\mathbb{E}(X | Y)$  est la projection orthogonale de  $X$  sur le sous-espace de  $L^2(\Omega, \mathcal{A}, P)$ , engendré par  $(1, Y)$ , lorsque  $X$  et  $Y$  sont de dimension 1 et plus généralement lorsque  $X$  (resp.  $Y$ ) est

<sup>3</sup>Ce n'est autre que la caractérisation par les produits scalaires, de la projection orthogonale sur un sous-espace de Hilbert.

<sup>4</sup>En particulier,  $\mathbb{E}[\mathbb{E}(Y | \mathcal{B})] = \mathbb{E}(Y)$ .

<sup>5</sup>Plus généralement  $\mathbb{E}(YZ | \mathcal{B}) = Z\mathbb{E}(Y | \mathcal{B})$ .

<sup>6</sup>Plus généralement si  $X$  est  $\mathcal{B}$ -mesurable et si  $Y$  est indépendante de  $\mathcal{B}$ :  $\mathbb{E}(f(X, Y) | \mathcal{B}) = \int f(X, y)P_Y(dy)$ , où  $P_Y$  est la loi de probabilité de  $Y$ .

de dimension  $n$  (resp.  $p$ ) on trouve la projection orthogonale de  $X$  sur le sous-espace de  $L^2(\Omega, \mathcal{A}, P)^n$  des vecteurs aléatoires de la forme  $AY + b$  où  $A$  est une matrice  $n \times p$  et  $b$  un vecteur de dimension  $n$ . Si  $\text{var}(Y)$  est inversible,  $\mathbb{E}(X | Y)$  s'écrit explicitement :

$$\mathbb{E}(X | Y) = \mathbb{E}(X) + \text{cov}(X, Y)\text{var}(Y)^{-1}[y - \mathbb{E}(Y)]. \quad (6.1)$$

**PROPOSITION 6.3.** Si  $(X, Y_1, Y_2)$  est un vecteur gaussien,  $Y_1$  et  $Y_2$  sont orthogonaux et  $x$  est centré on a :

$$\mathbb{E}(X | Y_1, Y_2) = \mathbb{E}(X | Y_1) + \mathbb{E}(X | Y_2). \quad (6.2)$$

## PROPRIÉTÉS DES RÉGULATEURS LQ

Le but de ce chapitre est de compléter l'étude de la méthode consistant à optimiser un critère quadratique pour calculer un régulateur d'un système linéaire dans le cas où on observe l'état. On montre que, de cette manière, on peut placer les pôles du système bouclé tout en respectant des marges de stabilité importantes. D'autre part l'introduction de la notion de système positif permet de montrer la robustesse du contrôleur à des nonlinéarités grâce à un théorème de Popov assurant la stabilité d'un système positif en contre réaction avec un système nonlinéaire vérifiant certaines propriétés. On donne d'autre part une introduction au calcul des régulateurs LQ dans le domaine spectral par une méthode qui est une version optimisation du filtrage de Wiener.

### 1. RAPPEL SUR LE RÉGULATEUR LQ

On considère le système linéaire suivant

$$\dot{x} = Ax + Bu, \quad x(t_0) = \xi,$$

où  $x(t)$  (resp.  $u(t)$ ) est un vecteur de dimension  $n$  (resp.  $m$ ),  $\xi$  est un vecteur de dimension  $n$ ,  $A$  est une matrice  $n \times n$  et  $B$  est une matrice  $n \times m$ . On détermine la commande en minimisant le critère

$$J(u) = \int_{t_0}^{t_f} \begin{pmatrix} x' & u' \end{pmatrix} \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} dt + x'(t_f)P(t_f)x(t_f),$$

où les matrices  $\begin{pmatrix} Q & S \\ S' & R \end{pmatrix}$  et  $P(t_f)$  sont positives et  $R$  est définie positive.

On a le résultat.

**THÉORÈME 1.1.** Si l'équation de Riccati

$$\dot{P} + A'P + PA - (PB + S)R^{-1}(B'P + S') + Q = 0,$$

a une solution (coincisant en  $t_f$  avec la matrice  $P(t_f)$  donné dans le critère) alors la commande optimale est donnée par

$$u_* = -R^{-1}(B'P + S')x,$$

et les fonctionnelles quadratiques de la trajectoire s'écrivent en fonction de la commande optimale sous la forme

$$J(u) - x'(t_0)P(t_0)x(t_0) = \int_{t_0}^{t_f} (u_* - u)'R(u_* - u)dt,$$

et donc  $\xi'P(t_0)\xi$  est le coût optimal  $J_*$ .

PREUVE. Supposons qu'il existe une solution à l'équation de Riccati (1.1) sur  $[t_o, t_f]$ . En utilisant les équations vérifiées par  $P$  et par  $u_*$  on a :

$$\begin{aligned} & \begin{pmatrix} x' & u' \end{pmatrix} \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \\ &= -x' \dot{P}x - x' P A x - x' A' P x + x' (P B + S) R^{-1} (B' P + S') x + 2x' S u + u' R u \\ &= -\frac{d}{dt} (x' P x) + 2x' (P B + S) u + x' (P B + S) R^{-1} (B' P + S') x + u' R u \\ &= -\frac{d}{dt} (x' P x) - 2u'_* R u + u'_* R u_* + u' R u . \end{aligned}$$

□

Nous nous intéressons surtout au cas particulier  $S = 0$  dans la cas stationnaire. Rappelons le résultat important suivant.

**THÉORÈME 1.2.** Sous les hypothèses  $Q = CC'$  avec la commandabilité et l'observabilité du système  $(A, B, C)$  et  $R$  définie positive l'équation de Riccati

$$A' P + P A - P B R^{-1} B' P + Q = 0$$

a une solution positive unique qui a l'interprétation

$$\xi' P \xi = \inf_u \left\{ \int_0^\infty (x' Q x + u' R u) dt \mid x(0) = \xi, \dot{x} = A x + B u \right\} ,$$

et la commande optimale a pour valeur

$$u_* = -R^{-1} B' P x .$$

Le système commandé par ce feedback est stable.

## 2. PLACEMENT DE PÔLES PAR LE RÉGULATEUR LQ

Nous reprenons l'exposé de [60]. Considérons le problème de régulation temps invariant en horizon infini mono-entrée mono-sortie suivant

$$\dot{x} = A x + B u, \quad x(0) = \xi, \quad y = C x ,$$

$$J(u) = \int_0^\infty (y^2 + r u^2) dt ,$$

et étudions l'évolution des pôles du système bouclé optimal (minimisant  $J(u)$  par rapport  $u$ ) lorsque  $r$  varie de 0 à l'infini. Les pôles du système bouclé optimal sont les zéros de  $\det(sI - A + BK)$  avec  $K = r^{-1} B' P$  où  $P$  est solution de l'équation de Riccati

$$A' P + P A - P B r^{-1} B' P + C' C = 0 .$$

**THÉORÈME 2.1.** Les pôles d'un système  $(A, B, C)$  mono-entrée mono-sortie commandable et observable, bouclé par un contrôleur LQ optimal minimisant  $\int_0^\infty (y^2 + r u^2) dt$  , sont les racines stables de

$$d(s)d(-s) + r^{-1}n(s)n(-s) = 0 ,$$

où les polynômes  $n$  et  $d$  sont définis par

$$H(s) = C(sI - A)^{-1}B = n(s)/d(s),$$

avec  $H$  irréductible.

PREUVE. A l'optimum le système satisfait

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & -Br^{-1}B' \\ -C'C & -A' \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix}.$$

Le changement de variable

$$\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} I & 0 \\ P & I \end{pmatrix} \begin{pmatrix} \xi \\ \pi \end{pmatrix},$$

où  $P$  est solution de l'équation de Riccati

$$A'P + PA - PBr^{-1}B'P + C'C = 0,$$

donne le système

$$\begin{pmatrix} \dot{\xi} \\ \dot{\pi} \end{pmatrix} = \begin{pmatrix} A - BK & -Br^{-1}B' \\ 0 & -A' + K'B' \end{pmatrix} \begin{pmatrix} \xi \\ \pi \end{pmatrix},$$

où  $K = r^{-1}B'P$ . Et donc les pôles du système bouclé (valeurs propres de  $A - BK$ ) sont des zéros de

$$p(s) = \det \begin{pmatrix} sI - A & Br^{-1}B' \\ C'C & sI + A' \end{pmatrix} = 0.$$

On a

$$\begin{aligned} p(s) &= \det(sI - A) \det(sI + A' - C'C(sI - A)^{-1}Br^{-1}B') \\ &= \det(sI - A) \det(I - C'C(sI - A)^{-1}Br^{-1}B'(sI + A')^{-1}) \det(sI + A') \\ &= d(s)d(-s)(-1)^n \det(1 - r^{-1}B'(sI + A')^{-1}C'C(sI - A)^{-1}B) \\ &= d(s)d(-s)(-1)^n \left[ 1 + \frac{n(s)n(-s)}{rd(s)d(-s)} \right], \end{aligned}$$

où nous avons utilisé  $\det(I_n - AB) = \det(I_m - B'A')$  si  $A$  et  $B'$  sont des matrices  $n \times m$ .  $\square$

On en déduit les deux résultats suivants qui donnent les deux extrémités du lieu des pôles lorsque  $r$  varie de 0 à l'infini. En pratique on commencera par  $r = 0$  et si on obtient des commandes trop grandes on augmentera  $r$ .

**COROLLAIRE 2.2.** La limite des pôles du système bouclé optimal (lorsque  $r$  tend vers l'infini) (contrôle cher) sont les pôles stables du système en boucle ouverte et les symétrisés par rapport à l'axe imaginaire du plan complexe des pôles instables du système en boucle ouverte.

PREUVE. Lorsque  $r$  tend vers l'infini les pôles du système bouclé optimal vont converger vers les racines stables de  $d(s)d(-s)$  d'après la continuité des racines d'un polynôme par rapport aux coefficients du polynôme. Les racines de  $d(s)d(-s)$  étant symétriques par rapport à l'axe imaginaire

on peut séparer ces racines en  $n$  stables et  $n$  instables. On en déduit le résultat.  $\square$

**COROLLAIRE 2.3.** La limite des pôles du système bouclé optimal (lorsque  $r$  tend vers 0) (contrôles bon marchés) sont les zéros stables du système en boucle ouverte, et les symétrisés par rapport à l'axe imaginaire du plan complexe des zéros instables du système en boucle ouverte, et les zéros à l'infini stables (c.a.d. les asymptotes dans le demi-plan gauche du plan complexe solution de la courbe

$$s^{2(n-m)} = (-1)^{n-m+1} n_0^2 / r$$

où nous avons noté  $m$  le degré de  $n$  et  $n_0$  son coefficient de plus haut degré), et les symétrisés par rapport à l'axe imaginaire des zéros à l'infini instables.

**PREUVE.** Lorsque  $r$  tend vers 0 les pôles du système bouclé restant à distance fini convergent vers les racines stables de  $n(s)n(-s)$ . Mais le degré de  $d$  est supérieur au degré de  $n$ . La courbe

$$d(s)d(-s) + r^{-1}n(s)n(-s) = 0 ,$$

se comporte pour les  $s$  grand comme

$$(-1)^n s^{2n} + r^{-1}(-1)^{-m} n_0^2 s^{2m} = 0 ,$$

ce qui donne  $2(n - m)$  racines à l'infini qui sont solutions de

$$(-1)^n s^{2(n-m)} + (-1)^{-m} n_0^2 / r = 0 ,$$

d'où le résultat.  $\square$

### 3. APPROCHE FRÉQUENTIELLE DU RÉGULATEUR LQ

Nous suivons ici [59] en changeant la démonstration du théorème. La relation entrée sortie du système temps invariant

$$\dot{x} = Ax + Bu, \quad x(0) = \xi, \quad y = Cx ,$$

où  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  et  $y(t) \in \mathbb{R}^p$  dans le domaine fréquentiel<sup>1</sup> vérifie la relation

$$Y = HU + \Xi, \quad \text{avec } H = C(s - A)^{-1}B, \quad \Xi = C(s - A)^{-1}\xi .$$

Le critère quadratique

$$J(u) = \int_0^\infty (y'Qy + u'Ru)dt ,$$

avec  $Q$  et  $R$  définies positives,  $(A, B, C)$  commandable et observable, s'écrit dans le domaine fréquentiel pour des commandes  $u$  stabilisant le

<sup>1</sup>Les transformées de Laplace bilatères de fonctions notées par des minuscules sont notées par les majuscules correspondantes.

système<sup>2</sup>

$$J(U) = 1/(2\pi) \int_{-\infty}^{\infty} [(U^b H^b + \Xi^b)Q(HU + \Xi) + U^b RU](i\omega)d\omega .$$

On veut minimiser  $J(U)$  parmi les transformées de Laplace de commandes causales ( $U$  est transformée de Laplace de fonction à support sur  $[0, \infty)$ ) stabilisant le système. Le résultat énoncé ne sera utile en pratique que dans le cas où le système en boucle ouverte est stable et d'inverse stable. Dans les autres cas des simplifications illicites peuvent intervenir (un zéro et un pôle du demi plan droit se simplifient). On sait résoudre cette difficulté en donnant une paramétrisation de toutes les commandes stabilisantes (voir [59] par exemple) mais on ne le fera pas dans ce paragraphe introductif.

Le fonction de transfert  $U_o$  telle que  $U = U_o \Xi$  est appelée contrôleur en boucle ouverte. On a alors  $Y = (I + HU_o)\Xi$  qui peut s'écrire  $Y = (I + HU_f)^{-1}\Xi$  pour une fonction de transfert  $U_f$  que l'on appellera contrôleur en boucle fermée. Mais alors  $-U_f Y = U_o \Xi$  et donc  $U_f = -(I + U_o H)^{-1}U_o$  donne explicitement  $U_f$  en fonction de  $U_o$ . Il reste donc à calculer  $U$  optimal.

**THÉORÈME 3.1.** La transformée de Laplace de la commande optimale est donnée par

$$U = -\tilde{F}^{-1}\{(\tilde{F}^b)^{-1}H^b Q \Xi\}_+, \text{ avec } \tilde{F}^b \tilde{F} = H^b QH + R .$$

où  $\{.\}_+$  désigne la partie causale et  $\tilde{F}$  est transformée de Laplace d'une fonction causale.

Dans le cas rationnel propre (degré du numérateur inférieur au degré du dénominateur) et stable, cadre auquel nous restreignons,  $\{.\}_+$  revient après une décomposition en éléments simples à conserver les fractions ayant un pôle dans le demi-plan gauche du plan complexe. Le résultat est analytique dans le demi-plan droit.

**PREUVE.** En revenant dans le domaine temporel mais avec un point de vue opératoire la relation entrée sortie s'écrit  $y = h * u + \xi$  où  $*$  dénote l'opérateur de convolution et  $\xi(\cdot)$  dénote ici abusivement la réponse du système pour une commande nulle et une condition initiale  $x(0) = \xi$ . Le contrôle  $u$  doit être vu ici comme la réponse impulsionnel d'un contrôleur, et donc, doit être causal, c.a.d. à support sur  $[0, \infty)$ . Le critère s'écrit alors

$$J(u) = \int_0^{\infty} [(\xi' + u' * h')Q(h * u + \xi) + u' Ru](t)dt ,$$

ce qui peut se réécrire puisque toutes les fonctions intervenant dans le critère sont causales

$$J(u) = \int_{-\infty}^{\infty} [(\xi' + u' * h')Q(h * u + \xi) + u' Ru](t)dt ,$$

<sup>2</sup>  $H^b(s) \stackrel{\text{def}}{=} H'(-s)$ .

et donc en utilisant les produits scalaires de  $L^2((-\infty, \infty); \mathbb{R}^n)$  noté  $V_n$  on a

$$J(u) = \langle h * u + \xi, Q(h * u + \xi) \rangle_{V_p} + \langle u, Ru \rangle_{V_m} .$$

On doit minimiser  $J(u)$  sur  $V_m^+ \stackrel{\text{def}}{=} \{u \in V_m : u(t) = 0, \forall t \leq 0\}$ . Ce problème admet une solution unique puisque  $Q$  est définie positive et  $R$  est positive.

Réécrivons  $J(u)$  sous une autre forme<sup>3</sup>

$$J(u) = u^b * (h^b * Qh + R) * u + 2\xi^b * Qh * u + \xi^b * Q\xi ,$$

qui devient en utilisant la factorisation<sup>4</sup>  $h^b * Qh + R = \tilde{f}^b * \tilde{f}$  avec  $\tilde{f} \in V_{m^2}^{+5}$  et en passant dans le domaine fréquentiel pour pouvoir justifier l'utilisation d'inverse de convolution

$$J(u) = J(U) = \|\tilde{F}U + (\tilde{F}^b)^{-1}H^bQ\xi\|^2 + \xi^bQ\xi - \|(\tilde{F}^b)^{-1}H^bQ\xi\|^2 .$$

Minimiser  $J(U)$  revient alors à minimiser

$$\begin{aligned} L(U) &= \|\tilde{F}U + (\tilde{F}^b)^{-1}H^bQ\xi\|^2 \\ &= \|\{\tilde{F}U + (\tilde{F}^b)^{-1}H^bQ\xi\}_+\|^2 + \|\{\tilde{F}U + (\tilde{F}^b)^{-1}H^bQ\xi\}_-\|^2 \\ &= \|\tilde{F}U + \{(\tilde{F}^b)^{-1}H^bQ\xi\}_+\|^2 + \|\{(\tilde{F}^b)^{-1}H^bQ\xi\}_-\|^2 , \end{aligned}$$

puisque  $\tilde{f} * u \in V_m^+$ . Le contrôleur optimal vaut alors

$$U = \tilde{F}^{-1}\{(\tilde{F}^b)^{-1}H^bQ\xi\}_+ .$$

□

Pour achever la démonstration précédente on a besoin du lemme suivant qui est également important pour obtenir des propriétés qualitatives des régulateurs LQ.

**LEMME 3.2 (factorisation).** Si  $(A, B, C)$  est commandable et observable on a

$$H^bQH + R = F^bRF ,$$

avec

$$F = I + K(sI - A)^{-1}B ,$$

où  $K$  est obtenu en résolvant l'équation de Riccati du problème LQ:

$$K = R^{-1}B'P, \quad A'P + PA - PBR^{-1}B'P + C'QC = 0 .$$

D'autre part  $F^{-1}$  est stable.

On appelle *erreur de retour* la fonction de transfert  $F$  car on peut la voir comme la différence de  $I$  et de  $-K(sI - A)^{-1}B$  retour de la boucle.

<sup>3</sup> $\langle f, g * h \rangle = \langle g^b * f, h \rangle$ .

<sup>4</sup>L'existence d'une telle factorisation sera prouvé dans le lemme suivant.

<sup>5</sup>On a noté  $V_{m^2} \stackrel{\text{def}}{=} L^2((-\infty, \infty); \mathbb{R}^{m \times m})$ .

PREUVE. L'équation de Riccati implique

$$(sI + A')P + P(A - sI) - PBR^{-1}B'P + C'QC = 0.$$

En multipliant cette équation à gauche par  $-B'(sI + A')^{-1}$  et à droite par  $(sI - A)^{-1}B$  et en ajoutant  $R - R$  on obtient

$$R + H^bQH = B'P(sI - A)^{-1}B - B'(sI + A')^{-1}PB \\ - B'(sI + A')PBR^{-1}B'P(sI - A)^{-1}B + R,$$

qui se réécrit

$$R(F - I) + (F^b - I)R + (F^b - I)R(F - I) + R = R + H^bQH,$$

d'où la première partie du théorème en remarquant que le membre de gauche vaut  $F^bRF$ .

La stabilité de  $F^{-1}$  provient de la stabilité du système régulé par un contrôleur LQ. En effet on a dans le cas mono-entrée

$$1 - K(sI - A + BK)^{-1}B = (1 + K(sI - A)^{-1}B)^{-1}, \quad (3.1)$$

qui montre que le système optimal peut être vu comme la mise en feedback avec une rétroaction unitaire du système  $K(sI - A)^{-1}B$  (qui correspond à ouvrir la boucle au niveau du contrôle).  $\square$

De ce lemme on déduit des résultats sur la marge de stabilité du système par un contrôleur LQ optimal.

**COROLLAIRE 3.3.** Le système mono-entrée mono-sortie commandé par le régulateur LQ optimal donne un lieu de Nyquist du système en boucle ouverte restant en dehors du cercle de rayon unité centré en  $(-1, 0)$  dans le plan complexe.

PREUVE. On a  $F^bRF = R + H^bQH$  grâce au lemme précédent et puisque que  $Q \geq 0$  on a  $[F^bRF](j\omega) \geq R$  et donc  $|F(j\omega)| \geq 1$  mais  $F(s) = 1 + K(sI - A)^{-1}B$  et donc le lieu de Nyquist de  $K(sI - A)^{-1}B$  reste en dehors du cercle de rayon unité centré en  $(-1, 0)$ . Le résultat découle alors de l'interprétation de l'égalité (3.1).  $\square$

**REMARQUE 3.4.** Il est intéressant de remarquer que l'on a obtenu ainsi la commande optimale en feedback sur la sortie et non pas sur l'état. Ce résultat ne sera applicable que dans le cas où le système en boucle ouverte est stable et d'inverse stable ce qui est rarement suffisant en pratique.

## 4. SYSTÈMES POSITIFS ET ROBUSTESSE DU RÉGULATEUR LQ

Ce paragraphe reprend des parties de [57, 56, 52]. On définit les systèmes positifs. On montre que la synthèse des contrôleurs par la méthode LQ conduit à des systèmes bouclés positifs. On montre que ces systèmes restent stables lorsqu'ils sont soumis à des perturbations nonlinéaires à condition qu'elles appartiennent à des cônes dont on précisera la frontière.

**DÉFINITION 4.1 (Positivité).** Le système linéaire<sup>6</sup>

$$\begin{cases} \dot{x} &= Ax + Bu, \\ y &= Cx + Du, \end{cases} \quad (4.1)$$

où

1. le triplet  $(A, B, C)$  est commandable et observable,
2.  $A$  est stable  $(\lambda(A) < 0)$ ,

est dit *positif* si  $\int_{-\infty}^{\infty} u'(t)y(t)dt \geq 0$  pour tout  $u$  à support compact.

**EXEMPLE 4.2.** Les systèmes électriques sans générateur de courant ou de tension sont positifs.

Par exemple une capacité  $C$  et une résistance en parallèle  $R$  avec comme entrée la tension  $u$  et comme sortie l'intensité  $i$  totale traversant le circuit ( $i = u/R + C\dot{u}$ ) est un système positif. En effet on a

$$\begin{aligned} \int_{-\infty}^{\infty} u i dt &= \int_{-\infty}^{\infty} \left( \frac{u^2}{R} + Cu \frac{du}{dt} \right) dt \\ &= C \frac{u^2(\infty) - u^2(-\infty)}{2} + \int_{-\infty}^{\infty} \frac{u^2}{R} dt = \int_{-\infty}^{\infty} \frac{u^2}{R} dt \geq 0. \end{aligned}$$

Le résultat suivant rassemble diverses caractérisations des systèmes positifs.

**THÉORÈME 4.3 (Caractérisation des systèmes positifs).** Si on appelle  $H(s)$  la fonction de transfert  $D + C(s - A)^{-1}B$  du système linéaire (4.1) et si l'on note  $T = H + H^b$  on a (1)  $\dots$  (5) sont équivalents, avec :

1. le système (4.1) est positif,
2.  $\Re(H)(s) \geq 0 \forall s : \Re(s) \geq 0$ ,<sup>7</sup>
3.  $T(j\omega) \geq 0, \forall \omega$  dans  $\mathbb{R}$ ,
4.  $\exists F$  fraction rationnelle réelle:  $T = FF^b$ ,
5.  $\exists P \geq 0 : \begin{bmatrix} Q = -A'P - PA & S = C' - PB \\ S' & R = D + D' \end{bmatrix} \geq 0$  avec

$$\int_{t_1}^{t_2} y u dt = [x P x]_{t_1}^{t_2} + \int_{t_1}^{t_2} \begin{bmatrix} x' & u' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt.$$

**PREUVE.** • (3)  $\Rightarrow$  (2)

La matrice de transfert étant propre elle n'a pas de pôles à l'infini. Le système étant exponentiellement stable n'a pas de pôles dans le demi-plan droit au sens large (y compris l'axe imaginaire). La matrice de transfert est donc analytique dans le demi-plan droit.

Grâce au principe du maximum  $\Re(H)(s)$  est positive dans tout le demi-plan droit (ce qui est le résultat cherché) puisqu'elle est positive sur l'axe imaginaire, nulle à l'infini, analytique à l'intérieur du demi-plan droit.

<sup>6</sup>Dans cette partie un système sera noté par un quadruplet  $(A, B, C, D)$  lorsque  $D \neq 0$ .

<sup>7</sup> $\Re(s)$  note la partie réelle de  $s$ ,  $\Re(H)(s) \stackrel{\text{def}}{=} H(s) + H'(\bar{s})$ .

- (2)  $\Rightarrow$  (1)

Le théorème de Parseval nous donne

$$\begin{aligned}
 & \int_{-\infty}^{\infty} u'(t)y(t)dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} u'(j\omega)y(-j\omega)d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} u'(j\omega)H(-j\omega)u(-j\omega)d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} u'(j\omega)H'(j\omega)u(-j\omega)d\omega \\
 &= \frac{1}{4\pi} \int_{-\infty}^{\infty} u'(j\omega)(H'(j\omega) + H(-j\omega))u(-j\omega)d\omega .
 \end{aligned}$$

On en déduit le résultat.

- (1)  $\Rightarrow$  (3)

De la formule précédente on déduit le résultat. En effet s'il en était autrement  $T(j\omega)$  serait négative au moins en un point et donc dans un voisinage de ce point puisque cette fonction est analytique. Alors en choisissant  $u$  à transformée de Laplace à support dans ce voisinage on obtiendrait une contradiction avec l'hypothèse  $\int u'y \geq 0$ .

- (5)  $\Rightarrow$  (1)

$$\begin{aligned}
 0 &\leq \int_{t_1}^{t_2} (x' \ u') \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} dt \\
 &= \int_{t_1}^{t_2} [-x'(A'P + PA)x + 2x'(C' - PB)u + u'Ru]dt \\
 &= \int_{t_1}^{t_2} (-x'Px - \dot{x}'Px + 2x'C'u + u'Ru)dt \\
 &= \int_{t_1}^{t_2} (2x'C'u + u'Ru)ds - [x'Px]_{t_1}^{t_2} \\
 &= 2 \int_{t_1}^{t_2} y'uds - [x'Px]_{t_1}^{t_2} .
 \end{aligned}$$

Et donc

$$0 \leq \int_{-\infty}^{\infty} [x' \ u'] \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = 2 \int_{-\infty}^{\infty} y'udt ,$$

pout tout  $u$  à support compact .

- (1)  $\Rightarrow$  (5)

Soit  $u^*$  la commande transférant le système de 0 à  $x$ , minimisant  $\int_{-\infty}^0 u'ydt$ . Le coût optimal associé est quadratique, positif. Notons

$x'P^*x$  ce coût. Il satisfait l'équation de la programmation dynamique

$$(-Ax, 2P^*x) + \inf_u [(-Bu, 2P^*x) + (Cx + Du, u)] = 0, \quad \forall x,$$

qui se réécrit

$$\inf_u \begin{pmatrix} x' & u' \end{pmatrix} \begin{pmatrix} -A'P^* - P^*A & C' - P^*B \\ C - B'P^* & (D + D')/2 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} = 0, \quad \forall x,$$

et donc

$$\begin{pmatrix} -A'P^* - P^*A & C' - P^*B \\ C - B'P^* & (D + D')/2 \end{pmatrix} \geq 0.$$

- (4)  $\Rightarrow$  (3)

Si  $T$  se factorise en  $FF^b$  alors le théorème de Parseval nous donne

$$\begin{aligned} \int_{-\infty}^{\infty} u'y dt &= \frac{1}{4\pi} \int_{-\infty}^{\infty} u'(j\omega)(H(j\omega) + H'(-j\omega))u(-j\omega)d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \|F(j\omega)u(j\omega)\|^2 d\omega \geq 0. \end{aligned}$$

- (2)  $\Rightarrow$  (4) (dans le cas mono entrée mono sortie)

$H$  est rationnelle et s'écrit donc  $A(s)/B(s)$  alors :

$$\Re(T(j\omega)) = \frac{A(j\omega)B(-j\omega) + B(j\omega)A(-j\omega)}{2|B(j\omega)|^2}.$$

Le numérateur est un polynôme en les puissances paires de  $\omega$ , positif pour tout  $\omega$ , donc n'a pas de racines réelles simples, et donc se factorise sous la forme

$$\prod_k (s^2 + \alpha_k + j\beta_k)(s^2 + \alpha_k - j\beta_k) = K(s)K(-s),$$

et donc

$$T(s) = \frac{A(s)}{B(s)} + \frac{A(-s)}{B(-s)} = \frac{K(s)K(-s)}{B(s)B(-s)} = F(s)F(-s).$$

Dans le cas général on montrera (5)  $\Rightarrow$  (4) en généralisant le lemme de factorisation appliqué à  $P^*$ . □

Ce théorème est souvent appelé lemme positif réel.

Muni de ces caractérisations des systèmes positifs on peut énoncer un théorème de stabilité de systèmes nonlinéaires.

**THÉORÈME 4.4 (Popov).** Etant donné le système dynamique positif (4.1) commandé par  $u = -G(y)$ , alors quelque soit  $G$  telle que  $u'y \leq 0$  le système bouclé est asymptotiquement stable au sens de Lyapounov.

PREUVE. On démontre ici que la stabilité au sens de Lyapounov. On utilise le théorème de Lyapounov avec la fonction de Lyapounov  $V = x'Px$  où  $P$  satisfait le lemme positif réel

$$\begin{aligned}\dot{V} &= x'P\dot{x} + \dot{x}'Px, \\ &= x'P(Ax + Bu) + (u'B' + x'A')Px, \\ &= -x'Qx + x'(C' - S)u + u'(C - S')x, \\ &\leq -x'Qx - x'Su - u'S'x - u'Ru \text{ puisque } y'u \leq 0, \\ &\leq 0 \text{ (par le lemme positif réel)}. \end{aligned}$$

□

**COROLLAIRE 4.5 (Critère du cercle).** Dans le cas mono-entrée mono-sortie, si la fonction de transfert  $[L + KH(s)]/[l + kH(s)]$  est positive réelle pour  $k, l, K, L$  réels positifs  $k'$  ou  $k$  différents de zero  $k/l \leq K/L$  réels positifs alors le système nonlinéaire obtenu par un bouclage nonlinéaire  $y \mapsto u$  vérifiant  $(lu + ky)(Lu + Ky) \leq 0$  est asymptotiquement stable au sens de Lyapounov.

PREUVE. Le système entrée sortie  $v \mapsto z$  défini par

$$v = lu + ky, \quad z = Lu + Ky,$$

a pour fonction de transfert  $[L + KH(s)]/[l + kH(s)]$  si  $H$  désigne la fonction de transfert  $u \mapsto y$ . D'autre part  $vz \leq 0$  correspond à  $(lu + ky)(Lu + Ky) \leq 0$ . On est donc ramené au théorème de Popov pour le système  $v \mapsto z$ . □

**REMARQUE 4.6.** Il suffit de remarquer que la zone du plan complexe où  $\Re[(1 + KZ)(1 + kZ)] \geq 0$  est l'extérieur du cercle  $C$  de diamètre  $(-1/k, -1/K)$  pour comprendre que le critère du cercle dit que "si le lieu de Nyquist d'un système stable en boucle ouverte laisse sur sa gauche le cercle  $C$ , le système en boucle fermé obtenu par le feedback statique  $u = \Psi(y)$  est stable pourvu que le graphe de  $\Psi$  soit à l'intérieur du secteur défini par  $u = -cy, k \leq c \leq K$ ."

**EXERCICE 4.7.** On considère un système dont l'actuateur réalisant physiquement la commande peut "saturer" (c.a.d. ne peut pas dépasser une valeur donnée). Ce système est commandé par une commande proportionnelle (dans la zone de non saturation). Il est stable en boucle ouverte. Donnez une condition sur le lieu de Nyquist de sa fonction de transfert en boucle ouverte pour que le système en boucle fermé reste stable.

**PROPOSITION 4.8.** Si l'on appelle  $H(s) \stackrel{\text{def}}{=} \hat{B}'P(s - A)^{-1}\hat{B}$  avec  $\hat{B} = BR^{-1/2}$  la fonction de transfert du système optimal  $(A, B, I, Q, R)$ <sup>8</sup>, supposé commandable, le transfert  $H/(1 + H/2)$  qui correspond au système bouclé  $S = (A - \hat{B}\hat{B}'P/2, \hat{B}, \hat{B}'P)$  est positif.

<sup>8</sup>Système  $(A, B, I)$  optimisé au sens du critère  $(Q, R)$ .

PREUVE. L'équation de Riccati du problème LQ stationnaire et du gain optimal associé s'écrit :

$$A'P + PA - PBR^{-1}B'P + Q = 0, \quad K = R^{-1}B'P.$$

On peut donc écrire l'équation de Riccati sous la forme

$$(A - BK/2)'P + P(A - BK/2) + Q = 0. \quad (4.2)$$

Le système  $(A - BK/2, \hat{B}, \hat{B}'P) = (A - \hat{B}\hat{B}'P/2, \hat{B}, \hat{B}'P)$  est positif en effet  $A - BK/2$  est stable (car  $P$  est une fonction de Lyapounov acceptable grâce à (4.2)) et positif puisqu'on peut lui appliquer le lemme positif réel en utilisant  $P$  solution de (4.2) pour  $S = 0$ . Il suffit alors de remarquer que la fonction de transfert de  $(A - BK/2, \hat{B}, \hat{B}'P)$  est bien  $H/(1 + H/2)$  pour pouvoir conclure.  $\square$

**COROLLAIRE 4.9.** La méthode LQ conduit à des contrôleurs qui restent stables pour toute perturbation éventuellement nonlinéaire modifiant le gain optimal de façon multiplicative d'un facteur compris entre  $1/2$  et  $+\infty$ .

PREUVE. Il suffit d'appliquer le critère du cercle avec  $L = 0$ ,  $K = 1$ ,  $l = 1$  et  $k = 1/2$ .  $\square$

## 5. RAPPEL SUR LA STABILITÉ

Cette exposé s'inspire largement de [56]. Nous étudions la stabilité des systèmes dynamiques en temps continu:

$$\dot{x} = f(x, t) \quad x \in \mathbb{R}^n, t \in \mathbb{R}^+, \quad (5.1)$$

Après avoir rappelé les principales définitions. Nous donnons trois critères de stabilité. Les critères de Lyapounov, Nyquist et Routh.

### 5.1. DÉFINITIONS DE LA STABILITÉ

Nous aurons besoin de quatre types différents de stabilité. A chacune de ces notions voisines est attachée une définition précise. Nous supposons que  $f(t, 0) = 0 \quad \forall t$  dans tout ce paragraphe.

**DÉFINITION 5.1 (LS).** Nous dirons qu'un système dynamique est *stable au sens de Lyapounov ssi*:

$$\forall t_0, \epsilon \exists \eta : \|x(t_0)\| < \eta \Rightarrow \|x(t)\| \leq \epsilon \quad \forall t > t_0. \quad (5.2)$$

Cette définition signifie qu' "une faible perturbation de la condition initiale entraîne une faible perturbation de la trajectoire".

**DÉFINITION 5.2 (LAS).** Nous dirons qu'un système dynamique est *asymptotiquement stable au sens de Lyapounov si*

1. le système est LS,
2.  $\forall t_0 \exists R : \|x(t_0)\| < R \Rightarrow \lim_{t \rightarrow \infty} x(t) = 0$ .

La deuxième condition indique que le système retourne à 0.

**DÉFINITION 5.3 (ES).** Nous dirons qu'un système dynamique est exponentiellement stable si

$$\forall t_0, \exists k, c, R > 0 : \|x(t_0)\| < R \Rightarrow \|x(t)\| \leq ke^{-c(t-\tau)} \|x(\tau)\|, \quad \forall t > \tau \geq t_0.$$

La condition indique que le système retourne à 0 avec une vitesse exponentielle.

**DÉFINITION 5.4 (EBSB).** Nous dirons qu'un système entrée sortie

$$y(\cdot) = \mathcal{S}(u(\cdot))$$

est stable au sens entrée bornée sortie bornée si

$$\forall t_0, \epsilon, \exists \eta : \|u(t)\| < \eta, \quad \forall t > t_0 \Rightarrow \|x(t)\| \leq \epsilon, \quad \forall t > t_0. \quad (5.3)$$

## 5.2. MÉTHODE DE LYAPOUNOV

Elle consiste à trouver une fonction qui mesure la "distance" à l'origine que l'on appelle fonction de Lyapounov, qui décroît le long de la trajectoire du système. Cette "distance" doit vérifier un certain nombre de propriétés. Pour cela on introduit quelques définitions.

**DÉFINITION 5.5 (classe K).** Une fonction  $g$  de  $\mathbb{R}^+$  dans  $\mathbb{R}^+$  est dite de classe K, ce qui est noté  $g \in K$ , si  $g$  est continue, strictement croissante et vérifie  $g(0) = 0$ .

**DÉFINITION 5.6.** Pour une fonction  $V$  d'un espace normé dans  $\mathbb{R}^+$  et  $g \in K$  on dit que :

1.  $V$  est *définie positive* si  $g(\|x\|) \leq V(x)$ , et  $V(0) = 0$ .
2.  $V$  est *K-majorée* si  $V(x) \leq g(\|x\|)$ .
3.  $V$  est *radialement non bornée* si  $g(\|x\|) \leq V(x)$  et  $\lim_{y \rightarrow \infty} g(y) = \infty$ .

Etant donnée une fonction  $V(x, t)$  définie positive, K-majorée, radialement non bornée de  $\mathbb{R}^n \times \mathbb{R}^+ \mapsto \mathbb{R}^+$  on définit :

$$W(x, t) = \left( \frac{\partial V}{\partial t} + f \cdot \nabla V \right)(x, t).$$

On a alors :

$$\frac{dV}{dt}(x(t), t) = W(x(t), t).$$

Le résultat s'énonce :

**THÉORÈME 5.7 (Lyapounov).** Avec  $W$  défini par (5.2), pour que le système (5.1) soit

1. stable au sens de Lyapounov il suffit que :

$$\forall x, t \quad W(x, t) \leq 0 ;$$

2. asymptotiquement stable au sens de Lyapounov il suffit que :

$$\forall x, t \quad W(x, t) \leq -\gamma(\|x\|) .$$

pour  $\gamma \in K$ .

Dans le cas des systèmes linéaires stationnaires le théorème de Lyapounov se particularise.

**COROLLAIRE 5.8.** Si les deux matrices  $P$  et  $Q$  symétriques définies positives vérifient l'équation de Lyapounov:

$$F'P + PF = -Q \quad (5.4)$$

le système :

$$\dot{x} = Fx \quad (5.5)$$

est asymptotiquement stable au sens de Lyapounov.

Inversement si (5.5) est exponentiellement stable alors pour toute matrice  $Q$  (5.4) a une solution unique.

**PREUVE.** Pour démontrer ce résultat dans le sens direct on utilise la fonction de Lyapounov :

$$V(x) = x'Px ,$$

et puisque  $Q$  est définie positive

$$\exists \delta : \delta \|x\|^2 \leq x'Qx \quad \forall x .$$

En définissant la fonction  $\gamma \in K$  par  $\gamma : y \in \mathbb{R}^+ \mapsto \delta y^2 \in \mathbb{R}^+$  on a

$$W(x) = x'(PF + F'P)x ,$$

et donc

$$W(x) \leq -x'Qx \leq -\gamma(\|x\|) .$$

Inversement si (5.5) est ES alors  $x(t) \rightarrow 0$  de façon exponentielle lorsque  $t \rightarrow \infty$  et donc

$$P = \int_0^\infty e^{F't} Q e^{Ft} dt ,$$

existe.

D'autre part,

$$F'P + PF = \int_0^\infty (F' e^{F't} Q e^{Ft} + e^{F't} Q e^{Ft} F) dt , \quad (5.6)$$

$$= \int_0^\infty \frac{d(e^{F't} Q e^{Ft})}{dt} dt , \quad (5.7)$$

$$= -Q . \quad (5.8)$$

Il nous reste à démontrer l'unicité de la solution de (5.4) sous la condition de stabilité de (5.5). S'il en était autrement l'équation de Sylvester  $F'P + PF = 0$  aurait une solution non nulle ce qui ne se produit que lorsque les spectres de  $F$  et  $-F'$  ont une intersection non nulle (par exemple dans le cas  $P$  inversible  $F'P + PF = 0$  se réécrit  $P^{-1}FP = -F'$ ). Cela ne peut arriver que si le système linéaire a des valeurs propres de part et d'autre de l'axe imaginaire et donc si le système est instable.  $\square$

## 5.3. CRITÈRE DE NYQUIST

Le critère de Nyquist donne une condition de stabilité, dans le domaine fréquentiel, des systèmes linéaires stationnaires bouclés par une rétroaction unitaire.

**THÉORÈME 5.9 (Nyquist).** Un système à contre réaction unitaire donc de fonction de transfert  $H/(1 + H)$  est stable ssi l'image par  $H(s)$  (fonction de transfert du système direct c.a.d. sans la contre réaction) du contour du demi-plan droit du plan complexe orienté dans le sens inverse, noté  $\mathcal{C}$ , entoure le point  $-1$ , dans le sens direct, un nombre de fois égal au nombre de pôles instables du système direct.

PREUVE. On note :

$$T(s) = \frac{H(s)}{1 + H(s)}$$

la fonction de transfert du système bouclé.  $T$  sera stable ssi elle n'a aucun pôle dans le demi-plan droit du plan complexe. Or

$$\arg[1 + H(s)]_{\mathcal{C}} = 2\pi(P - Z)$$

où on désigne par  $P$  le nombre de pôles de  $1 + H$  dans le demi-plan droit, et  $Z$  le nombre de zéros de  $1 + H$  dans le demi-plan droit.

Si  $T$  est stable on a

- 0 est égal au nbre de pôles dans le demi-plan droit de  $T$ ,
- qui est égal nbre de zéros de  $1 + H$  dans le demi-plan droit,
- qui est égal nbre de pôles de  $1 + H$  dans le demi-plan droit, moins le nbre de fois que  $H(\mathcal{C})$  entoure  $-1$  dans le sens direct.

Et donc le nbre de pôles de  $H$  doit être égal au nbre de fois que  $H(\mathcal{C})$  entoure  $-1$  dans le sens direct.  $\square$

On en déduit le corollaire suivant.

**COROLLAIRE 5.10.** Si le système direct est stable, le système bouclé est stable ssi le lieu de Nyquist c.a.d  $H(\mathcal{C})$  n'entoure pas  $-1$  dans le sens direct.

## 5.4. CRITÈRE DE ROUTH

Il suffit de savoir déterminer le nombre de racines à partie réelle positive d'un polynôme pour pouvoir disposer d'un test de stabilité d'un système linéaire stationnaire. C'est le critère de Routh.

Avant de donner le critère de Routh nous allons introduire des notions qui ont un intérêt en elle même et qui servent à démontrer la validité du critère — l'indice de Cauchy d'une fonction, — les suites de Sturm de polynômes. Nous suivons l'exposé de Gantamacher [58].

**DÉFINITION 5.11 (Indice de Cauchy).** L'indice de Cauchy d'une fraction rationnelle  $R(s)$  entre  $a$  et  $b$ , noté  $I_a^b R(s)$ , est la différence entre le nombre de sauts de  $-\infty$  à  $+\infty$  et le nombre de sauts de  $+\infty$  à  $-\infty$  lorsque  $x$  varie de  $a$  à  $b$ .

**PROPOSITION 5.12.** Si  $P(s)$  est un polynôme  $I_a^b(P'/P)$  est le nombre de racines réelles de  $P(s)$  dans l'intervalle  $(a,b)$ .

PREUVE.

$$\frac{P'(s)}{P(s)} = R(s) + \sum_{j=1}^m \frac{n_j}{s - s_j},$$

où —  $R$  est sans pôles réels, —  $s_j$  sont les racines réelles de  $P$  avec la multiplicité  $n_j$ . Le résultat en découle immédiatement.  $\square$

**DÉFINITION 5.13 (Suite de Sturm).** Une suite de Sturm dans l'intervalle  $(a, b)$  est une suite de polynômes réels  $P_1(s), \dots, P_m(s)$  vérifiant

1.  $\forall x \in (a, b), \forall k > 0 \quad P_k(s) = 0 \Rightarrow P_{k-1}(s)P_k(s) < 0$  ;
2.  $P_m(s) \neq 0 \quad \forall s \in (a, b)$  .

EXEMPLE 5.14. 1. Suite des dérivées successives d'un polynôme.  
2. Suite des restes de divisions successives de polynômes.

$P_1$  et  $P_2$  étant deux polynômes premiers entre eux, on définit la suite de Sturm par

$$P_i = -\text{reste}(P_{i-1}, P_{i-2}) \quad \forall i, i \geq 2, i \leq m .$$

Cette suite est bien une suite de Sturm puisque

$$\exists Q_i : P_i = Q_i P_{i+1} - P_{i+2} .$$

Alors

$$P_{i+1}(s) = 0 \Rightarrow P_i(s)P_{i+2}(s) = -P_{i+2}^2(s) \leq 0 .$$

Les polynômes  $P_1$  et  $P_2$  étant premiers entre eux  $P_{i+1}$  et  $P_{i+2}$  n'ont pas de racines communes et donc  $P_i(s)P_{i+2}(s) < 0$ .

Le théorème suivant établit un lien entre l'indice de Cauchy et le nombre de changements de signe en un point  $s$ , noté  $V(s)$ , d'une suite de Sturm.

**THÉORÈME 5.15 (Sturm).** Si  $P_1, \dots, P_m$  désigne une suite de Sturm dans l'intervalle  $(a, b)$  on a :

$$I_a^b\left(\frac{P_2}{P_1}\right) = V(a) - V(b) .$$

PREUVE. Pour  $i > 1$ ,  $s$  variant dans l'intervalle  $(a, b)$  lorsque  $P_i(s)$  s'annule, puisque  $P_{k-1}(s)P_{k+1}(s) < 0$  le nbre de changements de signe de la suite ne varie pas.

La variation du nbre de changements de signe ne provient donc que des changements de signe de  $P_1$ , donc des passages par l'infini de  $P_2/P_1$ .

Lorsque  $P_2/P_1$  passe de  $-\infty$  à  $\infty$  cela signifie que  $P_1$  et  $P_2$  étaient d'abord de signe opposé puis de même signe il y a donc perte d'un changement de signe, de façon analogue il y a gain de changement de signe dans l'autre situation. On en déduit alors facilement le théorème.  $\square$

**PROPOSITION 5.16.** Si  $P$  est un polynôme réel sans racine sur l'axe imaginaire

$$\arg_{-\infty}^{\infty} P(i\omega) = (n - 2k)\pi ,$$

où —  $n$  désigne le degré de  $P$  —  $k$  le nbre de racines de  $P$  dans le demi-plan droit.

PREUVE. On utilise le théorème de Cauchy le long du contour de Nyquist  $\mathcal{C}$  (bord demi-plan droit dans le sens retrograde). On obtient

$$\arg_{-i\infty}^{i\infty} P + \arg_{R=\infty} P = -2\pi k .$$

Mais

$$\arg_{R=\infty} P = \arg_{R=\infty} s^n = -n\pi$$

d'où le résultat. □

Etant donné un polynôme  $P(s)$  de degré  $n$

$$P(s) = a_0 s^n + b_0 s^{n-1} + a_1 s^{n-2} + b_1 s^{n-3} + \dots ,$$

on définit  $Q$  et  $R$  par

$$P(i\omega) = Q(\omega) + iR(\omega) .$$

On peut alors utiliser  $\arg(P(i\omega)) = \arctan R(\omega)/Q(\omega)$  pour montrer

$$\arg_{-\infty}^{\infty} P(i\omega) = I_{-\infty}^{\infty} \frac{T(\omega)}{S(\omega)} ,$$

avec

$$\begin{aligned} S(\omega) &= a_0 \omega^n - a_1 \omega^{n-2} + \dots , \\ T(\omega) &= b_0 \omega^{n-1} - b_1 \omega^{n-3} + \dots . \end{aligned}$$

Pour cela il faut discuter selon la parité de  $n$ .

On en déduit alors, en utilisant la proposition 5.16, le résultat suivant.

**PROPOSITION 5.17.**

$$I_{-\infty}^{\infty} \frac{T(\omega)}{S(\omega)} = n - 2k .$$

On peut énoncer le théorème de Routh.

**THÉORÈME 5.18 (Routh).** Le nombre de variations de signe en  $\infty$ , noté  $V(\infty)$  de la suite de Sturm

$$P_1 = S, \quad P_2 = T, \quad P_{i+1} = -\text{reste}(P_{i-1}, P_i) ,$$

est égal au nombre de racines de  $P$  dans le demi-plan droit, noté  $k$

$$V(\infty) = k .$$

**PREUVE.** On applique le théorème de Sturm à la suite de Sturm  $P_i$  dans l'intervalle  $(-\infty, \infty)$ . On a donc

$$V(-\infty) - V(\infty) = n - 2k .$$

Mais

$$V(-\infty) = n - V(\infty) ,$$

car la suite des degrés  $(P_i)$  est  $(n, \dots, 0)$  dans le cas régulier où aucun coefficient de tête ne s'annule. □

On définit le tableau de Routh construit sur  $P$

**DÉFINITION 5.19 (Tableau de Routh).**

$$\begin{array}{cccc} a_0 & a_1 & a_2 & \dots \\ b_0 & b_1 & b_2 & \dots \\ c_0 & c_1 & c_2 & \dots \\ d_0 & d_1 & d_2 & \dots \\ \vdots & \vdots & \vdots & \end{array}$$

avec

$$\begin{aligned} c_i &= a_{i+1} - b_{i+1} \frac{a_0}{b_0} , \\ d_i &= b_{i+1} - c_{i+1} \frac{b_0}{c_0} . \end{aligned}$$

Le tableau de Routh est dit *régulier* lorsque les coefficients d'indice 0 ne s'annulent pas au cours de la construction du tableau.

On a alors le résultat le plus important de ce paragraphe.

**COROLLAIRE 5.20 (Critère de Routh).** Dans le cas régulier, le nombre de changements de signe de la première colonne du tableau de Routh construit sur le polynôme  $P$  est égal au nombre de racines dans le demi-plan droit de  $P$ .

**PREUVE.** Il suffit de remarquer que les signes de la première colonne du tableau de Routh donnent les signes des polynômes de la suite de Sturm construite sur  $P$  en  $\infty$ , puisque — la construction du tableau de Routh revient à calculer la suite de Sturm (5.18), — les coefficients d'indice 0 sont les coefficients des monômes de plus haut degré.  $\square$

**EXEMPLE 5.21.** 1. Le tableau de Routh de  $P(s) = (s-1)^3 = s^3 - 3s^2 + 3s - 1$  s'écrit

$$\begin{array}{cc} 1 & 3 \\ -3 & -1 \\ 8/3 & \\ -1 & \end{array}$$

La première colonne présente trois changements de signe, ce qui est le résultat souhaité puisque  $P$  a trois fois la racine 1 qui est dans le demi-plan droit.

2. Le tableau de Routh de  $P(s) = (s+1)^3 = s^3 + 3s^2 + 3s + 1$  s'écrit

$$\begin{array}{cc} 1 & 3 \\ 3 & 1 \\ 8/3 & \\ 1 & \end{array}$$

La première colonne n'a pas de changements de signe.  $P$  n'a pas de racines dans le demi-plan droit.

## CHAPITRE 8

### PROBLÈMES

#### 1. UNE GESTION DE STOCK

##### 1.1. ENONCÉ

On considère la gestion, en temps discret, d'un stock d'un produit soumis à des demandes aléatoires notées  $v_t$  (indépendantes et de mêmes lois)  $P(dv)$  où  $t$  représente la période de temps considérée.

L'état du stock est noté  $X_t$  il représente respectivement : — la quantité en stock si  $X_t \geq 0$ , — moins la demande cumulée non satisfaite si  $X_t < 0$  (appelée *défaillance d'approvisionnement*).

Le stock est alimenté par des commandes du produit faites éventuellement à chaque période de temps. La quantité commandée faite au début d'une période  $t$  est reçue avant la fin de cette période. Elle est notée  $U_t \geq 0$ .

On considère deux coûts associés au fonctionnement du système sur une période de temps

1. le *coût de stockage et de défaillance* noté  $f(x)$ 
  - (a) lorsque  $x \geq 0$ ,  $f(x)$  est interprété comme le coût de stockage de la quantité  $x$ ,
  - (b) lorsque  $x < 0$ ,  $f(x)$  est interprété comme le coût de défaillance d'approvisionnement de niveau  $-x$ ;
2. le *coût de commande* noté  $g(u)$  (défini sur  $u \geq 0$ ) est supposé concave croissant nul en 0.

1.1.1. QUESTION 1. Formulez le problème de gestion précédent sur  $T$  périodes de temps en terme de commande optimale stochastique.

Définissez la fonction de la programmation dynamique associée.

Donnez l'équation récurrente en temps satisfaite par cette fonction (équation de la programmation dynamique).

1.1.2. QUESTION 2. On considère l'équation suivante :

$$(1 + \lambda)V(x) = \min_{u \geq 0} \left\{ \int_v V(x + u - v)P(dv) + g(u) \right\} + f(x) ,$$

où  $\lambda$  désigne un nombre strictement positif. De quel problème de contrôle stochastique cette équation est elle l'équation de la programmation dynamique?

Expliquez la méthode de gestion du stock déduite de la résolution de cette équation.

Supposons  $f$  et  $g$  bornées et que l'on puisse en déduire que  $V$  soit bornée, démontrez que le commande solution de cette équation est optimale (c.a.d. qu'une autre décision conduirait à un coût supérieur).

1.1.3. QUESTION 3. On se place dans la situation de la première question. On fait tendre vers l'infini le nombre de périodes de gestion  $T$  on aimerait comprendre le comportement asymptotique de la fonction de la programmation dynamique. La solution de l'équation de la programmation dynamique admet des solutions de la forme comme  $-\mu T + W(x)$  pour une certaine condition finale, où  $\mu$  est un nombre réel positif.  $\mu$  s'interprète comme le coût moyen par période. On pourra numéroter les périodes par des nombres négatifs. On numérotera 0 la dernière période.

Quelle est l'équation que doit satisfaire le couple  $(\mu, W)$ ?

Quelle est la condition finale qui convient?

Ce couple est-il défini de façon unique?

1.1.4. QUESTION 4. On considère qu'une période de temps a une durée  $h$ .  $\mu$  s'interprète ici comme un coût par unité de temps.

A quel cas particulier de l'équation précédente correspond l'équation:

$$h\mu = \min \left\{ \inf_{u>0} [W(x+u-h)] - W(x) + 1, W(x-h) - W(x) \right\} + hx^2 ?$$

Ecrivez l'équation limite lorsque  $h \rightarrow 0$ .

Résolvez explicitement cette dernière équation dans la classe des  $W$  continues.

Vérifiez que la solution  $W$  correspondante est  $C^1$ .

Résumez en une phrase la stratégie obtenue.

## 1.2. CORRIGÉ

1.2.1. QUESTION 1.

1. On veut résoudre :

$$\min_{\substack{U_t \\ t=0, \dots, T-1}} \mathbb{E} \sum_{t=0}^{T-1} [f(X_t) + g(U_t)],$$

sous les contraintes dynamiques:

$$X_{t+1} = X_t + U_t - V_t, \quad t = 0, \dots, T-1.$$

2. La fonction de la programmation dynamique est définie par :

$$V(t, x) = \min_{\substack{U_s \\ s=t, \dots, T-1}} \mathbb{E} \left\{ \sum_{s=t}^{T-1} [f(X_s) + g(U_s)] \mid X_t = x \right\}.$$

3. L'équation de la programmation dynamique s'écrit alors :

$$V(t, x) = \min_{u \in R^+} \left[ \int_v V(t+1, x+u-v) P(dv) + f(u) + g(x) \right],$$

$$V(T, x) = 0.$$

### 1.2.2. QUESTION 2.

1. L'équation :

$$(1 + \lambda)V(x) = \min_{u \geq 0} \left\{ \int_v V(x+u-v) P(dv) + g(u) \right\} + f(x),$$

est l'équation de la programmation dynamique du problème de commande stochastique en horizon infini avec coût actualisé suivant :

$$\min_{\substack{U_t \\ t=0, \dots, \infty}} \mathbb{E} \sum_{t=0}^{\infty} \frac{1}{(1 + \lambda)^{t+1}} [f(X_t) + g(U_t)],$$

sous les contraintes dynamiques :

$$X_{t+1} = X_t + U_t - V_t, \quad t = 0, \dots, \infty.$$

2. La méthode de gestion consiste en :

- la résolution de l'équation de la programmation dynamique précédente qui nous donne le coût optimal  $V$  et la stratégie de décision markovienne :

$$s : x \mapsto u \in \arg \min_{u \geq 0} \left[ \int_v V(x+u-v) P(dv) + g(u) \right],$$

- l'utilisation de la stratégie optimale en commandant à chaque période la quantité  $s(X_t)$  si le stock à l'instant  $t$  est  $X_t$ .

3. Pour montrer que cette stratégie est optimale on considère une autre stratégie  $Z_{t,\omega}$  ne dépendant que du passé  $(X_s)_{s \leq t}$  des observations (ici l'état du système). Etudions alors, l'évolution de  $1/(1 + \lambda)^t V(X_t^Z)$  où  $V$  est la solution de l'équation de la programmation dynamique et  $X^Z$  désigne la niveau de stock en supposant que le système évolue avec la politique  $Z$ . On a :

$$\begin{aligned} & \mathbb{E} \left\{ \frac{1}{(1 + \lambda)^{t+1}} V(X_{t+1}^Z) - \frac{1}{(1 + \lambda)^t} V(X_t^Z) \mid X_s^Z, s \leq t \right\} \\ &= \frac{1}{(1 + \lambda)^{t+1}} \left[ \int_v V(X_t^Z + Z_t - v) P(dv) - \frac{1}{(1 + \lambda)^{t+1}} V(X_t^Z) \right] \\ &\geq \frac{1}{(1 + \lambda)^{t+1}} [g(Z_t) + f(X_t^Z)], \end{aligned}$$

grâce successivement à la définition de l'espérance conditionnelle puis à l'équation satisfaite par  $V$ .

En sommant sur  $t$  allant de 0 à l'infini ces inégalités et en utilisant le fait que  $V$  est bornée et en prenant l'espérance du résultat on obtient :

$$V(x) \leq \mathbb{E} \left\{ \sum_{t=0}^{\infty} \frac{1}{(1+\lambda)^{t+1}} [f(X_t^Z) + g(Z_t)] \mid X_0^Z = x \right\} .$$

Le même raisonnement montre que l'égalité est atteinte pour la stratégie optimale.

### 1.2.3. QUESTION 3.

1. On substitue  $V(t, x)$  par  $-\mu T + W(x)$  dans l'équation de la programmation dynamique de la question 1, on obtient :

$$\mu + W(x) = \min_{u \geq 0} \left[ \int_v W(x + u - v) P(dv) + g(u) \right] + f(x) .$$

En prenant pour condition finale  $V(0, x) = W(x)$  on vérifie par récurrence arrière que  $-\mu T + W(x)$  est solution.

2.  $W$  est définie à une constante près, en effet si  $W$  est solution,  $W + k$ , où  $k$  est une constante quelconque, est encore une solution.

### 1.2.4. QUESTION 4.

1. L'équation correspond au cas particulier :

- $P(dv) = \delta_h$  ou  $\delta_h$  désigne la masse de Dirac en  $h$ ;
- $f(x) = hx^2$ ;
- $g(u) = 0$  si  $u \neq 0$ ,  $g(u) = 1$  si  $u = 0$ ;
- on a également changé  $\mu$  en  $h\mu$ .

2. L'équation limite est :

$$0 = \min \left\{ \inf_{u > 0} W(x + u) - W(x) + 1, -W'(x) + x^2 - \mu \right\} .$$

3. La solution comporte deux régions. Dans la première région l'équation

$$-W'(x) + x^2 - \mu = 0 ,$$

est satisfaite ( $W$  étant défini au moins à une constante près on impose  $W(0) = 0$ ). On admet que cette région est connexe. On a alors dans cette zone :

$$W(x) = x^3/3 - \mu x .$$

$\mu$  étant positif cette fonction admet un minimum local en  $x = \mu^{1/2}$  (la valeur du minimum étant  $m = -(2/3)\mu^{3/2}$ ). Alors il y a trois possibilités

- $2|m| \leq 1$  il n'existe pas de solutions continues;

- $2|m| = 1$  alors

$$W(x) = \begin{cases} |m| & \text{pour } x \leq s = -\mu^{1/2}, \\ x^3/3 - \mu x & \text{ailleurs,} \end{cases}$$

est continue et convient donc;

- $2|m| > 1$  alors les candidats possibles sont

$$W(x) = \begin{cases} |m| & \text{pour } x \leq s, \\ x^3/3 - \mu x & \text{ailleurs,} \end{cases}$$

avec  $s \leq S$  à déterminer. Mais il y a toujours une zone où l'équation n'est pas satisfaite et donc ces candidats ne conviennent pas.

Dans le seul cas possible, c.a.d le deuxième cas,  $\mu$  est déterminée par l'équation:  $-2m = 1$  et donc  $\mu = (3/4)^{2/3}$ .

4. Le raccord entre les deux zones a lieu sur le maximum local de  $x^3/3 - \mu x$ . Puisque avant ce maximum la fonction est constante le raccord est bien  $C^1$ .
5. On a ainsi obtenu une politique  $(s, S)$ . Lorsque le stock descend en dessous de  $s$  on commande de façon à se ramener à  $S$ .

## 2. MAINTENANCE D'UNE AUTOMOBILE

### 2.1. ENONCÉ

2.1.1. INTRODUCTION. On considère la maintenance, en temps discret, d'une automobile dont l'état de fonctionnement est modélisé par une chaîne de Markov à états en nombre fini. Dans une première partie on étudie l'évolution de l'état pour une politique de maintenance donnée. Dans une deuxième partie on optimise la politique de maintenance.

On modélise l'évolution de l'état de notre automobile par une chaîne de Markov. Chaque état représente le niveau de dégradation. L'état de fonctionnement parfait est noté 0. L'état d'épave inutilisable est noté  $E$ . Les états intermédiaires sont représentés par un entier compris entre 0 et  $E$ . La probabilité de passer de l'état  $x$  à l'état  $x + 1$  pour  $x = 0, 1, \dots, E - 1$  est supposée être indépendante de  $x$  et égale à  $\lambda$  que l'on appelle taux de dégradation. Les transitions vers les autres états sont considérées comme impossibles.

Une politique de maintenance consiste à décider des instants où l'on va chez le garagiste. On décide d'un niveau de dégradation inacceptable qui lorsqu'il est atteint entraîne la réparation. On suppose qu'après chaque réparation (qui est instantanée) la voiture revient dans l'état de fonctionnement parfait. On saute donc de l'état qui précède l'état inacceptable à l'état neuf avec la probabilité  $\lambda$ .

Dans la première partie on considère la politique de maintenance où l'état inacceptable est l'état d'épave. On utilise la voiture sans aller chez le garagiste jusqu'à ce qu'elle soit devenue une épave (état  $E$  que l'on

suppose être capable d'observer). On remplace alors la voiture instantanément et on recommence avec cette dernière, supposée avoir le même taux de dégradation que la voiture précédente. On qualifie de naïve cette politique de maintenance. Dans la deuxième partie on améliore cette politique en allant éventuellement plus tôt chez le garagiste.

On considère les coûts suivants :

1. un coût  $c_x$  associé au mauvais fonctionnement de la voiture, dépendant de son niveau de dégradation  $x$ , que l'on paie tous les jours;
2. un coût de réparation  $k_x$  qui est la somme d'un coût fixe et d'un coût dépendant du niveau de dégradation du véhicule (ce coût est payé chaque fois que l'on a recours au service d'un garagiste; lorsqu'on va chez le garagiste on néglige le coût de mauvais fonctionnement du véhicule devant celui de réparation).

### 2.1.2. ETUDE DE LA CHAÎNE DE MARKOV ASSOCIÉE À LA POLITIQUE NAÏVE.

1. Donnez la matrice de transition de la chaîne de Markov représentant l'évolution de l'état la voiture sans politique de maintenance puis avec la politique de maintenance naïve. On appelle  $M$  cette dernière matrice. Dans la suite de ce paragraphe on est toujours dans cette situation.
2. Calculez explicitement la probabilité  $p_x^n$  d'être dans l'état  $x$  à l'instant  $n$ . A l'instant 0 la voiture est supposée neuve donc dans l'état 0 avec une probabilité 1.
3. Montrez que  $p^n$  converge vers une unique mesure d'équilibre que l'on déterminera.
4. Déduisez des questions précédentes la formule :

$$1/E = \lim_{n \rightarrow \infty} \sum_{\substack{k=0 \\ [n-l/E]}}^{(kE+l)} C_n^{(kE+l)} \lambda^{kE+l} (1-\lambda)^{n-kE-l}$$

pour  $l = 0, 1, \dots, E-1$  où  $C_m^l$  désigne le nombre de combinaisons de  $l$  éléments parmi  $m$  et  $[a]$  la partie entière de  $a$ .

### 2.1.3. OPTIMISATION DE LA POLITIQUE DE MAINTENANCE.

1. Formulez en terme de commande optimale stochastique le problème d'entretien optimal dans le cas d'un horizon infini avec un taux d'actualisation  $\mu$ .
2. Donnez explicitement (le système d'équations scalaires) l'équation de la programmation dynamique à résoudre permettant de calculer la politique d'entretien optimale.
3. Donnez un algorithme de résolution de ce système d'équations.
4. Dans le cas  $c_x = x$ ,  $k_x = m$  ramenez la résolution de l'équation de la programmation dynamique à un problème d'optimisation scalaire

(à expliciter) dans lequel le paramètre à optimiser est l'état de dégradation au dessus duquel on va chez le garagiste.

5. Etudiez l'asymptotique  $\mu \rightarrow 0$ .

## 2.2. CORRIGÉ

### 2.2.1. ETUDE DE LA CHAÎNE DE MARKOV ASSOCIÉE À LA POLITIQUE NAÏVE.

1. La matrice de transition de la chaîne de Markov représentant l'évolution de l'automobile sans politique de maintenance est la matrice  $(E + 1) \times (E + 1)$  suivante :

$$\begin{pmatrix} 1 - \lambda & \lambda & 0 & \cdot & \cdot \\ 0 & 1 - \lambda & \lambda & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 1 - \lambda & \lambda \\ 0 & \cdot & \cdot & 0 & 1 \end{pmatrix}.$$

La matrice de transition de la chaîne de Markov représentant l'évolution de l'automobile lorsque la politique de maintenance naïve est utilisé est la matrice  $E \times E$  :

$$M = \begin{pmatrix} 1 - \lambda & \lambda & 0 & \cdot & \cdot \\ 0 & 1 - \lambda & \lambda & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 1 - \lambda & \lambda \\ \lambda & 0 & \cdot & 0 & 1 - \lambda \end{pmatrix}.$$

2. Notant  $P$  la matrice associée à la permutation

$$0 \rightarrow 1 \rightarrow 2 \rightarrow \dots (E - 1) \rightarrow 0,$$

c.a.d.

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot \\ 0 & 0 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & 1 \\ 1 & 0 & \cdot & \cdot & 0 \end{pmatrix},$$

on a :

$$M = (1 - \lambda)I + \lambda P,$$

où  $I$  désigne l'identité. En utilisant l'équation de Fokker-Planck on obtient :

$$p^n = \delta M^n$$

où :

$$\delta = (1 \ 0 \ \dots \ 0).$$

En utilisant alors la commutativité de  $I$  et  $P$  on obtient :

$$p_x^n = \left[ \sum_{k=0}^n C_n^k (1 - \lambda)^{n-k} \lambda^k P^k \right]_{0x}.$$

En utilisant la propriété  $P_{0x}^k = 1$  si  $x = k$  modulo  $E$ ,  $P_{0x}^k = 0$  sinon, on obtient

$$p_x^n = \sum_{k=0}^{\lfloor (n-x)/E \rfloor} C_n^{(kE+x)} (1-\lambda)^{n-kE-x} \lambda^{kE+x}.$$

3. Les valeurs propres de la matrice  $M$  sont

$$(1-\lambda) + \lambda e^{k2j\pi/E}$$

en utilisant les propriétés spectrales des matrices de permutation. La matrice  $M$  est diagonalisable et admet la valeur propre 1 avec la multiplicité 1. Les autres valeurs propres ont des modules strictement inférieurs à 1. On voit, en se plaçant dans la base propre, que  $M^n$  converge vers le projecteur spectral sur le vecteur propre associé à la valeur propre 1. Or ce vecteur propre est  $p = (1/E \ \cdots \ 1/E)$ .  $p^n$  converge donc vers l'unique mesure de probabilité vérifiant  $q = qM$  c.a.d.  $p$ .

4. En faisant  $x = l$  dans la formule donnant explicitement  $p_x^n$  le résultat de convergence de la question précédente donne précisément la formule cherchée.

### 2.2.2. OPTIMISATION DE LA POLITIQUE DE MAINTENANCE.

1. La commande  $u$  peut prendre deux valeurs  $u = 0$  signifie pas d'entretien,  $u = 1$  signifie un entretien. La matrice de transition de la chaîne de Markov contrôlée  $X_n$  est alors

$$M^u = \begin{pmatrix} 1-\lambda+u\lambda & (1-u)\lambda & 0 & \cdot & \cdot \\ u\lambda & 1-\lambda & (1-u)\lambda & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u\lambda & 0 & \cdot & 1-\lambda & (1-u)\lambda \\ \lambda & 0 & \cdot & 0 & 1-\lambda \end{pmatrix}.$$

Notant  $d_x^u = k_x$  si  $u = 1$  et 0 sinon, on veut résoudre

$$\min_U \mathbb{E} \sum_{i=0}^{\infty} \frac{1}{(1+\mu)^{i+1}} (c_{X_n} + d_{X_n}^{U_n}).$$

2. Pour résoudre ce problème on introduit la fonction de la programmation dynamique

$$v_x = \min_U \mathbb{E} \left\{ \sum_{i=0}^{\infty} \frac{1}{(1+\mu)^{i+1}} [c_{X_n} + d_{X_n}^{U_n}] \mid X_0 = x \right\}.$$

$v$  vérifie l'équation de la programmation dynamique

$$0 = \min_{u \in \{0,1\}} [(M^u - (1+\mu)I)v + d^u] + c.$$

Explicitant cette équation de la programmation dynamique, on obtient :

$$(\lambda + \mu)v_x = \min \{ \lambda v_{x+1} + c_x, \lambda v_0 + k_x \}, \quad \forall x = 0, \dots, E-2.$$

$$(\lambda + \mu)v_{E-1} = \lambda v_0 + k_{E-1} .$$

3. On peut résoudre ce système d'équations non linéaires par l'algorithme de Howard qui consiste à faire l'itération suivante : partant du feedback  $s$ , fixant pour chaque état la décision d'entretien, on améliore la stratégie (au sens de la diminution du coût à optimiser) en résolvant, dans un premier temps, pour tout  $x = 0, \dots, E - 2$  :

$$(\lambda + \mu)w_x = \lambda w_{x+1} + c_x, \text{ si } s_x = 0 ,$$

$$(\lambda + \mu)w_x = \lambda w_0 + k_x, \text{ si } s_x = 1 ,$$

$$(\lambda + \mu)w_{E-1} = \lambda w_0 + k_{E-1} ,$$

puis en calculant une nouvelle stratégie en résolvant

$$\min\{\lambda w_{x+1} + c_x, \lambda w_0 + k_x\} .$$

L'application qui à chaque  $x$  donne la commande réalisant le minimum définit la nouvelle stratégie.

4. Par la physique du problème il est clair que la stratégie optimale prend la valeur 1 au dessus d'un niveau de dégradation  $l$  et 0 en dessous. Pour une telle stratégie on peut évaluer le coût moyen actualisé en résolvant l'équation de Kolmogorov sur la zone où on n'entretient pas

$$(\lambda + \mu)v_x = \lambda v_{x+1} + x, \quad \forall x < l ,$$

$$(\lambda + \mu)v_l = \lambda v_0 + m .$$

Ce système s'élimine facilement pour se ramener à une équation linéaire

$$v_k = v_0 \rho^k - \Phi(k)/\lambda ,$$

avec

$$\rho = (\lambda + \mu)/\lambda ,$$

et

$$\Phi(l) = \sum_{i=1}^{l-1} i \rho^{l-1-i} ,$$

et donc

$$v_0 = \frac{m}{\rho} + \frac{\Phi(l)}{\lambda(\rho^l - \frac{1}{\rho})} .$$

Il reste alors à optimiser cette fonction de  $l$  par rapport à  $l$  (qui paramétrise la stratégie) pour avoir résolu notre problème.

5. Lorsque  $\mu$  tend vers 0,  $\rho^{l+1} - 1$  se comporte comme  $(\rho - 1)(l + 1)$  et  $\phi(l)$  comme  $l(l - 1)/2$ . En approximant  $l - 1$  et  $l + 1$  par  $l$  on obtient la formule simplifiée  $l \simeq \sqrt{2m}$ .

### 3. TRANSPORT

On considère un problème de gestion de voitures de location. On suppose disposer de deux parkings dans lesquels les voitures peuvent être louées et laissées après usage. Le contrôle s'exerce uniquement par des transports de voitures d'un parking à l'autre. On étudie un système dans une version extrêmement simplifiée. Certaines parties de l'analyse proposée ci dessous s'étendent à des situations plus compliquées. Les problèmes de transport, un peu réaliste, restent très souvent des problèmes ouverts.

#### 3.1. ENONCÉ

3.1.1. DÉPLACEMENTS ALÉATOIRES. On dispose au total de  $N$  voitures. On considère dans cette première partie les évolutions aléatoires, dues aux seuls usagers, des nombres de voitures de chacun des parkings, appelés A et B. Pour étudier ces déplacements on discrétise le temps. On suppose que le pas de discrétisation en temps est tel qu'au plus une voiture puisse être déplacée par un usager en une étape de temps. On appelle  $\mu > 0$  la probabilité qu'une voiture ait le temps d'être déplacée de A vers B, en une étape de temps, si le nombre de voitures de A n'est pas nul. La quantité  $\mu$  est supposée indépendante du nombre de voitures de A<sup>1</sup>. Si le nombre de voitures de A est nul, la probabilité d'un déplacement de A vers B est 0. De même on introduit la probabilité  $\nu > 0$  d'un déplacement de B vers A. On a  $\mu + \nu \leq 1$ . On appelle  $X_k$  le nombre de voitures dans A à l'instant  $k$ .

1. Donnez la matrice de transition ( $M$ ) de la chaîne de Markov  $X_k$ .
2. Soit  $x$  le nombre de voitures dans A à l'instant 0, donnez l'équation vérifiée par les probabilités de  $X_k = y$ ,  $0 \leq y \leq N$ ,  $k > 0$ , vues comme une fonction de  $k$  et de  $y$ .
3. Montrez que  $X_k$  admet une seule mesure invariante. Donnez l'équation satisfaite par cette mesure invariante. Calculez explicitement cette mesure invariante<sup>2</sup>.

3.1.2. TRANSPORTS. On suppose maintenant que les voitures ne puissent plus être déplacées aléatoirement mais qu'elles soient seulement transportées par les responsables de la gestion des parkings. A chaque étape de temps on doit décider de transporter une voiture de A vers B (cela coûte  $a$ ) ou de B vers A (cela coûte  $b$ ) ou de ne pas transporter (cela coûte 0). On veut transporter les voitures au moindre coût. On appelle  $X_k$  la population de A à l'instant  $k$ .

1. Donnez la matrice des coûts transition ( $C$ ) de la chaîne de Bellman  $X_k$ .
2. Soit  $x$  le nombre de voitures dans A à l'instant 0, donnez l'équation vérifiée par les coûts minimaux de transport pour que  $X_k = y$ ,  $0 \leq y \leq N$ ,  $k > 0$ , vus comme une fonction de  $k$  et de  $y$ .

<sup>1</sup>On sait traiter le cas où  $\mu$  et dépend du nombre de voitures de A.

<sup>2</sup>Le même problème est résoluble également pour  $n$  parkings.

- Donnez l'expression, en fonction de  $C$ , du coût minimal des transports, effectués en un nombre arbitraire d'étapes, faisant passer le nombre de voitures de  $A$  de  $x$  à  $y$ . Calculez explicitement ce coût par des calculs dans l'algèbre min-plus. Comparez le résultat à celui obtenu grâce à la stratégie optimale triviale que l'on précisera<sup>3</sup>.

3.1.3. DÉPLACEMENTS ALÉATOIRES ET TRANSPORTS. On suppose maintenant que chaque période de temps soit décomposée en deux parties dans la première partie on transporte les voitures dans la deuxième partie les voitures sont déplacées de façon aléatoire. Le transport et le déplacement se font selon les protocoles indiqués aux sections précédentes.

- On aimerait transporter les voitures de façon à ce que la variable aléatoire donnant le nombre de voitures dans  $A$  soit aussi concentrée que possible (en régime stationnaire) autour d'une valeur donnée  $y$ ,  $0 < y < N$ . Quelle politique stationnaire peut-on utiliser? Quel est le support de la mesure invariante de la chaîne de Markov correspondante à cette politique?
- Pour obtenir ce résultat qualitatif, on peut résoudre le problème de commande stochastique sur un horizon infini avec actualisation  $\alpha$ , et, avec, en plus du coût de transport, un coût sur l'état  $X_k$  égal à 0 pour  $X_k = y - 1, y, y + 1$  et un coût égal à  $q$ , grand, ailleurs. Explicitiez l'équation de la programmation dynamique correspondante. On la donnera dans l'algèbre ordinaire et également sous forme matricielle en utilisant l'algèbre min-plus. Donnez un algorithme de résolution de ce système d'équations.

## 3.2. CORRIGÉ

3.2.1. DÉPLACEMENTS ALÉATOIRES.

- La matrice de transition ( $M$ ) de la chaîne de Markov  $X_k$  est :

$$M = \begin{pmatrix} 1 - \nu & \nu & 0 & \cdot & 0 \\ \mu & 1 - \nu - \mu & \nu & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \mu & 1 - \nu - \mu & \nu \\ 0 & \cdot & 0 & \mu & 1 - \mu \end{pmatrix}$$

- Si l'on  $p_y^k$  la probabilité de  $X_k = y$  on a

$$p_{k+1} = p_k M, \quad p_0 = (0, \dots, 0, 1, 0, \dots, 0)$$

où le 1 est en position  $x + 1$ .

- $X_k$  admet une seule mesure invariante car le graphe associé à  $M$  a une seule composante fortement connexe puisque  $\mu > 0$  et  $\nu > 0$ . L'équation satisfaite par cette mesure invariante est  $p = pM$ . Pour calculer  $p$  il faut donc résoudre la récurrence linéaire du second ordre

$$\mu p_{x+1} + \nu p_{x-1} = (\mu + \nu) p_x,$$

<sup>3</sup>Le cas de  $n$  parkings est moins trivial.

$$\mu p_1 = \nu p_0$$

$$\nu p_{N-1} = \mu p_N .$$

La solution de cette récurrence est :

$$p_x = (\nu/\mu)^x (1 - \nu/\mu) / (1 - (\nu/\mu)^{N+1}) .$$

### 3.2.2. TRANSPORTS.

1. La matrice des coûts transition ( $C$ ) de la chaîne de Bellman  $X_k$  est donnée par :

$$C = \begin{pmatrix} 0 & b & +\infty & \cdot & +\infty \\ a & 0 & b & \cdot & +\infty \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ +\infty & \cdot & a & 0 & b \\ +\infty & \cdot & +\infty & a & 0 \end{pmatrix}$$

2. Si l'on  $v_y^k$  les coûts optimaux pour que  $X_k = y$  on a

$$v_{k+1} = v_k \otimes C, \quad v_0 = (+\infty, \dots, +\infty, 0, +\infty, \dots, +\infty)$$

où le 0 est en position  $x + 1$ .

3. L'expression, en fonction de  $C$ , du coût minimal de transport en un nombre arbitraire d'étapes, faisant passer le nombre de voitures  $A$  de  $x$  à  $y$  est donné par  $C_{xy}^*$  avec

$$C^* = E \oplus C \oplus C^2 \oplus \dots \oplus C^N .$$

où  $E$  est la matrice diagonale ayant des 0 sur la diagonale et des  $+\infty$  ailleurs. Le calcul de  $C^*$  se fait facilement on obtient :

$$C^* = \begin{pmatrix} 0 & b & b^2 & \cdot & b^N \\ a & 0 & b & \cdot & b^{N-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a^{N-1} & \cdot & a & 0 & b \\ a^N & \cdot & a^2 & a & 0 \end{pmatrix} .$$

Supposons que  $x > y$  le coefficient  $C_{xy}^* = a^{x-y}$  dans l'algèbre min-plus soit  $a(x-y)$  dans l'algèbre ordinaire qui correspond bien au coût de transport des  $x - y$  voitures en trop de  $A$ . La politique optimale consistant dans ce cas, bien sûr, à transporter un à un  $x - y$  voitures de  $A$  vers  $B$  puis de ne plus rien transporter. Tout autre politique nécessitera plus de transports. Tout transport ayant un coût positif. Une autre politique ne peut pas être optimale.

### 3.2.3. DÉPLACEMENT ALÉATOIRES ET TRANSPORTS.

1. On peut utiliser la politique consistant à transporter une voiture de  $A$  vers  $B$  dès que l'état est supérieur à  $y$ , ne rien transporter si l'état vaut  $y$ , transporter de  $B$  vers  $A$  si l'état est inférieur à  $y$ . La chaîne de Markov admet alors une seule classe finale dont les états sont  $\{y - 1, y, y + 1\}$ . Son unique mesure invariante a pour support cette classe finale.

2. Considérons la matrice  $D$  ayant pour coefficient  $D_{xx'} = 0$  si  $x = y - 1, y, y + 1$  tout  $x'$  et  $q$  ailleurs. A partir de  $C$  et de  $D$  formons la matrice  $C'$  dont les coefficients sont obtenus en faisant la somme terme à terme des éléments de  $C$  et de  $D$ . On remarque que la matrice  $C'$  est tridiagonale au sens de l'algèbre min-plus. L'équation de la programmation dynamique actualisée s'écrit alors simplement

$$v \times (1 + \alpha) = C' \otimes (Mv) .$$

L'opérateur  $\times$  signifie ici la multiplication habituelle au sens de l'algèbre ordinaire.

On peut écrire cette équation de façon classique au prix de beaucoup plus de lourdeur :

$$(1 + \alpha)v_x = \min\{a + \mu v_{x-2} + (1 - \mu - \nu)v_{x-1} + \nu v_x, \quad (3.1)$$

$$\mu v_{x-1} + (1 - \mu - \nu)v_x + \nu v_{x+1}, \quad (3.2)$$

$$b + \mu v_x + (1 - \mu - \nu)v_{x+1} + \nu v_{x+2}\} + c_x , \quad (3.3)$$

pour tout  $x, 0 \leq x \leq N$ . où  $c_x = 0$  si  $x = y - 1, y, y + 1$  et  $q$  ailleurs.

On peut résoudre cette équation par l'algorithme de Howard.

## 4. GESTION DE RÉSERVOIR

On considère la gestion d'un barrage destiné à produire de l'électricité. On étudie d'abord l'évolution de l'état pour une commande ne dépendant pas de l'état du barrage. Dans une deuxième partie on optimise la politique de gestion du réservoir de façon à maximiser l'énergie produite. On se place dans une situation simple où des calculs explicites peuvent être effectués.

On modélise l'évolution du stock d'eau dans le barrage par une chaîne de Markov. Chaque état représente un niveau d'eau. On a  $E$  niveaux:  $\{0, 1, \dots, E - 1\}$ . La probabilité de passer de l'état  $x$  à l'état  $x + 1$  pour  $x = 0, 1, \dots, E - 2$  est supposée être indépendante de  $x$  et égale à  $\rho$ . Elle correspond à la probabilité d'un apport d'eau. Lorsque le stock est plein ( $x = E - 1$ ) un tel apport conduit à un débordement et l'état reste à  $E - 1$ . La probabilité de passer de l'état  $x$  à  $x - 1$  vaut  $u$  avec  $0 \leq u \leq 1 - \rho$  et sera considérée comme une commande à optimiser dans la deuxième partie. Ce type de transitions correspond aux situations de turbinage du barrage. Lorsque le barrage est vide on ne peut plus turbiner ( $u=0$ ). La quantité d'énergie produite à chaque période de temps vaut  $xu$  si le niveau d'eau est  $x$  et le turbinage  $u$ . On veut maximiser l'énergie produite sachant que lorsque le stock est plein on risque de déborder et lorsque ce stock est vide on perd de l'énergie potentielle.

### 4.1. ENONCÉ

4.1.1. ETUDE DE LA CHAÎNE DE MARKOV POUR UNE STRATÉGIE CONSTANTE. On suppose ici que la stratégie de gestion est constante en temps et indépendante du niveau d'eau  $s_x^n = u, \forall x, n$  où  $n$  désigne le temps.

1. Donnez la matrice de transition de la chaîne de Markov représentant l'évolution du stock d'eau pour cette politique. On appellera  $M$  cette matrice.
2. Donnez l'équation satisfaite par la loi marginale  $p_x^n$  pour que le stock soit au niveau  $x$  à l'instant  $n$ . On supposera donnée une loi initiale  $p^0$ .
3. Lorsque  $n$  augmente on suppose que cette loi marginale se stationnarise. On appelle  $p$  cette loi de probabilité en régime stationnaire. Quelle équation satisfait elle?
4. Montrez qu'il y a unicité de ce régime stationnaire. On s'aidera de la section "base du  $\mathcal{N}(A)$ " du chapitre 1 du polycopié.
5. Discutez du support de cette mesure invariante en fonction de  $u$ .
6. En cherchant  $p_x$  sous la forme  $p_x = C\beta^x$  (avec  $C$  et  $\beta$  deux constantes positives), dans le cas  $u > 0$ , on donnera une formule explicite pour  $p$ .
7. Etant donné un taux d'actualisation  $\lambda$ , donnez l'équation satisfaite par:

$$v_x^\lambda = E\left\{\sum_0^\infty \frac{1}{(1+\lambda)^{n+1}} X^n U^n \mid X^0 = x\right\},$$

où  $X^n$  désigne l'état de la chaîne de Markov et  $U^n$  sa commande à l'instant  $n$ .

8. Montrez que  $\lambda v^\lambda$  reste borné lorsque  $\lambda$  tend vers 0.
9. On fait un développement de  $v^\lambda$  sous la forme  $v/\lambda + w^0 + \lambda w^1 + \dots$ . Donnez un système d'équations caractérisant  $v$ .
10. En déduire que  $v$  est un vecteur constant.
11. Quelle est l'interprétation probabiliste de  $v$ ?
12. Calculez explicitement  $v$ .

#### 4.1.2. OPTIMISATION DE LA POLITIQUE DE GESTION.

1. Formulez en termes de commande optimale stochastique le problème de gestion optimale dans le cas d'un horizon infini avec un taux d'actualisation  $\lambda$ .
2. Explicitez l'équation de la programmation dynamique correspondante?
3. Montrez que la stratégie est du type bang-bang. On supposera, dans la suite du problème, avoir démontré qu'en dessous d'un certain stock  $\alpha$  on ne turbine pas et qu'au dessus de ce seuil on turbine au maximum  $u = 1 - \rho$ .
4. On fait tendre à nouveau  $\lambda$  vers 0 et on fait à nouveau un développement en  $\lambda$  de la fonction valeur. Donnez l'équation définissant le premier terme du développement de la fonction de la programmation dynamique.
5. Dans le cas où on a une contrainte de turbinage minimum ( $u \geq \gamma$ , pour  $\gamma > 0$ ) est imposée, indiquez une façon de calculer le seuil  $\alpha$  déterminant la stratégie optimale.

## 4.2. CORRIGÉ

## 4.2.1. ETUDE DE LA CHAÎNE DE MARKOV POUR UNE STRATÉGIE CONSTANTE.

1. La matrice de transition de la chaîne de Markov représentant l'évolution du stock d'eau pour la politique constante est:

$$M = \begin{pmatrix} 1 - \rho & \rho & 0 & \cdot & 0 \\ u & 1 - \rho - u & \rho & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & u & 1 - \rho - u & \rho \\ 0 & \cdot & 0 & u & 1 - u \end{pmatrix}.$$

2. L'équation satisfaite par la loi marginale  $p_x^n$  pour que le stock soit au niveau  $x$  à l'instant  $n$  est donnée par:

$$p^{n+1} = p^n M .$$

3. Lorsque  $n$  augmente on cette loi marginale se stationnarise sur  $p$  satisfaisant:

$$p = pM .$$

4. Il y a unicité de ce régime stationnaire car la chaîne de Markov admet une seule classe finale quel que soit  $u$  dès que  $\rho > 0$ .  
5. On suppose  $\rho > 0$ .

Si  $u > 0$  le support de  $p$  est l'espace tout entier.

Si  $u = 0$  le support de  $p$  est l'état  $E - 1$ .

6. En cherchant  $p_x$  sous la forme  $p_x = C\beta^x$  on voit que:

$$\beta = \rho/u .$$

et donc:

$$C = (\rho/u - 1)/((\rho/u)^E - 1) .$$

7. L'équation satisfaite par  $v_x^\lambda$  est l'équation de Kolmogorov arrière:

$$(1 + \lambda)v_x^\lambda = (Mv^\lambda)_x + xu , \forall x .$$

8. Le fait que  $\lambda v^\lambda$  reste borné lorsque  $\lambda$  tend vers 0 provient de la bornitude de  $xu$  et de:

$$\sum_n \lambda/(1 + \lambda)^{n+1} = 1 .$$

9. En identifiant les termes en  $\lambda$  et en conservant les deux premiers termes on obtient le système d'équations:

$$v = Mv ,$$

$$v_x + w_x^0 = (Mw^0)_x + xu ,$$

qui caractérise complètement  $v$ . En effet la première équation implique que  $v$  est dans le noyau de  $M - I$  dont on sait qu'il est de dimension 1. La deuxième équation multipliée à gauche par  $p$  implique  $pv = u \sum_x xp_x$  qui détermine complètement  $v$ .

10. Puisque  $v$  est dans le noyau de  $M - I$  et que  $\sum_{x'} M_{xx'} = 1$ ,  $v$  un vecteur constant. On appellera encore  $v$  la constante correspondante
11. Puisque  $v = \sum_x xup_x$ ,  $v$  s'interprète comme la moyenne du coût au sens de la mesure invariante  $p$ .
12. Le calcul explicite de  $v$  revient à calculer la moyenne d'une loi géométrique tronquée. Notons:

$$S(\beta) = \sum_{x=0, E-1} \beta^x .$$

On a:

$$\beta S'(\beta) = \sum_{0, E-1} x\beta^x .$$

On en déduit :

$$v = u\beta S'(\beta)/S(\beta) ,$$

avec  $\beta = \rho/u$  et  $S(\beta) = (\beta^E - 1)/(\beta - 1)$ .

#### 4.2.2. OPTIMISATION DE LA POLITIQUE DE GESTION.

1. Le problème de commande optimale stochastique s'écrit:

$$v_x^\lambda = \max_{\{U^n, n=0, \dots, \infty\}} \mathbb{E}\left\{\sum_{n=0}^{\infty} 1/(1+\lambda)^{n+1} X^n U^n | X^0 = x\right\} ,$$

où  $X^n$  est la chaîne de Markov de matrice de transition  $M^u = M$ .

2. L'équation de la programmation dynamique correspondante s'écrit:

$$(1+\lambda)v_x^\lambda = \max_{0 \leq u \leq 1-\rho} (M^u v^\lambda)_x + xu , \forall x .$$

3. La stratégie est du type bang-bang puisque

$$(M^u v^\lambda)_x + xu$$

est linéaire en  $u$  pour tout  $x$ .

4. L'équation définissant le premier terme du développement de la fonction de la programmation dynamique s'obtient comme précédemment en conservant les deux premiers termes du développement:

$$v = \max_{0 \leq u \leq 1-\rho} M^u v ,$$

$$v_x + w_x^0 = \max_{u \in U_x^*} ((M^u w^0)_x + xu) , \forall x ,$$

où  $U^*$  désigne l'ensemble des  $u$  réalisant le minimum dans la première équation. La première équation implique que  $v$  est un vecteur constant et que

$$U_x^* = [0, 1 - \rho]$$

(la première équation indique seulement que  $v$  est constant et ne donne pas d'information sur  $u$ ). Ces 2 équations peuvent donc se réécrire en une seule: rechercher la constante  $v$  et  $w$  tels que

$$v + w_x = \max_{0 \leq u \leq 1-\rho} ((M^u w)_x + xu), \forall x.$$

5. Dans le cas où on a une contrainte de turbinage minimum ( $u \geq \gamma$ , pour  $\gamma > 0$ ) est imposée on pourra utiliser l'algorithme de Howard ou calculer explicitement la mesure invariante en utilisant le caractère bang-bang de la stratégie.

## 5. JEUX DE PILE OU FACE

On considère le jeu de pile ou face dans le cas borné. Le joueur s'arrête lorsqu'il a perdu une somme donnée à l'avance ou lorsqu'il a obtenu un gain donné. Après quelques calculs sur ce problème. On vérifie qu'une politique d'arrêt optimal du jeu avant sa fin normale ne conduit à aucune amélioration de la moyenne des gains. On étudie l'influence d'un biais dans le tirage sur l'espérance de gain. On estime le biais. On résout le problème d'arrêt optimal en observation incomplète dans le cas où le biais n'est pas connu mais doit être estimé pendant le jeu.

### 5.1. ENONCÉ

5.1.1. ETUDE DE LA CHAÎNE DE MARKOV ASSOCIÉ AU JEU DE PILE OU FACE. On considère une pièce ayant la probabilité  $p$  de tomber sur face. On fait des tirages indépendants de la pièce. On commence le jeu avec une somme  $x_0$ . Lorsque la pièce tombe sur face on gagne 1. Lorsque la pièce tombe sur pile on perd 1. Lorsque la fortune du joueur a atteint 0 ou  $F$  on n'a plus le droit de jouer. La fortune reste donc à la valeur atteinte pendant les étapes suivantes.

1. Donnez la matrice de transition de la chaîne de Markov représentant l'évolution de la fortune du joueur. On appellera  $M$  cette matrice.
2. Donnez l'équation satisfaite par la loi marginale  $p_x^n$  pour que la fortune soit au niveau  $x$  à l'instant  $n$ .
3. Décrivez les classes finales et transitoires de cette chaîne de Markov.
4. Explicitez les mesures de probabilités invariantes extrémales de cette chaîne (voir le paragraphe base du  $\mathcal{N}(A')$  du chapitre 1 du polycopié).
5. On s'intéresse aux probabilités de gagner (terminer en  $F$ ) et de perdre (terminer en 0). Donner les équations permettant de calculer ces probabilités (voir le paragraphe base du  $\mathcal{N}(A)$  du chapitre 1 du polycopié).
6. Calculer explicitement les probabilités de gagner et de perdre dans le cas  $p = 1/2$ .

7. Calculez explicitement le projecteur spectral sur l'espace propre associé à la valeur propre 1 de la matrice  $M$  c.a.d. la matrice  $P$  stochastique satisfaisant

$$PM = MP = P = P^2$$

dans le cas  $p = 1/2$ .

8. En déduire l'espérance mathématique de gain du joueur dans le cas  $p = 1/2$ .

5.1.2. MARTINGALES. Les joueurs ont toujours rêvé de trouver la bonne "martingale" leur permettant de gagner de l'argent grâce à une procédure d'arrêt astucieuse. On se propose d'optimiser la procédure d'arrêt du jeu (supposée restreinte ici à des temps de sortie d'ensemble d'états). On se place dans le cas  $p = 1/2$ . A chaque instant on a donc maintenant la possibilité de s'arrêter ou de continuer (décision qu'il faut optimiser). Si on s'arrête on conserve la fortune dont on dispose à cet instant.

1. Formulez le problème commande optimale stochastique correspondant.
2. Explicitiez l'équation de la programmation dynamique correspondante.
3. Calculez la solution de l'équation de la programmation dynamique correspondante.
4. En déduire qu'il n'y a pas de procédure d'arrêt permettant d'améliorer les gains.

5.1.3. SENSIBILITÉ AU BIAIS. On suppose que la pièce a un biais  $p = (1 + \epsilon)/2$ . On est dans la situation du paragraphe 2 (on n'optimise pas l'instant d'arrêt).

1. Donnez l'équation de Kolmogorov permettant de calculer l'espérance mathématique du gain.
2. Calculez la solution de cette équation.
3. Calculez la sensibilité autour de  $\epsilon = 0$  de l'espérance de gain c.a.d. la dérivée par rapport à  $\epsilon$  de l'espérance de gain.

5.1.4. BIAIS INCONNU. On considère la nouvelle version du jeu de pile ou face suivant. On prend une pièce dont on ne connaît pas  $p$  puis on joue en optimisant l'arrêt.

1. Donnez l'estimateur du maximum de vraisemblance de  $p$ .
2. Donnez l'équation de la programmation dynamique du problème en observation incomplète correspondant en supposant que l'estimateur du maximum de vraisemblance est égal à l'espérance conditionnelle de  $p$  connaissant le passé des observations.
3. Montrez que la politique optimale d'arrêt n'est plus triviale.

## 5.2. CORRIGÉ

## 5.2.1. ETUDE DE LA CHAÎNE DE MARKOV ASSOCIÉ AU JEU DE PILE OU FACE.

1.

$$M = \begin{pmatrix} 1 & 0 & 0 & \cdot & 0 \\ 1-p & 0 & p & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 1-p & 0 & p \\ 0 & \cdot & 0 & 0 & 1 \end{pmatrix}.$$

2.

$$p^n = p^{n-1}M,$$

$$p^0 = (0 \cdot 0 \ 1 \ 0 \cdot 0),$$

où le 1 est en position  $x_0$ .

3. Il y a 2 classes finales :

$$f_1 = \{0\},$$

$$f_2 = \{F\},$$

et 1 classe transitoire :

$$t = \{1, \dots, F-1\}.$$

4. Il y a 2 mesures de probabilités invariantes extrêmes :

$$q^1 = (1 \ 0 \cdot 0),$$

$$q^2 = (0 \cdot 0 \ 1).$$

5. La probabilités de gagner (terminer en  $F$ ) est solution de :

$$A\chi^2 = 0 \text{ sur } t, \quad \chi_0^2 = 0, \quad \chi_F^2 = 1.$$

La probabilité de perdre (terminer en 0) est solution de :

$$A\chi^1 = 0 \text{ sur } t, \quad \chi_0^1 = 1, \quad \chi_F^1 = 0.$$

6. Dans le cas  $p = 1/2$  on a :

$$\chi_x^2 = x/F,$$

$$\chi_x^1 = 1 - x/F.$$

7. Le projecteur spectral sur l'espace propre associé à la valeur propre 1 de la matrice  $M$  vaut :

$$P = \chi^1 \otimes q^1 + \chi^2 \otimes q^2.$$

Dans le cas  $p = 1/2$  on a :

$$P = \begin{pmatrix} 1 & 0 & \cdot & 0 & 0 \\ 1 - 1/F & 0 & \cdot & 0 & 1/F \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1/F & 0 & \cdot & 0 & 1 - 1/F \\ 0 & 0 & \cdot & 0 & 1 \end{pmatrix}.$$

8. L'espérance mathématique de gain du joueur dans le cas  $p = 1/2$  est :

$$F \chi_{x_0}^2 = x_0.$$

### 5.2.2. MARTINGALES.

1. Soit  $B \subset E = \{0, 1, \dots, F\}$  on doit résoudre :

$$v_x = \max_B \mathbb{E}(X_{\nu_B} | X_0 = x),$$

où  $\nu_B$  désigne le temps de sortie de  $B$  et  $X_n$  la chaîne de Markov décrivant l'évolution de la fortune du joueur.

2. L'équation de la programmation dynamique correspondante s'écrit :

$$v_x = \max\{x, [Mv]_x\}, \quad \forall x \in E.$$

3. La solution de cette équation est :

$$v_x = x.$$

4. Il est donc indifférent de s'arrêter ou de continuer puisque le max est atteint simultanément par  $x$  et  $Mv$ .

### 5.2.3. SENSIBILITÉ AU BIAIS.

1. L'équation de Kolmogorov donnant l'espérance mathématique du gain est :

$$v = Mv \text{ sur } t, \quad v_0 = 0, \quad v_F = F.$$

2. Cette équation vectorielle s'explique complètement :

$$v_x = 1/2[v_{x+1} + v_{x-1} + \epsilon(v_{x+1} - v_{x-1})] \text{ sur } t.$$

et donc en posant  $\alpha = (1 - \epsilon)/(1 + \epsilon)$  on a :

$$v_x = F \frac{1 - \alpha^x}{1 - \alpha^F}.$$

3. Le développement asymptotique à 2 l'ordre est :

$$v_x = x + \epsilon x(F - x) + o(\epsilon).$$

## 5.2.4. BIAIS INCONNU.

1. L'estimateur du maximum de vraisemblance de  $p$  est obtenu en maximisant par rapport  $p$  la vraisemblance  $C_n^y p^y (1-p)^{n-y}$  avec  $y$  le nombre de fois où face a été obtenu dans un tirage de longueur  $n$  dans lequel 0 ou  $F$  n'ont pas été atteint. On obtient  $\hat{p} = y/n$ .
2. L'équation de la programmation dynamique du problème en horizon fini et en observation incomplète (en supposant que l'estimateur du maximum de vraisemblance est égal à l'espérance conditionnelle de  $p$  connaissant le passé des observations) est alors :

$$v^n[x, p] = \max\{x, pv^{n+1}[x+1, (np+1)/(n+1)] \\ + (1-p)v^{n+1}[x-1, np/(n+1)]\}, \\ \forall n < N, \forall x \in [0, F], \forall p \in [0, 1],$$

$$v^n[0, p] = 0, v^n[F, p] = F, \forall p, \forall n < N,$$

$$v^N[x, p] = x, \forall x, p.$$

3. Cette équation n'admet plus  $x$  pour solution et donc la politique optimale d'arrêt n'est plus triviale. Il n'est plus indifférent de s'arrêter ou de continuer.

## 6. PROCESSUS DE CALCUL

On modélise un processus de calcul. On établit le lien avec un problèmes de contrôle d'une chaîne de Markov. On évalue la relation entre le nombre entrées et le nombre de sorties par des calculs dans l'algèbre min-plus.

## 6.1. LE PROCESSUS DE CALCUL

On considère un processeur constitué d'un multiplieur et d'un additionneur séparé par des mémoires appelées A. Ce processeur est chargé de faire le calcul  $ae_n + b$  pour une suite infinie d'entrées  $e_n$ . On suppose que les  $e_n$  peuvent s'accumuler dans des mémoires (appelées B) au fur et à mesure de leurs arrivées avant leurs traitements. Le processus de calcul fonctionne de la façon suivante.

1. Le multiplieur dès qu'il a terminé la multiplication précédente regarde s'il existe une donnée  $e_n$  à traiter et s'il a l'autorisation de commencer. Si c'est la cas il fait la multiplication qui dure 2 tops d'horloge et il met le résultat dans une des mémoires A. Ce faisant il a consommé une autorisation de multiplier.
2. L'additionneur dès qu'il a terminé l'addition précédente regarde s'il y a une donnée dans A. Si c'est le cas il fait l'addition qui dure 1 top d'horloge, affiche le résultat et envoie une nouvelle autorisation au multiplieur dès que l'addition est terminée.
3. A l'instant initial le multiplieur dispose de 2 autorisations.

Le but de ces autorisations de calcul est d'éviter l'écrasement des données dans la mémoire A car sa capacité est limitée à 2.

On appelle compteurs les fonctions du temps discret  $t$  (nbre de tops d'horloge) donnant le nombre de données arrivées dans B ( $U_t$ ), le nombre de multiplications commencées ( $X_t^1$ ), le nombre d'additions commencées ( $X_t^2$ ) et le nombre de résultats affichés ( $Y_t$ ) du top 0 jusqu'au top  $t$ .

On appelle de même dateurs les fonctions du numéro d'événement  $n$  donnant la date ( $U_n$ ) à laquelle est disponible  $e_n$  dans B, la date ( $X_n^1$ ) à laquelle la multiplication numérotée  $n$  commence, la date ( $X_n^2$ ) à laquelle l'addition numérotée  $n$  commence et la date ( $Y_n$ ) de l'affichage numéroté  $n$ .

## 6.2. ENONCÉ

1. Montrez que le système d'équations régissant les compteurs est

$$X_t^1 = \min\{U_t, 2 + X_{t-1}^2, 1 + X_{t-2}^1\},$$

$$X_t^2 = \min\{X_{t-2}^1, 1 + X_{t-1}^2\},$$

$$Y_t = X_{t-1}^2.$$

avec les conditions initiales:

$$X_0^1 = X_{-1}^1 = X_0^2 = 0.$$

2. Après avoir fait une augmentation d'états et une inversion du temps interprétez ce système d'équations en terme de contrôle optimal d'une chaîne de Markov (états, contrôles, matrices de transition, coût à optimiser, type du problème).
3. Ecrivez ce système d'équations sous forme matricielle en utilisant l'algèbre min-plus.
4. Montrez que le système s'écrit  $X_t = A \otimes X_{t-1}$  lorsque  $U_0 = +\infty$ . Calculez  $A^n$  pour  $n = 2, \dots, 5$ . Par quelle relation sont liés  $A^{\otimes 3}$  et  $A^{\otimes 5}$ ?
5. Qu'en deduez vous sur le comportement asymptotique de  $A^t$  lorsque  $t$  tend vers l'infini. Interprétez ce résultat physiquement.
6. Réécrivez le système récurrent initial dans l'algèbre min-plus.
7. En utilisant les séries génératrices min-plus définies pour une suite  $\{X_t, t \in N\}$  par

$$\widehat{X}(\delta) = \bigoplus_{t \geq 0} X_t \delta^t$$

montrez qu'il existe une série  $\widehat{H}$  telle que  $\widehat{Y} = \widehat{H} \otimes \widehat{U}$  (on précisera le sens de ce produit  $\otimes$  de séries). On calculera explicitement  $\widehat{H}$  en utilisant la notation  $S^* = 0 \oplus S \oplus S^{\otimes 2} \oplus \dots$  pour  $S$  une série en  $\delta$  (cet opérateur  $*$  sert à résoudre certaines équations linéaires que l'on explicitera).

8. En déduire que la relation entrée ( $U$ ) sortie ( $Y$ ) est une inf-convolution<sup>4</sup> de l'entrée avec une fonction que l'on interprétera physiquement. On indiquera l'expérience physique permettant d'obtenir cette fonction sans faire de calcul.
9. Donnez le système d'équations régissant les dateurs.

### 6.3. CORRIGÉ

1. Il s'obtient en exprimant les contraintes imposées sur les nombres de tâches effectuées jusqu'à un instant  $t$  par le multiplieur et l'additionneur

$$\begin{aligned} X_t^1 &\leq U_t, \\ X_t^1 &\leq 2 + X_{t-1}^2, \\ X_t^1 &\leq 1 + X_{t-2}^1, \\ X_t^2 &\leq X_{t-2}^1, \\ X_t^2 &\leq 1 + X_{t-1}^2. \end{aligned}$$

et en remarquant que fonctionnant au plus vite une des trois premières contraintes et une des deux dernières contraintes sont nécessairement saturées.

2. En introduisant l'état supplémentaire  $X_t^3 = X_{t-1}^1$  et en faisant le changement de temps  $t' = -t$  on obtient le nouveau système

$$\begin{aligned} X_{t'}^1 &= \min\{U_{t'}, 2 + X_{t'+1}^2, 1 + X_{t'+1}^3\}, \\ X_{t'}^3 &= X_{t'+1}^1, \\ X_{t'}^2 &= \min\{X_{t'+1}^3, 1 + X_{t'+1}^2\}, \\ Y_{t'} &= X_{t'+1}^2. \end{aligned}$$

Les 3 premières équations sont les équations de la programmation dynamique d'un problème de contrôle de chaîne de Markov à trois états (l'état est noté  $Z$ ), 2 contrôles (le contrôle est noté  $V$ ) et la possibilité de s'arrêter, en horizon fini, de matrices de transition

$$M^1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad M^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

de coût

$$c^1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \quad c^2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Le critère à minimiser s'écrit

$$X_{t'}^i = \min_{V, v} E \left\{ \sum_{s=t'}^{\inf(v, 0)} c_{Z_s}^{V_s} + 1_{v \leq 0} U_v \mid Z_{t'} = i \right\},$$

<sup>4</sup>On rappelle que l'inf-convolution de 2 fonctions  $f$  et  $g$  notée  $f \star g$  est définie par  $[f \star g](z) = \inf_x (f(x) + g(z - x))$ .

où  $\nu$  est un temps d'arrêt que l'on optimise (chaque fois que la chaîne passe dans l'état 1 on peut décider de s'arrêter et de payer  $U_s$  (si l'instant correspondant est  $s$ ) ou de continuer).

3. On peut écrire l'équation de la programmation dynamique en utilisant l'algèbre min-plus sous la forme:

$$X_t = A \otimes X_{t-1} \oplus B \otimes U_t ,$$

$$Y_t = C \otimes X_{t-1} .$$

avec

$$A = \begin{pmatrix} +\infty & 2 & 1 \\ +\infty & 1 & 0 \\ 0 & +\infty & +\infty \end{pmatrix} B = \begin{pmatrix} 0 \\ +\infty \\ +\infty \end{pmatrix} X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$C = ( +\infty \ 0 \ +\infty ) .$$

4. Lorsque  $U_0 = +\infty$  on a  $X_t = A^{\otimes t} X_0$  puisque  $B \otimes U_t = +\infty$  puisque les  $U_t$  sont croissants.

$$A^3 = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix} , A^5 = \begin{pmatrix} 3 & 4 & 3 \\ 2 & 3 & 2 \\ 2 & 4 & 3 \end{pmatrix} .$$

On constate que  $A^5 = 1 \otimes A^3$ .

5. Pour tout  $n \geq 3$  on a  $A^{n+2} = 1 \otimes A^n$ . Physiquement cette périodicité correspond à un remplissage du pipeline de calcul. Une fois ce pipeline rempli, le processeur va à la vitesse du multiplieur qui est la ressource la plus lente du système.
6. Le système récurrent initial s'écrit dans l'algèbre min-plus

$$X_t^1 = U_t \oplus 2 \otimes X_{t-1}^2 \oplus 1 \otimes X_{t-2}^1 ,$$

$$X_t^2 = X_{t-2}^1 \oplus 1 \otimes X_{t-1}^2 ,$$

$$Y_t = X_{t-1}^2 .$$

7. En multipliant par  $\delta^t$  et en sommant en  $t$  on obtient le système

$$\widehat{X}^1 = \widehat{U} \oplus 2 \otimes \delta \otimes \widehat{X}^2 \oplus 1 \otimes \delta^2 \otimes \widehat{X}^1 ,$$

$$\widehat{X}^2 = \delta^2 \otimes \widehat{X}^1 \oplus 1 \otimes \delta \otimes \widehat{X}^2 ,$$

$$\widehat{Y} = \delta \otimes \widehat{X}^2 .$$

En utilisant le fait que  $a*b$  est solution de  $X = a \otimes X \oplus b$  on peut éliminer successivement  $\widehat{X}^2$  puis  $\widehat{X}^1$ . On obtient alors (l'opérateur  $\otimes$  des séries <sup>5</sup> ne sera plus noté dans la suite)  $\widehat{Y} = \widehat{H}(\delta)\widehat{U}$ , avec

$$\widehat{H}(\delta) = \delta(1\delta^2 \oplus 2\delta(1\delta)^*\delta^2)^*(1\delta)^*\delta^2 ,$$

$$= \delta^3(1\delta)^*(1\delta^2)^* .$$

<sup>5</sup>défini par  $\widehat{S} \otimes \widehat{T} = \bigoplus_t [\bigoplus_s S_s \otimes T_{t-s}] \delta^t$ .

8. Par identification avec

$$\widehat{H}(\delta) = \bigoplus_t H_t \otimes \delta^t$$

on peut déterminer les coefficients  $H_t$ . L'expression  $\widehat{Y} = \widehat{H}(\delta) \otimes \widehat{U}$  montre alors que

$$Y_t = \inf_{s \geq 0} \{H_{t-s} + U_s\}$$

c.a.d. que  $Y$  est obtenu en faisant l'inf-convolution de l'entrée  $U$  avec  $H$  qui peut être considérée comme la réponse impulsionnelle du système.

Physiquement les coefficients  $H_t$  de la réponse impulsionnelle sont obtenus en rendant disponible une suite infini de données à l'entrée à l'instant 0 et en observant le nombre de sorties obtenu jusqu'à l'instant  $t$ .

9. La modélisation dateur du processeur s'écrit

$$\begin{aligned} X_n^1 &= \max\{U_n, 1 + X_{n-2}^2, 2 + X_{n-1}^1\}, \\ X_n^2 &= \max\{2 + X_n^1, 1 + X_{n-1}^2\}, \\ Y_n &= 1 + X_n^2. \end{aligned}$$

## 7. FILTRAGE ET COMMANDE EN OBSERVATION INCOMPLÈTE

### 7.1. ENONCÉ

Les deux questions sont indépendantes bien qu'il y ait une progression dans les notions utilisées.

7.1.1. FILTRAGE. On considère une chaîne de Markov à 2 états dont la probabilité de sauter de l'état 1 à l'état 2 est égale à celle de sauter de l'état 2 à l'état 1 et vaut  $p$  inconnu pouvant prendre les valeurs  $p^1$  ou  $p^2$  avec  $0 < p^i < 1$  pour  $i = 1, 2$ . On observe la trajectoire de la chaîne. On suppose qu'a priori il y ait une probabilité de 1/2 pour que la valeur de  $p$  soit  $p^1$  et donc une valeur de 1/2 pour qu'elle soit  $p^2$ . Au départ la chaîne de Markov est dans l'état 1.

1. Ramenez le calcul de la probabilité a posteriori que  $p$  soit égal à  $p^1$  (après  $n$  observations de l'état) à un problème de filtrage d'une chaîne de Markov sur un autre espace d'états.
2. Donnez l'équation récurrente satisfaite par l'espérance conditionnelle pour que  $p = p^1$  ou  $p^2$  connaissant la trajectoire passée des états de la chaîne initiale.
3. Supposons avoir observé la suite d'états (1,2,1,1). Quelle est la probabilité conditionnelle que  $p = p^1$ .
4. Supposons avoir observé  $k$  changements d'états pour  $N + 1$  observations quelle est la probabilité conditionnelle pour que  $p = p^1$ .

5. Supposons que  $p = p^1$  montrer que la probabilité conditionnelle précédente tend vers 1 lorsque le nombre d'observations tend vers l'infini.

7.1.2. MAINTENANCE. On veut assurer la maintenance d'un groupe électrogène d'une centrale nucléaire servant à assurer l'existence d'une source d'électricité en cas de panne de la centrale. En régime normal le groupe ne fonctionne pas donc on ne sait s'il est en état de fonctionner ou non. Pour le savoir on doit le tester. A chaque période de temps on doit donc décider de le tester ou pas. On considère qu'il n'a que deux états possibles soit il fonctionne soit il ne fonctionne pas. Si on le teste on observe l'état mais cela coûte le prix du test ( $k_1$ ) plus le coût de remplacement ( $k_2$ ) s'il ne fonctionne pas (remplacement que l'on fait systématiquement en cas de panne). A la suite du remplacement le système fonctionne nécessairement. Si on ne le remplace pas le système à une probabilité  $p$  ( $0 < p < 1$ ) d'être en panne à la fin du test. On prend donc un risque que l'on voudrait éviter. Pour cela on affecte un coût  $k_0$  à l'état de panne. L'état peut donc prendre 2 valeurs: 1) fonctionne, 2) fonctionne pas. L'observation peut donc prendre trois valeurs: 1) on n'a pas testé et donc rien observé 2) on a testé et observé que le système ne fonctionne pas, 3) on a testé et observé que le système fonctionne. Au début de la gestion le système est en état de fonctionnement.

1. Donnez les 6 matrices de transitions (de sauter d'un état à un autre) indexées par l'observation faite et la commande utilisée.
2. La probabilité conditionnelle  $q^n$  d'être dans un état sachant le passé des observations et des commandes (où  $n$  indique la période de temps) est une chaîne de Markov dont l'état vit sur l'espace des probabilités chargeant les deux états possibles ( $q \in \mathbb{R}^2$ ,  $q_1 + q_2 = 1$ ,  $q_1, q_2 \geq 0$ ). A chaque étape elle saute dans cet ensemble selon la valeur de l'observation. Caractérisez le point d'arrivée en fonction du point de départ ainsi que la probabilité pour que ce saut ait lieu.
3. Exprimez le critère à optimiser en terme de cette espérance conditionnelle lorsque l'on gère le système sur  $N$  périodes.
4. Donnez l'équation de la programmation dynamique permettant de déterminer la décision optimale de tester ou non le matériel en fonction du passé des observations et des commandes (ou ce qui revient au même en fonction de la valeur de l'espérance conditionnelle de l'état connaissant le passé des observations et des commandes).
5. Donnez les grandes lignes d'une méthode effective de calcul de cette commande optimale.

## 7.2. CORRIGÉ

### 7.2.1. FILTRAGE.

1. Le problème de filtrage consiste à considérer la chaîne observé à 4 états  $\mathcal{F} = \{(1, p^1), (2, p^1), (1, p^2), (2, p^2)\}$  les observations possibles étant au nombre de 2. L'observation vaut 1 si la chaîne est

dans l'ensemble d'états

$$\{(1, p^1), (1, p^2)\}$$

et vaut 2 si la chaîne est dans l'ensemble

$$\{(2, p^1), (2, p^2)\}.$$

La matrice de transition de sauter d'un état dans un autre et d'observer 1 vaut

$$M^1 = \begin{pmatrix} 1 - p^1 & 0 & 0 & 0 \\ p^1 & 0 & 0 & 0 \\ 0 & 0 & 1 - p^2 & 0 \\ 0 & 0 & p^2 & 0 \end{pmatrix},$$

celle d'observer 2 vaut

$$M^2 = \begin{pmatrix} 0 & p^1 & 0 & 0 \\ 0 & 1 - p^1 & 0 & 0 \\ 0 & 0 & 0 & p^2 \\ 0 & 0 & 0 & 1 - p^2 \end{pmatrix}.$$

On vérifie que  $M^1 + M^2$  est une matrice stochastique.

2. La loi initiale  $q^0$  est donnée par la matrice

$$q^0 = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

où la première ligne correspond à l'observation  $y = 1$  et la deuxième à  $y = 2$ .

Pour obtenir la probabilité conditionnelle (notée  $q^3$ ) que  $p = p^1$  ou  $p = p^2$  après les observations (1,2,1,1) il faut faire les produits matricielles

$$q^{01} M^2 M^1 M^1$$

et renormaliser le résultat pour en faire une probabilité soit

$$q^3 = \frac{1}{c} \begin{pmatrix} (p^1)^2(1 - p^1) & 0 & (p^2)^2(1 - p^2) & 0 \end{pmatrix},$$

avec  $c = (p^1)^2(1 - p^1) + (p^2)^2(1 - p^2)$  et donc la probabilité pour que  $p = p^1$  est la probabilité pour que la chaîne étendu soit dans un des 2 premiers états c.a.d.

$$(p^1)^2(1 - p^1)/c.$$

3. Les formules précédentes se généralisent au  $k$  changements d'états pour  $N$  observations. La probabilité conditionnelle pour que  $p = p^1$  vaut

$$\frac{(p^1)^k(1 - p^1)^{N-k}}{(p^1)^k(1 - p^1)^{N-k} + (p^2)^k(1 - p^2)^{N-k}}.$$

On appellera  $\alpha_N$  cette probabilité.

4. Si  $p = p^1$  d'après la loi des grands nombres on a

$$\lim_N k/N = p^1$$

presque sûrement.

On peut réécrire  $\alpha_N$  sous la forme

$$\frac{1}{1 + \left[ \frac{(p^2)^{k/N} (1-p^2)^{(N-k)/N}}{(p^1)^{k/N} (1-p^1)^{(N-k)/N}} \right]^N}.$$

Mais

$$\frac{(p^2)^{k/N} (1-p^2)^{(N-k)/N}}{(p^1)^{k/N} (1-p^1)^{(N-k)/N}}$$

et presque sûrement aussi voisin de

$$\frac{(p^2)^{p^1} (1-p^2)^{p^1}}{(p^1)^{p^1} (1-p^1)^{1-p^1}}$$

qu'on le veut.

La fonction

$$p \mapsto p^{p'} (1-p)^{1-p'}$$

atteint son maximum au même point que la fonction

$$p \mapsto p' \log(p) + (1-p') \log(1-p)$$

qui est concave. Le maximum est atteint pour  $p = p'$  et donc

$$\left[ \frac{(p^2)^{k/N} (1-p^2)^{(N-k)/N}}{(p^1)^{k/N} (1-p^1)^{(N-k)/N}} \right]^N$$

tend vers 0 presque sûrement.

### 7.2.2. MAINTENANCE.

1. L'espace d'états est  $\mathcal{E} = \{1, 2\}$  où  $x = 1$  correspond à l'état de fonctionnement et  $x = 2$  à l'état de panne.

L'espace des commandes est  $\mathcal{F} = \{1, 2\}$  où la commande  $u = 1$  correspond à la décision de ne pas tester et  $u = 2$  à la décision de tester.

L'espace des observations est  $\mathcal{G} = \{0, 1, 2\}$  où  $y = 0$  correspond à l'absence d'observation,  $y = 1$  correspond à l'observation d'un état de panne,  $y = 2$  correspond à l'observation de l'état de marche.

Si on note par  $M^{uy}$  les 6 matrices de transition on a

$$M^{10} = \begin{pmatrix} 1-p & p \\ 0 & 1 \end{pmatrix}, \quad M^{11} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad M^{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$M^{20} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad M^{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad M^{22} = \begin{pmatrix} 1-p & p \\ 0 & 0 \end{pmatrix}.$$

On vérifie que

$$\sum_{x', y} M_{xx'}^{uy} = 1, \quad \forall x, u.$$

2. Si on appelle  $q$  la probabilité conditionnelle d'un état connaissant le passé des commandes et des observations jusqu'à l'étape  $n$ , la probabilité conditionnelle ( $q'(u)$ ) à l'étape  $n + 1$  dépend de la valeur de la commande  $u$  prise à l'instant  $n$  et prend les valeurs
- $q'(1) = qM^{10}$  avec la probabilité 1 (probabilité pour que  $y = 0$ );
  - $q'(2) = \begin{pmatrix} 1 & 0 \end{pmatrix}$  avec la probabilité  $q_2^n$  (probabilité d'observer une panne compte tenu de l'observation du passé, dans le cas où on a testé le système);
  - $q'(2) = qM^{22}/q_1$  avec la probabilité  $q_1$  (probabilité d'observer le système en état de fonctionnement compte tenu des observations passées, dans le cas où on a testé le système).
3. Le critère à optimiser

$$\min_U \mathbb{E} \left[ \sum_0^n c_{X^n}^{U^n} \right],$$

où  $U^n$  désigne la suite des décisions prises et  $X^n$  la suite des états de la chaîne de Markov avec

$$c = \begin{pmatrix} 0 & k_1 \\ k_0 & k_0 + k_1 + k_2 \end{pmatrix},$$

où la première colonne correspond à la commande 1 et la deuxième colonne à la commande 2. Ce critère se réécrit en utilisant l'espérance conditionnelle  $q^n$  calculée à la question précédente

$$\min_U \mathbb{E} \left[ \sum_0^n \sum_x q_x^n c_x^{U^n} \right].$$

4.  $q^n$  étant un processus de Markov l'équation de la programmation dynamique en horizon fini s'écrit

$$v^n(q) = \min\{v^{n+1}(qM^{10}) + q_2k_0, \\ v^{n+1}(\begin{pmatrix} 1 & 0 \end{pmatrix})q_2 + v^{n+1}(qM^{22}/q_1)q_1 + q \cdot c^2\},$$

où

$$v^n(q) = \min_U \mathbb{E} \left[ \sum_k^N \sum_x q_x^k c_x^{U^k} \mid q^n = q \right].$$

5. Pour résoudre ce problème il va falloir se ramener au contrôle d'une chaîne de Markov ayant un nombre fini d'états et donc discretiser l'ensemble  $q \in \mathbb{R}^2$ ,  $q^1 + q^2 = 1$ ,  $q^1, q^2 \geq 0$  en un ensemble fini à  $p + 1$  valeurs en  $\{[0, 1], [h, (p - 1)h], \dots, [1, 0]\}$  avec  $h = 1/p$  et approximer les sauts faits par l'espérance conditionnelle de telle sorte qu'elle reste sur le maillage. On est ensuite ramené à un problème classique de chaînes de Markov à états finis dans le cas de l'observation complète (l'espérance conditionnelle est en effet observé).

## 8. STRATÉGIES

On veut étudier la maintenance d'un système de production comprenant une suite infini d'étapes. On suppose que le système n'a que deux états possibles: soit il fonctionne ( $x = 1$ ), soit il ne fonctionne pas ( $x = 0$ ). Si le système ne fonctionne pas, on considère que cela coûte 1 (manque à gagner pour une étape). La politique de maintenance du système consiste à décider de tester ou non le système avant chaque étape de fonctionnement. Le prix du test (et de la remise en état éventuelle) est  $\alpha$ . Si on ne teste pas le système ( $u = 0$ ) avant une étape, il a une probabilité  $\epsilon$  de tomber en panne à la fin de cette étape (pendant l'étape correspondante il fonctionne). Si on le teste ( $u = 1$ ) et s'il ne fonctionne pas, il est remis en état pendant une étape (il ne fonctionne donc pas pendant cette étape mais il sera en état de marche à l'étape suivante). Si au cours du test on s'aperçoit qu'il fonctionne normalement, le système est entretenu, et on est assuré (probabilité 1) qu'il ne tombera pas en panne à la fin de l'étape du test.

Une politique stationnaire consiste à décider, en fonction de l'état du système à l'étape précédente, de tester ou non le système, et ceci indépendamment du numéro de l'étape considérée.

### 8.1. ENONCÉ

Pour chaque politique stationnaire donnez :

1. la matrice de transition de l'état du système,
2. la ou les mesures invariantes extrémales des chaînes de Markov correspondantes (celles dont le support est le plus petit possible lorsqu'il y en a plusieurs),
3. les équations de Kolmogorov donnant le coût de chacune de ces politiques, dans le cas d'une infinité d'étapes actualisées avec un taux  $\lambda > 0$ ,
4. les comportements asymptotiques (premier terme) de ces coûts lorsque le taux d'actualisation tend vers 0.

Quelle est la meilleure politique, dans le cas actualisé, avec un taux d'actualisation tendant vers 0, lorsque  $\epsilon$  est petit, beaucoup plus petit que  $\alpha$  (on supposera également que  $\alpha < 1$ )?

On détermine maintenant la politique optimale par la programmation dynamique.

1. Explicitez l'équation de la programmation dynamique déterminant la politique optimale, pour un nombre infini d'étapes, dans le cas actualisé avec le taux  $\lambda$ .
2. Donnez le système d'équations déterminant la politique optimale lorsque le taux d'actualisation tend vers 0 (on pourra supposer qu'à l'optimum la chaîne de Markov a une mesure invariante unique).

3. Calculez explicitement la politique optimale en adaptant l'algorithme de Howard au cas où le taux d'actualisation tend vers 0. On initialisera l'algorithme avec la politique consistant à ne jamais faire de maintenance, et on fera les itérations à la main.

## 8.2. CORRIGÉ

1. Il y a 4 stratégies stationnaires définies par les applications  $S$  de  $\{0, 1\}$  dans  $\{0, 1\}$ . Si  $S_{ij}$  désigne la fonction  $S$  prenant la valeur  $i$  en 0 et  $j$  en 1, et si l'on note  $M^{ij}$  la matrice de transition  $M^{S_{ij}}$ , on a les 4 matrices de transition :

$$M^{00} = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{pmatrix}, \quad M^{01} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$M^{10} = \begin{pmatrix} 0 & 1 \\ \epsilon & 1 - \epsilon \end{pmatrix}, \quad M^{11} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

donnant l'évolution du système pour chacune des politiques possibles.

2. Pour déterminer le nombre de mesures invariantes extrémales il suffit de compter le nombre de classes finales. Pour la politique  $S_{01}$  il y en a 2. Dans les autres cas il y en a une. Notons les, sous forme de vecteurs lignes (lorsqu'il y en a 2 on les donnera sous forme de matrice (2 lignes)). On appelle  $p^{ij}$  la matrice des mesures invariantes extrémales associées à la stratégie  $S_{ij}$ . Elle vérifie  $p^{ij} = p^{ij} M^{ij}$ . Ces mesures sont extrémales si et seulement si leurs supports sont des classes finales. On a :

$$p^{00} = (1 \quad 0), \quad p^{01} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$p^{10} = (\epsilon/(1 + \epsilon) \quad 1/(1 + \epsilon)), \quad p^{11} = (0 \quad 1).$$

3. Si l'on note  $v^{ij}$  la fonction valeur associée à la chaîne  $M^{ij}$  et aux coûts  $c^{ij}$  définis par

$$c^{00} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c^{01} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}, \quad c^{10} = \begin{pmatrix} 1 + \alpha \\ 0 \end{pmatrix}, \quad c^{11} = \begin{pmatrix} 1 + \alpha \\ \alpha \end{pmatrix},$$

les équations de Kolmogorov pour un problème actualisé avec un taux  $\lambda$  sont alors :

$$(1 + \lambda)v^{ij} = M^{ij} v^{ij} + c^{ij}.$$

4. Le premier terme du développement (lorsque  $\lambda$  tend vers 0) est donné (voir petite classe) par la moyenne au sens de la ou des mesures invariantes des coûts instantanés. Il est de la forme  $v^{ij}/\lambda$  pour la politique  $S_{ij}$  avec  $v^{ij}$  donné par  $p^{ij} c^{ij}$  (dans le cas d'une seule classe finale les

composantes de  $v^{ij}$  sont identiques, dans le cas de deux classes finales le coefficient dépend de la classe finale considérée). On obtient donc :

$$v^{00} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v^{01} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix},$$

$$v^{10} = \begin{pmatrix} (1 + \alpha)\epsilon/(1 + \epsilon) \\ (1 + \alpha)\epsilon/(1 + \epsilon) \end{pmatrix}, \quad v^{11} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}.$$

La meilleure politique est donc la politique  $S_{10}$ , c.a.d. que l'on répare lorsque le système est en panne, mais que l'on n'entretient pas lorsque le système marche. Cela paraît normal puisque le taux de panne est faible par rapport au prix de l'entretien.

1. L'équation de la programmation dynamique du problème actualisé s'écrit :

$$(1 + \lambda)v_0 = \min\{v_0 + 1, v_1 + 1 + \alpha\},$$

$$(1 + \lambda)v_1 = \min\{\epsilon v_0 + (1 - \epsilon)v_1, v_1 + \alpha\},$$

où  $v$  est la fonction valeur définie par

$$v_k = \min_U \mathbb{E} \left\{ \sum_{n=0}^{\infty} 1/(1 + \lambda)^{n+1} c_{X^n}^{U_n} \mid X_0 = k \right\},$$

$U = (U_0, \dots, U_n, \dots)$  et

$$c^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c^1 = \begin{pmatrix} 1 + \alpha \\ \alpha \end{pmatrix}.$$

2. On cherche un développement de  $v$  en  $\lambda$  sous la forme

$$v = v/\lambda + w + \lambda w^1 \dots$$

On obtient en identifiant les termes de même puissance en  $\lambda$

$$v_k = \min_u [M^u v]_k,$$

$$v_k + w_k = \min_{u \in U_k^*} [M^u w + c^u]_k,$$

avec

$$M^0 = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{pmatrix}, \quad M^1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

et  $U_k^* = \operatorname{argmin}_u [M^u v]_k$ . Dans le cas où la politique optimale conduit à une chaîne de Markov possédant une seule classe finale,  $v_k$  ne dépend plus de  $k$  car le seul vecteur propre à droite de la chaîne est

$$v \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

D'autre part  $U_k^* = \{0, 1\}$ . La condition d'optimalité s'écrit alors

$$v + w_k = \min_{u \in U_k^*} [M^u w + c^u]_k,$$

avec  $v$  constant,  $w$  défini à une constante près (on peut prendre  $w_1 = 0$ ).

3. L'algorithme de Howard s'adapte facilement.

1) Pour une stratégie  $S$  donnée on calcule un couple  $(v, w)$  satisfaisant

$$v = M^S v, v + w = M^S w + c^S.$$

2) A  $w$  on associe une nouvelle stratégie par

$$S : k \rightarrow u \in \operatorname{argmin}_u [M^u w + c^u]_k.$$

Sur l'exemple on obtient la suite

$$\begin{aligned} S^{00} \rightarrow v = 1, w_0 = 1/\epsilon, w_1 = 0 \rightarrow S^{11} \rightarrow v = \alpha, w_0 = 1, w_1 = 0 \rightarrow S^{10} \\ \rightarrow v = (1 + \alpha)\epsilon/(1 + \epsilon), w_0 = (1 + \alpha)/(1 + \epsilon), w_1 = 0 \rightarrow S^{10}. \end{aligned}$$

## 9. CLOWN ÉQUILIBRISTE

### 9.1. ENONCÉ

Dans ce problème, on étudie la dynamique d'un clown en équilibre sur un monocycle ainsi que ses façons de se stabiliser et de se déplacer.

Cette étude utilise les méthodes variationnelles pour construire un modèle simplifié linéaire. Les problèmes de stabilisation et de déplacement sont résolus, d'abord en calculant des commandes admissibles, puis optimisant les commandes par les méthodes de Pontryaguine et par la programmation dynamique.

Le système considéré est, bien sûr, une idéalisation du problème réel. La roue, appelée  $M$ , est supposée de rayon infinitésimal et de masse  $m$ . Le clown a une tête massive concentrant toute sa masse — supposée égale à 1 — en un point appelé  $P$ . Le corps est une barre rigide, sans masse, liant la tête  $P$  à la roue  $M$ , de longueur 1. Le clown, justement nommé "Pédalvit", vit dans l'univers plan de la roue — supposée verticale — et se déplace, en ligne droite, selon un axe orienté appelé  $Ox$ . On appelle:  $x$  l'abscisse de  $M$ ,  $\vartheta$  l'angle du corps avec la verticale  $Oy$ , orienté dans le sens des aiguilles d'une montre dans le repère  $(Ox, Oy)$ .

Moyennant ces simplifications, il est possible de résoudre ce problème assez complètement.

Dans la suite, l'accélération de la pesanteur est supposée égale à 1. Le temps est noté  $t$  ou  $s$ . On note  $\dot{f}(t)$  la dérivée de la fonction du temps  $f(t)$ .

9.1.1. LE CALCUL DES VARIATIONS COMME MOYEN DE MODÉLISATION. Rappelons que l'on peut déterminer la dynamique d'un système mécanique en calculant l'extremum sur l'ensemble des trajectoires possibles d'une "action". Cette action est définie comme l'intégrale, en temps le long d'une trajectoire, de l'énergie cinétique moins l'énergie potentielle du système considéré.

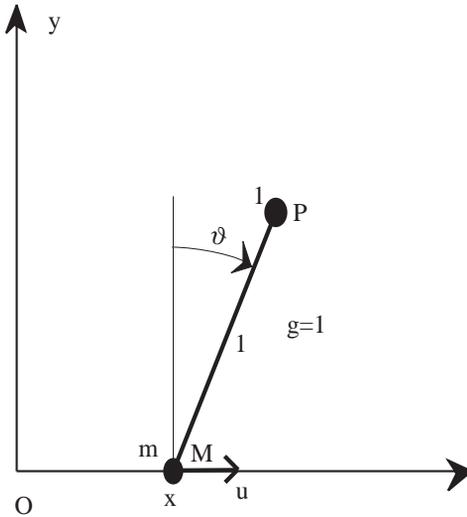


FIGURE 1. Le système mécanique.

En pédalant, le clown exerce une force  $F$  sur la roue  $M$ . Cette force est supposée — dans cette partie uniquement — dériver d'un potentiel  $U$  ne dépendant que de la position  $x$  de la roue.

MODÉLISATION SIMPLIFIÉE. Sachant que le potentiel  $U$  choisi par Pédalvit est  $ux$ , l'action du système est donnée par :

$$\mathcal{A}_c(x, \vartheta) = \frac{1}{2} \int_0^t \left\{ \left[ \frac{d}{dt}(x + \sin(\vartheta)) \right]^2 + \left[ \frac{d}{dt} \cos(\vartheta) \right]^2 + m\dot{x}^2 - 2 \cos(\vartheta) - 2ux \right\} (s) ds ,$$

définie sur l'espace des couples de fonctions  $s \mapsto (x(s), \vartheta(s))$ .

La dynamique du système complet étant trop compliquée pour être manipulée facilement, on considère les mouvements dans un voisinage de la verticale. L'action complète est alors approximée par l'action simplifiée suivante :

$$\mathcal{A}_s(x, \vartheta) = \frac{1}{2} \int_0^T [(\dot{x} + \dot{\vartheta})^2 + m\dot{x}^2 + \vartheta^2 - 2ux](s) ds .$$

1. Indiquer brièvement le sens physique de chacun des termes apparaissant dans l'action complète  $\mathcal{A}_c$ .
2. Discuter brièvement les approximations faites pour obtenir l'action simplifiée  $\mathcal{A}_s$ .

COERCIVITÉ DE L'ACTION SIMPLIFIÉE. On appelle  $H^1(0, T; \mathbb{R}^n)$  l'espace de Hilbert des fonctions  $f : [0, T] \mapsto \mathbb{R}^n$  de norme finie au sens du produit scalaire

$$(f, g) = \int_0^T \sum_{i=1}^n [f_i g_i + \dot{f}_i \dot{g}_i](t) dt .$$

On rappelle qu'une fonction de  $H^1(0, T; \mathbb{R})$  est continue que

$$f(t) = f(0) + \int_0^t \dot{f}(s) ds$$

et que les fonctions régulières sont denses dans  $H^1(0, T; \mathbb{R})$ . On a de plus les inégalités suivantes :

$$|f(t_2) - f(t_1)| \leq \sqrt{|t_2 - t_1|} \|f\|_{H^1},$$

$$\|f\|_{\infty} \leq \left( \sqrt{T} + \sqrt{1/T} \right) \|f\|_{H^1},$$

où  $\|\cdot\|_{\infty}$  désigne la norme du sup.

1. Si l'on appelle

$$H_0^1(0, T; \mathbb{R}^2) = \{X = (x, \vartheta) \mid X(0) = 0, X \in H^1(0, T; \mathbb{R}^2)\},$$

montrer que l'application  $X \in H_0^1 \mapsto \mathcal{A}_s(X) \in \mathbb{R}$  est  $\alpha$ -convexe pour  $T$  pas trop grand.

2. Montrer que cette action n'est plus  $\alpha$ -convexe pour  $T = +\infty$  en exhibant une suite de fonctions  $(X_n)$  vérifiant :

$$X_n \in H_0^1(0, \infty; \mathbb{R}^2), \quad \lim_n \|X_n\|_{H^1} = +\infty \quad \lim_n \mathcal{Q}(X_n) = 0,$$

où

$$\mathcal{Q}(X) \stackrel{\text{def}}{=} \frac{1}{2} \int_0^{\infty} [(\dot{x} + \dot{\vartheta})^2 + m\dot{x}^2 + \vartheta^2](s) ds.$$

3. En déduire que le problème du calcul de la moindre action admet une solution unique pour  $T$  pas trop grand.

LES ÉQUATIONS DU MOUVEMENT. En minimisant les actions correspondantes déduire :

1. les équations du mouvement simplifié :

$$(1 + m)\ddot{x} + \ddot{\vartheta} = -u, \quad \ddot{x} + \ddot{\vartheta} - \vartheta = 0.$$

2. les équations du mouvement complet.

Dans la suite, on ne s'intéressera qu'au mouvement simplifié et le mot "mouvement" signifiera toujours "mouvement simplifié".

Dans les sections suivantes, on fait dépendre la variable de commande  $u$  — la force de pédalage de Pédalvit — de  $x$  et de  $\vartheta$ . Cette force ne dérive plus d'un potentiel. Les équations du mouvement restent valables — mécanique newtonienne — bien qu'elles ne puissent plus s'obtenir par le principe de la moindre action.

9.1.2. LA COMMANDE NON OPTIMALE. STABILITÉ DU MOUVEMENT. On appelle mode propre du système les nombres complexes  $\lambda_i$  tels qu'il existe  $(x_i(t), \vartheta_i(t))(t)$  polynômes en  $t$  avec

$$X_i(t) \stackrel{\text{def}}{=} (x_i(t), \vartheta_i(t))e^{\lambda_i t}$$

solution du mouvement.

Toutes les solutions du mouvement étant des combinaisons linéaires de quatre de ces solutions particulières indépendantes, il est clair que, si tous les modes ont des parties réelles strictement négatives, toutes les composantes du mouvement convergent vers 0 — lorsque  $t$  tend vers l'infini — avec une vitesse exponentielle. On dira alors que le mouvement est *exponentiellement stable*. Si un des modes a une partie réelle strictement positive, on dira que *le système est exponentiellement instable*.

Si tous les modes sont imaginaires et les polynômes correspondants sont des constantes, on dira que le mouvement est purement *oscillant*.

On considère des commandes de la forme  $u = cx + d\vartheta$ .

1. Donner un système d'équations caractérisant les modes propres (on se placera dans le cas de 4 modes propres distincts (les polynômes  $x_i(t)$  et  $\vartheta_i(t)$  sont alors des constantes)).
2. Montrer que, quelque soient  $c$  et  $d$ , le mouvement correspondant est, soit exponentiellement instable, soit purement oscillant.

Bien qu'ayant compris comment ne pas tomber en obtenant un régime purement oscillant, Pédalvit aimerait pouvoir amortir les oscillations. Pour cela, il est donc obligé d'adapter sa force de pédalage non seulement à  $x$  et  $\vartheta$  mais aussi à  $\dot{x}$  et  $\dot{\vartheta}$ .

FORMES CANONIQUES ET APPLICATIONS. Le système commandé peut se mettre sous une forme canonique, dite de Brunowski, grâce à laquelle il est facile de stabiliser le système ou de déterminer une commande transférant le système d'un point à un autre dans l'espace d'état.

1. Montrer, par un changement de commande — défini par  $u = f(v, y)$  où  $y = (x, \dot{x}, \vartheta, \dot{\vartheta})$  désigne le vecteur d'état du système et  $v$  la nouvelle commande — que l'on peut ramener les équations du mouvement à :

$$z^{(4)} = v, \tag{9.1}$$

où  $z$  désigne l'abscisse de la tête  $P$  de Pédalvit. On explicitera le changement de commande utilisé.

2. Grâce à cette forme canonique, caractériser une commande<sup>6</sup>  $u(t)$  assurant le déplacement de Pédalvit de la position verticale arrêtée à l'abscisse 1, à l'instant 0, à la position verticale arrêtée à l'abscisse 0, à l'instant 1.

<sup>6</sup>Lorsque la commande  $u$  ne dépend que du temps on dit que la commande est en *boucle ouverte*.

3. Grâce à cette forme canonique, calculer une commande<sup>7</sup>

$$u = \phi(z, \dot{z}, \ddot{z}, z^{(3)}) ,$$

avec  $\phi$  linéaire, telle que tous les modes propres, du système ainsi commandé, soient  $-1$ .

Dans la suite, on utilisera cette forme canonique pour optimiser les déplacements et stabiliser le système de façon optimale au sens de critères que l'on précisera.

### 9.1.3. DÉPLACEMENT OPTIMAL PAR LA MÉTHODE DE PONTRYAGUINE.

En utilisant la forme canonique (9.1), on veut calculer la commande assurant le déplacement de Pédalvit de la position verticale arrêtée, à l'abscisse 1, à la position verticale arrêtée à l'abscisse 0, en un temps  $T$ , tout en minimisant ses efforts c.à.d. en minimisant :

$$J(v) = \frac{1}{2} \int_0^T v^2(s) ds .$$

Pour simplifier les calculs, on a supposé, que Pédalvit reste à peu près vertical et donc que  $v^2$  est voisin de  $u^2$  — mesurant plus précisément son aversion à faire des efforts.

1. En utilisant la méthode de Pontryaguine donner les conditions nécessaires et suffisantes d'optimalité du mouvement correspondant.
2. Caractériser la commande optimale en boucle ouverte,  $v(t)$ , assurant le déplacement souhaité.

### 9.1.4. STABILISATION OPTIMALE PAR LA PROGRAMMATION DYNAMIQUE.

En utilisant la forme canonique (9.1), Pédalvit veut calculer une commande le stabilisant à l'origine de façon optimale au sens d'un critère faisant un compromis entre ses efforts et la qualité de la stabilisation. Pour cela, il minimise le critère

$$J(v) = \frac{1}{2} \int_0^\infty [v^2 + \gamma z^2](s) ds, \quad \gamma > 0 .$$

1. Montrer que ce critère est  $\alpha$ -convexe sur  $H_0^4(0, \infty; \mathbb{R})$  sous espace fermé des fonctions de  $H^4(0, \infty; \mathbb{R})$  ayant leurs valeurs ainsi que leur 3 premières dérivées nulles à l'origine. L'espace de Hilbert  $H^4(0, \infty; \mathbb{R})$  est l'espace des fonctions de norme finie au sens du produit scalaire

$$(f, g) = \int_0^\infty \left( \sum_{i=0}^4 f^{(i)} g^{(i)} \right) (t) dt .$$

De ce résultat découle que l'optimum existe et est unique.

---

<sup>7</sup>Lorsque la commande dépend de l'état on dit qu'elle est en *boucle fermée* ou en "feedback".

2. Dans la suite le vecteur d'état du système sous forme canonique est noté  $Z$  :

$$Z(t) = (z, \dot{z}, \ddot{z}, z^{(3)})(t) .$$

En adaptant le théorème correspondant du cours, montrer que la fonction :

$$w(z_0, z_1, z_2, z_3) \stackrel{\text{def}}{=} \inf_v \{J(v) \mid Z(0) = (z_0, z_1, z_2, z_3)\} ,$$

vérifie l'équation de la programmation dynamique :

$$\frac{\partial w}{\partial z_0} z_1 + \frac{\partial w}{\partial z_1} z_2 + \frac{\partial w}{\partial z_2} z_3 - \frac{1}{2} \left( \frac{\partial w}{\partial z_3} \right)^2 + \frac{\gamma}{2} (z_0)^2 = 0 , \quad (9.2)$$

$$w(0, 0, 0, 0) = 0 .$$

Pour cela on admettra que (9.2) admet une unique solution positive régulière.

EQUATION DE RICCATI ALGÈBRE. Montrer que la solution positive de l'équation de la programmation dynamique (9.2) est quadratique :  $w(\zeta) = (1/2)\zeta' P \zeta$  avec  $P \geq 0$  et  $\zeta = (z_0, z_1, z_2, z_3)$ .

1. Par substitution dans (9.2) donner l'équation — dite de Riccati — vérifiée par  $P$ .
2. Montrer que la commande optimale est :

$$v^*(z_0, z_1, z_2, z_3) = -p_{14}z_0 - p_{24}z_1 - p_{34}z_2 - p_{44}z_3 .$$

3. Montrer que le système dynamique optimal — gouverné par ce feedback optimal — s'écrit :

$$\dot{Z} = AZ , \quad (9.3)$$

avec  $A$  une matrice — dont les coefficients dépendent de  $P$  — que l'on explicitera .

4. Calculer le polynôme caractéristique de  $A$  en fonction des coefficients de  $P$ .

INTERPRÉTATION DES VARIABLES DUALES.

1. Si l'on note

$$W(t) = \text{grad}(w)(Z(t)) ,$$

et  $\Xi(t) = (Z, W)(t)$ , montrer que  $\Xi$  satisfait le système hamiltonien  $\dot{\Xi} = \mathcal{H}\Xi$ , avec

$$\mathcal{H} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix} .$$

2. Montrer que ce système hamiltonien est le système d'optimalité obtenu par la méthode de Pontryaguine appliquée à ce même problème de stabilisation.
3. Calculer les modes propres de ce système hamiltonien.
4. Montrer que le changement de variables  $Z = \chi$ ,  $W = P\chi + \lambda$ , où  $P$  est solution de l'équation de Riccati, conduit à un système dynamique en  $(\chi, \lambda)$  bloc-triangulaire.
5. Expliciter la factorisation du polynôme caractéristique de  $\mathcal{H}$  obtenu par ce changement de variable.

FEEDBACK OPTIMAL ET PLACEMENT OPTIMAL DES MODES PROPRES. Pour que le coût optimal soit fini quelque soit la condition initiale il est nécessaire et suffisant que tous les modes propres du système optimal (9.3) soient à partie réelle strictement négative.

1. Dédire de cette remarque et de la factorisation précédente les valeurs des coefficients de  $P$  nécessaires au calcul explicite de la commande optimale  $v^*$ .
2. Expliciter la commande optimale  $v^*$  comme fonction de l'état.
3. Expliquer comment évoluent les modes propres du système optimal en fonction du paramètre  $\gamma$ .

## 9.2. CORRIGÉ

9.2.1. LE CALCUL DES VARIATIONS COMME MOYEN DE MODÉLISATION. MODÉLISATION SIMPLIFIÉE.

1. Le terme  $1/2([\frac{d}{dt}(x + \sin(\vartheta))]^2 + [\frac{d}{dt} \cos(\vartheta)]^2)$  correspond à l'énergie cinétique de la tête,  $1/2m\dot{x}^2$  à l'énergie cinétique de la roue,  $\cos(\vartheta)$  à l'énergie potentielle de la tête,  $ux$  à l'énergie potentielle modélisant la force de pédalage.
2. Les approximations faites consistent à négliger la partie de l'énergie cinétique correspondant à la vitesse verticale de la tête et à approximer  $\sin(\vartheta)$  par  $\vartheta$  et  $\cos(\vartheta)$  par  $1 - \frac{1}{2}\vartheta^2$ . Ces approximations sont raisonnables pour des mouvements dans le voisinage de la verticale puisqu'elles ne gardent que les termes d'ordre 2 dans un développement par rapport à  $\vartheta$ .

COERCIVITÉ DE L'ACTION. Les preuves des deux estimations <sup>8 9</sup> données se font facilement bien qu'elle ne soient pas demandées.

1. L'application  $X \in H^1(0, T; \mathbb{R}^2) \mapsto \mathcal{A}_s(X) \in \mathbb{R}$  est  $\alpha$ -convexe pour  $T$  trop grand. En effet, de l'égalité  $2ab = a^2 + b^2 - (a - b)^2$  on déduit :

$$\begin{aligned} \mathcal{A}_s(\lambda X_1 + (1 - \lambda)X_2) &= \lambda \mathcal{A}_s(X_1) + (1 - \lambda)\mathcal{A}_s(X_2) \\ &\quad - \lambda(1 - \lambda)\mathcal{Q}(X_1 - X_2), \end{aligned}$$

où  $\mathcal{Q}$  désigne la partie quadratique de  $\mathcal{A}_s$ . Il faut alors montrer la minoration

$$\mathcal{Q}(X_1 - X_2) \geq \alpha \|X_1 - X_2\|_{H^1}$$

pour pouvoir conclure. On utilise alors  $(a+b)^2 + a^2 \geq 1/4(a^2 + b^2)$  et l'inégalité :  $\|x\|_{L^2} \leq T \|\dot{x}\|_{L^2}$  pour  $x \in H^1(0, T; \mathbb{R})$  avec  $x(0) = 0$ , pour montrer qu'il existe  $k > 0$  tel que :

$$\mathcal{Q}(X) \geq k \int_0^T (x^2 + \vartheta^2 + (\dot{x})^2 + (\dot{\vartheta})^2)(t) dt .$$

2. La suite de fonctions  $X_n(t)$  telles  $\vartheta_n(t) = 0$  et

$$x_n(t) = \begin{cases} t/n & \text{si } 0 \leq t \leq n, \\ 1 & \text{si } n \leq t \leq 2n, \\ 1 - (t - 2n)/n & \text{si } 2n \leq t \leq 3n, \\ 0 & \text{si } t \geq 3n, \end{cases}$$

vérifie  $\lim_n \|X_n\|_{H^1(0, \infty)} = +\infty$ ,  $\lim_n \mathcal{Q}(X_n) = 0$ . Elle appartient clairement à  $H_0^1(0, \infty; \mathbb{R}^2)$ .

3. Le problème de la moindre action, dans cet exemple, et pour  $T$  fini, et  $X(0) = 0$  se ramène à minimiser une fonction  $\alpha$ -convexe sur  $H_0^1(0, T; \mathbb{R}^2)$  qui est un ensemble convexe fermé (grâce au fait que l'application qui à  $X \in H^1 \mapsto X(0) \in \mathbb{R}^2$  est continue) non vide. Il existe donc un minimum unique.

<sup>8</sup>Preuve de  $|f(t_2) - f(t_1)| \leq \sqrt{|t_2 - t_1|} \|f\|_{H^1}$ . On a :

$$(f(t_2) - f(t_1))^2 = \left( \int_{t_1}^{t_2} \dot{f}(s) ds \right)^2 \leq |t_2 - t_1| \int_{t_1}^{t_2} (\dot{f}(s))^2 ds \leq |t_2 - t_1| \|f\|_{H^1}^2 .$$

<sup>9</sup>Preuve de  $|f|_\infty \leq (\sqrt{T} + \sqrt{1/T}) \|f\|_{H^1}$ .

$$\begin{aligned} \left( f(t) - 1/T \int_0^T f(s) ds \right)^2 &= \left( 1/T \int_0^T (f(t) - f(s)) ds \right)^2 \leq 1/T \int_0^T [f(t) - f(s)]^2 ds \\ &\leq 1/T \int_0^T |t - s| \|\dot{f}\|_{L^2}^2 ds \leq |T| \|\dot{f}\|_{L^2}^2 . \end{aligned}$$

D'autre part, on a

$$\left( 1/T \int_0^T f(s) ds \right)^2 \leq 1/T \int_0^T (f(s))^2 ds .$$

L'inégalité triangulaire donne alors le résultat.

## LES ÉQUATIONS DU MOUVEMENT.

1. On obtient les équations du mouvement simplifié en minimisant l'action simplifiée. Notons

$$H_{00}^1(0, T; \mathbb{R}^2) = \{X \mid X \in H^1(0, T), X(0) = 0, X(T) = 0\}.$$

On cherche  $X \in H^1(0, T; \mathbb{R}^2)$ , vérifiant  $X(0) = (x_0, \vartheta_0)$  et  $X(T) = (x_T, \vartheta_T)$ , telle que

$$(\dot{\mathcal{A}}_s(X), Y) = 0, \quad \forall Y \in H_{00}^1(0, T; \mathbb{R}^2),$$

avec

$$(\dot{\mathcal{A}}_s(X), Y) \stackrel{\text{def}}{=} \int_0^T [(\dot{x} + \dot{\vartheta})(\dot{Y}_1 + \dot{Y}_2) + m\dot{x}\dot{Y}_1 + \vartheta Y_2 - uY_1](s) ds.$$

En supposant que les fonctions soient suffisamment régulières pour que l'on puisse intégrer par partie, on obtient :

$$\int_0^T [(1+m)\ddot{x} + \ddot{\vartheta} + u]Y_1 + (\ddot{x} + \ddot{\vartheta} - \vartheta)Y_2(s) ds = 0.$$

Ce qui entraîne les équations annoncées du mouvement .

2. Les équations du mouvement complet, obtenues par la même technique, sans se préoccuper des conditions aux limites, sont :

$$(1+m)\ddot{x} + \frac{d}{dt}(\cos(\vartheta)\dot{\vartheta}) = -u, \quad \ddot{x} \cos \vartheta + \ddot{\vartheta} - \sin(\vartheta) = 0.$$

## 9.2.2. LA COMMANDE NON OPTIMALE. STABILITÉ DU MOUVEMENT.

1. On cherche des solutions de la forme  $(x, \vartheta)e^{\lambda t}$  avec  $(x, \vartheta) \in \mathbb{R}^2$ . En substituant dans les équations du mouvement, on obtient :

$$\begin{aligned} ((1+m)x + \vartheta)\lambda^2 &= -cx - d\vartheta, \\ (x + \vartheta)\lambda^2 - \vartheta &= 0. \end{aligned}$$

ce qui se réécrit

$$A \begin{pmatrix} x \\ \vartheta \end{pmatrix} = 0,$$

avec

$$A \stackrel{\text{def}}{=} \begin{pmatrix} (1+m)\lambda^2 + c & d + \lambda^2 \\ \lambda^2 & \lambda^2 - 1 \end{pmatrix}$$

et donc  $\lambda$  devra être solution de  $\det A = 0$ .

2. La fonction  $\det A$  est un polynôme du second degré en  $\mu \stackrel{\text{def}}{=} \lambda^2$ , et donc :

- soit les racines en  $\mu$  sont toutes les deux réelles négatives et le système est oscillant,
- soit une des racines n'est pas réelle négative, et on a alors, forcément, une racine en  $\lambda$  à partie réelle positive, et le système est instable.

1. En utilisant les variables  $z = x + \vartheta$  et  $\vartheta$  les équations du mouvement s'écrivent

$$\ddot{z} + m(\ddot{z} - \ddot{\vartheta}) = -u, \quad \ddot{z} - \ddot{\vartheta} = 0 .$$

ce système se réécrit

$$(1 + m)\ddot{\vartheta} - m\ddot{\vartheta} = -u, \quad \ddot{z} = \ddot{\vartheta} ,$$

puis

$$\ddot{\vartheta} = (u + (1 + m)\ddot{\vartheta})/m, \quad \ddot{z} = \ddot{\vartheta} ,$$

et donc avec le changement de commande  $v = (u + (1 + m)\ddot{\vartheta})/m$

$$z^{(4)} = v .$$

On a de plus la relation intéressante  $\ddot{z} = \ddot{\vartheta}$ .

2. Pour déplacer Pédalvit de la position verticale arrêtée à l'abscisse 1 à la position verticale arrêtée à l'abscisse 0, on peut chercher une trajectoire de  $z(t)$  polynômiale de degré 7. Les conditions à l'instant 0 imposent  $z(t) = 1 + pt^4 + qt^5 + rt^6 + st^7$ . Les conditions à l'instant 1 imposent aux inconnues ( $p, q, r, s$ ) de vérifier le système linéaire :

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 5 & 6 & 7 \\ 12 & 20 & 30 & 42 \\ 24 & 60 & 120 & 210 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \\ s \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

Il suffit de dériver 4 fois ce polynôme pour obtenir  $v$ , le changement de commande donne alors la commande  $u$ .

3. Pour calculer une commande  $u = \phi(x, \dot{x}, \vartheta, \dot{\vartheta})$ , linéaire, tel que tous les modes propres du système commandé deviennent  $-1$ . On commence par le calculer comme fonction de  $z$  et de ses dérivées. Considérons le feedback  $v = -(fz + g\dot{z} + h\ddot{z} + iz^{(3)})$ , les modes propres du système deviennent les solutions de

$$q(\lambda) \stackrel{\text{def}}{=} \lambda^4 + i\lambda^3 + g\lambda^2 + h\lambda + f = 0 .$$

Mais on veut qu'il y ait un seul mode propre  $-1$  il faut donc

$$q(\lambda) = (\lambda + 1)^4$$

d'où on déduit les coefficients. On a  $u = mv - (1 + m)\ddot{\vartheta}$  et donc le feedback cherché est  $u = -m(z + 4\dot{z} + 6\ddot{z} + 4z^{(3)}) - (1 + m)\ddot{z}$  avec  $z = x + \vartheta$ . Ainsi, on a obtenu un feedback dépendant que de la position de la tête de Pedalvit et de ses dérivées.

### 9.2.3. DÉPLACEMENT OPTIMAL PAR LA MÉTHODE DE PONTRYAGUINE.

On minimise :

$$J(v) = \frac{1}{2} \int_0^1 v^2(s) ds .$$

sous la contrainte

$$z^{(4)} = v .$$

1. Pour utiliser la méthode de Pontryaguine il faut mettre l'équation  $z^{(4)} = v$  sous forme d'un système du premier ordre

$$\frac{d}{dt} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} v .$$

La condition nécessaire et suffisante d'optimalité, en notant  $(p^{(3)}, \ddot{p}, \dot{p}, p)$  les variables duales, s'écrit :

$$\frac{d}{dt} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} v ,$$

$$z(0) = 1, \quad z(1) = \dot{z}(0) = \dot{z}(1) = \ddot{z}(0) = \ddot{z}(1) = z^{(3)}(0) = z^{(3)}(1) = 0 ,$$

$$v = -p ,$$

$$\frac{d}{dt} \begin{pmatrix} p^{(3)} \\ \ddot{p} \\ \dot{p} \\ p \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} p^{(3)} \\ \ddot{p} \\ \dot{p} \\ p \end{pmatrix} .$$

2. Le système d'optimalité se réécrit donc  $z^{(8)} = 0$  avec les conditions aux limites

$$z(0) = 1, \quad z(1) = \dot{z}(0) = \dot{z}(1) = \ddot{z}(0) = \ddot{z}(1) = z^{(3)}(0) = z^{(3)}(1) = 0 .$$

$z$  est donc un polynôme de degré 7. La commande calculée au paragraphe précédent est donc optimale.

On remarque qu'en utilisant le calcul de variation on aurait obtenu plus rapidement le résultat  $z^{(8)} = 0$ .

9.2.4. STABILISATION OPTIMALE PAR LA PROGRAMMATION DYNAMIQUE. On minimise :

$$J(v) = \frac{1}{2} \int_0^\infty [v^2 + \gamma z^2](s) ds .$$

sous la contrainte

$$z^{(4)} = v .$$

par la méthode de la programmation dynamique.

1. Pour montrer la coercivité de  $J(v)$  il suffit de montrer que  $J(v)$  définit une norme équivalente à  $H_0^4(0, \infty; \mathbb{R})$  et donc il suffit de montrer qu'il existe

$$\int_0^\infty |\ddot{f}|^2(t) dt \leq kJ(f), \quad \int_0^\infty |\dot{f}|^2(t) dt \leq kJ(f) ,$$

$$\int_0^\infty |f^{(3)}|^2(t) dt \leq kJ(f) .$$

mais on a

$$\int_0^{+\infty} |\ddot{f}|^2(t) dt = \int_0^{+\infty} [ff^{(4)}](t) dt .$$

Mais par l'inégalité de Cauchy-Schwarz on a :

$$\left\{ \int_0^{+\infty} [ff^{(4)}](t) dt \right\}^2 \leq \|f\|_{L^2}^2 \|f^{(4)}\|_{L^2}^2 \leq k(J(f))^2 .$$

De même on a :

$$\int_0^{+\infty} |\dot{f}|^2(t) dt = - \int_0^{+\infty} [f\ddot{f}](t) dt .$$

D'où la deuxième majoration en utilisant l'inégalité de Cauchy-Schwarz. Le même genre de raisonnement pour la troisième estimation montre le résultat.

2. Soit la fonction  $w(z_0, z_1, z_2, z_3)$  solution de l'équation de la programmation dynamique

$$\frac{\partial w}{\partial z_0} z_1 + \frac{\partial w}{\partial z_1} z_2 + \frac{\partial w}{\partial z_2} z_3 - \frac{1}{2} \left( \frac{\partial w}{\partial z_3} \right)^2 + \frac{\gamma}{2} z_0^2 = 0 ,$$

$$w(0, 0, 0, 0) = 0 ,$$

et soit  $v$  une commande admissible, c.à.d. telle que  $z \in H^{(4)}(0, \infty; \mathbb{R})$ , on a alors

$$\begin{aligned} \frac{d}{dt} w(Z(t)) &= \left[ \frac{\partial w}{\partial z_0} \dot{z} + \frac{\partial w}{\partial z_1} \ddot{z} + \frac{\partial w}{\partial z_2} z^{(3)} + \frac{\partial w}{\partial z_3} v \right](t) \\ &\geq -\frac{1}{2} (v^2 + \gamma z^2)(t) . \end{aligned}$$

Cette dernière inégalité provient de l'équation dynamique en se rappelant que

$$-\frac{1}{2} \left( \frac{\partial w}{\partial z_3} \right)^2 \leq \frac{\partial w}{\partial z_3} v + \frac{v^2}{2} .$$

On obtient alors :

$$w(Z(\infty)) - w(Z(0)) \geq - \int_0^{\infty} \frac{1}{2} (v^2 + \gamma z^2)(t) dt .$$

Et donc  $w(Z(0)) \leq \int_0^{\infty} \frac{1}{2} (v^2 + \gamma z^2)(t) dt$  puisque la trajectoire  $Z$  est nulle à l'infini. Si on avait pris pour  $v$  la commande optimale on aurait obtenu une égalité.

Enfin si  $Z(0) = (0, 0, 0, 0)$  la commande  $v = 0$  conduit à un coût nul et donc  $w(0, 0, 0, 0) = 0$ .

#### EQUATION DE RICCATI ALGÈBRIQUE.

1. On cherche une solution de la forme  $w(\zeta) = (1/2)\zeta' P \zeta$ . La matrice  $P$  doit vérifier alors l'équation de Riccati algébrique :

$$A'P + PA - PB'BP + \gamma C'C = 0 ,$$

avec

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad C = (1 \ 0 \ 0 \ 0).$$

2. La commande optimale vaut

$$v^* = -B'P \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix},$$

ce qui donne le résultat demandé puisque  $P$  est symétrique.

3. Le système dynamique à l'optimum s'écrit alors

$$\frac{d}{dt} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -p_{14} & -p_{24} & -p_{34} & -p_{44} \end{pmatrix} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \end{pmatrix}.$$

4. Le polynôme caractéristique de  $A$  vaut

$$s^4 + p_{44}s^3 + p_{34}s^2 + p_{24}s + p_{14}.$$

#### INTERPRÉTATION DES VARIABLES DUALES.

1. Le système d'équation linéaire  $\dot{\Xi} = \mathcal{A}\Xi$  vérifié par le couple  $\Xi(t) = (Z, W)(t)$  avec

$$Z(t) = (z, \dot{z}, \ddot{z}, z^{(3)})(t), \quad W(t) = \text{grad}(w)(Z(t)),$$

est obtenu en dérivant l'équation de la programmation dynamique par rapport  $z_i$  et en remarquant que

$$\frac{d}{dt} W_i(t) = \frac{d}{dt} \frac{\partial w}{\partial z_i}(Z(t)) = \frac{\partial^2 w}{\partial z_0 \partial z_i} \dot{z} + \frac{\partial^2 w}{\partial z_1 \partial z_i} \ddot{z} + \dots - \frac{\partial^2 w}{\partial z_3 \partial z_i} \frac{\partial w}{\partial z_3},$$

puisque la commande optimale est  $v^* = -\frac{\partial w}{\partial z_3}$ . On obtient alors

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

2. Le système dynamique obtenu par la méthode de Pontryaguine appliquée à ce même problème (en utilisant les notations  $(w^{(3)}, \ddot{w}, \dot{w}, w)$  pour les variables duales) est :

$$\frac{d}{dt} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \\ w^{(3)} \\ \ddot{w} \\ \dot{w} \\ w \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} z \\ \dot{z} \\ \ddot{z} \\ z^{(3)} \\ w^{(3)} \\ \ddot{w} \\ \dot{w} \\ w \end{pmatrix}.$$

3. Les modes propres de ce système dynamique sont les racines de

$$s^8 + \gamma = 0,$$

qui se calculent aisément à partir des racines huitièmes de  $-1$ .

4. Le changement de variables  $Z = \chi$ ,  $W = P\chi + \lambda$ , où  $P$  est solution de l'équation de Riccati conduit au système

$$\frac{d}{dt} \begin{pmatrix} \chi \\ \chi_1 \\ \chi_2 \\ \chi_3 \\ \lambda \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -p_{14} & -p_{24} & -p_{34} & -p_{44} & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{14} \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & p_{24} \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & p_{34} \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & p_{44} \end{pmatrix} \begin{pmatrix} \chi \\ \chi_1 \\ \chi_2 \\ \chi_3 \\ \lambda \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

5. La factorisation du polynôme caractéristique de  $\mathcal{A}$  obtenue par ce changement de variable est :

$$s^8 + \gamma = (s^4 + p_{44}s^3 + p_{34}s^2 + p_{24}s + p_{14})(s^4 - p_{44}s^3 + p_{34}s^2 - p_{24}s + p_{14}).$$

FEEDBACK OPTIMAL ET PLACEMENT OPTIMAL DES MODES PROPRES.

Tous les modes propres du système optimal doivent être à partie réelle négative. Les racines de  $s^4 + p_{44}s^3 + p_{34}s^2 + p_{24}s + p_{14} = 0$  doivent donc avoir une partie réelle négative.

1. Après avoir explicité les racines de  $s^8 + \gamma = 0$ , la factorisation précédente donne

$$\begin{aligned} s^4 + p_{44}s^3 + p_{34}s^2 + p_{24}s + p_{14} &= s^4 + \left( \sqrt{2 + \sqrt{2}} + \sqrt{2 - \sqrt{2}} \right) \gamma^{1/8} s^3 \\ &\quad + (2 + \sqrt{2}) \gamma^{1/4} s^2 \\ &\quad + \left( \sqrt{2 + \sqrt{2}} + \sqrt{2 - \sqrt{2}} \right) \gamma^{3/8} s \\ &\quad + \gamma^{1/2}. \end{aligned}$$

2. La commande optimale est donc :

$$\begin{aligned}
 v^* = & - \left( \sqrt{2 + \sqrt{2}} + \sqrt{2 - \sqrt{2}} \right) \gamma^{1/8} \ddot{z}^{(3)} \\
 & - (2 + \sqrt{2}) \gamma^{1/4} \ddot{z} \\
 & - \left( \sqrt{2 + \sqrt{2}} + \sqrt{2 - \sqrt{2}} \right) \gamma^{3/8} \dot{z} \\
 & - \gamma^{1/2} z .
 \end{aligned}$$

3. Les modes propres du système optimal partent à l'infini avec  $\gamma$  en restant sur un cercle du plan complexe de rayon  $\gamma^{1/8}$  et avec des angles  $(5\pi/8, 7\pi/8, 9\pi/8, 11\pi/8)$ .



## Notations

$\mathbb{N}$	nombres entiers
$\mathbb{Z}$	nombres entiers relatifs
$\mathbb{R}$	nombres réels
$\mathbb{R}_{\max}$	dioïde $\{\mathbb{R}, \max, +\}$
$\mathbb{R}_{\min}$	dioïde $\{\mathbb{R}, \min, +\}$
$\oplus$	addition dans $\mathbb{R}_{\max}$
$\otimes$	multiplication dans $\mathbb{R}_{\max}$ pour les matrices $(A \otimes B)_{ij} = \inf_k (A_{ik} + B_{kj})$
$\emptyset$	ensemble vide
$\varepsilon$	zero de $\mathbb{R}_{\max}$
$e$	element identité dans $\mathbb{R}_{\max}$ ou $\mathbb{R}_{\min}$ ou dans les matrices carrées valeur dans $\mathbb{R}_{\max}$ ou $\mathbb{R}_{\min}$ , entrée equivalente dans le chapitre sur la forme produit
$\mathcal{T}$	ensemble des temps
$\mathcal{F}$	ensemble des commandes
$\mathcal{G}$	ensemble des observations
$\mathcal{E}$	ensemble des états
$\mathcal{R}$	ensemble des stratégies relaxées
$\mathcal{S}$	ensemble des stratégies déterministes
$\mathcal{S}_F$	ensemble des stratégies markoviennes
$\mathcal{S}_O$	ensemble des commandes boucle ouverte
$\mathcal{S}_{FP}$	ensemble des feedback à priori
$\mathcal{R}(A)$	image de $A$
$\mathcal{N}(A)$	noyau de $A$
$A^*$	$e \oplus A \oplus A^2 \oplus \dots$
$F$	nombre de commandes possibles
$G$	nombre d' observations possibles
$E$	nombre d'états possibles
$X^n$	état d'une chaîne de Markov
$Y^n$	observation d'une chaîne de Markov observée
$U^n$	commande d'une chaîne de Markov commandée
$M_{xx'}$	probabilité conditionnelle de transiter de l'état $x$ à l'état $x'$
$M_{xx'}^\varepsilon$	probabilité conditionnelle de transiter de l'état $x$ à l'état $x'$ d'une chaîne perturbée par une perturbation de valeur $\varepsilon$
$M_{xx'}^u$	probabilité conditionnelle de transiter de l'état $x$ à l'état $x'$ lorsque la commande est $u$

$M_{xx'}^y$	probabilité conditionnelle de transiter de l'état $x$ à l'état $x'$ et d'observer $y$
$M_{xx'}^{uy}$	probabilité conditionnelle de transiter de l'état $x$ à l'état $x'$ et d'observer $y$ lorsque la commande est $u$
$p^o$	loi initiale
$p^\infty$	mesure invariante
$p^{s\infty}$	mesure invariante pour la stratégie $s$
$H(v)$	hamiltonien
$I$	matrice identité ou nombre de files d'attente dans un système
$A$	générateur d'une chaîne de Markov $M - I$
$\lambda$	taux d'actualisation
$\mu$	$\lambda = \mu\varepsilon$ taux d'actualisation dans le cas perturbé
$r$	matrice de routage
$r^*$	si $r$ est une matrice, $r^* = e + r + r^2 + \dots$
$r^+$	si $r$ est une matrice, $r^+ = rr^* = r + r^2 + \dots$
$T^{ij}$	$T^{ij}(x^1, \dots, x^I) = (x^1, \dots, x^i - 1, x^{i+1}, \dots, x^j + 1, \dots, x^I)$
$v_x^n$	composante $x$ de la fonction de la programmation dynamique à l'étape $n$
$v_x$	composante $x$ de la fonction de la programmation dynamique dans le cas stationnaire ou ergodique
$w$	coût moyen par unité de temps
$c_x^n$	composante $x$ du coût instantané à l'étape $n$
$c_x$	composante $x$ du coût instantané dans le cas stationnaire ou ergodique
$F^n$	$\sigma$ -algèbre rendant mesurable les applications coordonnées jusqu'à l'étape $n$
$\mathbb{P}$	mesure de probabilité
$\mathbb{E}$	espérance mathématique
$1_A$	fonction indicatrice fonction de l'ensemble $A$
$Q_{m,\sigma}(x)$	forme quadratique $1/2[(x - m)/\sigma]^2$
$\mathcal{N}(m, \sigma)(x)$	densité d'une loi gaussienne de moyenne $m$ et d'écart type $\sigma$
$\mathbb{K}$	mesure de coût
$c$	souvent densité de coût
$\mathbb{M}$	valeur d'une variable de décision
$\mathbb{O}$	optimum d'une variable de décision
$\mathbb{F}$	fonction caractéristique d'une variable de décision
$\mathcal{F}$	transformée de Fenchel
$\mathcal{L}$	transformé de Laplace
$\mathcal{C}$	transformé de Cramer
$(A, B, C)$	matrices définissant un système dynamique linéaire temps invariant en temps continu ou en temps discret
$Q, S, R$	matrices définissant les coûts quadratiques pour les problèmes LQ ou LQG
$H$	fonction de transfert d'un système linéaire temps invariant
$H^b(s)$	transposition de $H$ et changement du signe de son argument

- $P$  matrice symétrique définissant une fonction de Bellman ou de Lyapounov quadratique
- $V, v, W$  souvent fonction de Bellman ou de Lyapounov
- $K$  gain optimal pour les problèmes LQ, LQG, ou Kalman



## Bibliographie

- [1] D. DACUNHA-CASTELLE, M. DUFLO (1982) *Probabilité et statistiques*. Masson.
- [2] F.DELEBECQUE, P. KOKOTOVIC, J.-P. QUADRAT (1984) *Aggregation and coherency in networks and Markov chains*. Int. J. of Control N. 35.
- [3] W. FELLER (1957) *An introduction to probability theory and its application*. J.Wiley.
- [4] M. FRECHET (1938) *Théorie des évènements en chaîne dans le cas d'un nombre fini d'états possibles*. Gauthier-Villars.
- [5] F. R. GANTMACHER (1966) *Théorie des matrices*. Dunod.
- [6] J. G. KEMENY, J. L. SNELL, A. W. KNAP (1976) *Finite Markov chains*. Springer-Verlag.
- [7] A. A. MARKOV (1910) Recherches sur un cas remarquable d'épreuves dépendantes. *Acta Mathematica* t.38 p.87-104.
- [8] J. NEVEU(1972) *Martingales à temps discret*. Masson.
- [9] R. PALLU DE LA BARRIERE (1966) *Cours d'automatique théorique*. Dunod.
- [10] D. REVUZ (1984) *Markov chains*. North-Holland.
- [11] M. AKIAN (1998)*Density of idempotent measures and large deviations*. Transaction of the American Mathematical Society and INRIA report N. 2534 and 2611 (1995).
- [12] M. AKIAN, J.P. QUADRAT, M.VIOT (1994) *Bellman Processes*. In LN in Cont. and Inf. Sc. N. 199, Springer-Verlag.
- [13] F. BACCELLI, G. COHEN, G.J. OLSDER, J.P. QUADRAT (1992) *Synchronization and Linearity*. J. Wiley & Sons.
- [14] G. BIRKHOFF (1967) *Lattice theory* Am. Math. Soc. Colloq. Pub. N. 25.
- [15] Z.-Q. CAO, K.H. KIM, F.W. ROUSH (1984) *Incline algebra and applications* J. Wiley & Sons.
- [16] G. COHEN, D.DUBOIS, J.P. QUADRAT, M.VIOT (1985) *A linear system-theoretic view of discrete event processes and its use for performance evaluation in manufacturing*. IEEE Transaction on Automatic Control 30.
- [17] R.A. CUNNINGHAME-GREEN (1979) *Minimax Algebras*. L.N. in Economics and Math. Systems, Springer Verlag.
- [18] M.L. DUBREIL-JACOTIN, L. LESIEUR, R. CROISOT (1953) *Théorie des treillis des structures algébriques ordonnées et des treillis géométriques* Gauthier-Villars.
- [19] S. GAUBERT and M. PLUS (1997) *Methods and applications of max-plus linear algebra*. In LN in Comp. Sc. N. 1200, Springer Verlag.
- [20] M. GONDRAN, M. MINOUX (1986) *Graphes et Algorithms*. J. Wiley & Sons.
- [21] J. GUNAWARDENA (editor) (1998) *Idempotency*. Cambridge University Press.
- [22] V.N. KOLOKOLTSOV, V.P. MASLOV (1997) *Idempotent analysis and its applications*. Kluwer.

- [23] V.P. MASLOV, S.N. SAMBORSKI (1992) *Idempotent Analysis*. Advances in Soviet Math. 13, Amer. Math. Society.
- [24] E. PAP (1995) *Null-additive set functions*. Math. and Appl. N.337, KLUWER.
- [25] J.S. Golan (1992) *The theory of semirings with applications in mathematics and theoretical computer sciences*. Pitman.
- [26] J.-P. QUADRAT (1989) *Théorèmes asymptotiques en programmation dynamique*. CRAS 311.
- [27] J.-P. QUADRAT and MAX-PLUS working group (1997) *Min-plus linearity and statistical mechanics*. Markov Processes and Related Fields 3.
- [28] U. ZIMMERMANN (1981) *Linear and combinatorial optimization in ordered algebraic structures*. Annals of discrete math. N. 10 North-Holland.
- [29] R. BELLMAN (1957) *Dynamic programming*. Princeton University Press.
- [30] A. BENSOUSSAN (1982) *Stochastics control by functional analysis method*. North-holland.
- [31] A. BENSOUSSAN, J. L. LIONS (1978) *Application des inéquations variationnelles en contrôle stochastique*. Dunod.
- [32] D. BERTSEKAS (1976) *Dynamic programming and stochastic control*. Academic Press.
- [33] J.P. BICKEL, N. EL KAROUI, M. YOR (1979) *Ecole d'été de probabilités de Saint Flour, Lecture Notes in Math. n. 876*. Springer-Verlag.
- [34] J. M. BISMUT (1981) *Mécanique aléatoire, Lecture Notes in Math. n. 866*. Springer-Verlag.
- [35] C. DERMAN (1970) *Finite state markovian decision processes*. Academic Press.
- [36] F. DELEBECQUE, J.-P. QUADRAT (1978) *Contribution of stochastic control singular perturbation averaging and team theory to an example of large system management of hydropower production*, IEEE Transaction on Automatic Control, April.
- [37] F. DELEBECQUE, J.-P. QUADRAT (1981) *Optimal control of Markov chains admitting weak and strong interactions*. Automatica april.
- [38] W. H. FLEMING, R. W. RISHEL (1975) *Deterministic and stochastic optimal control*. Springer-Verlag.
- [39] M. FLIESS (1973) *Automates et séries rationnelles non commutatives* dans "Automata languages programming", édité par Nivat. North Holland.
- [40] M. GOURSAT, J. P. QUADRAT (1987) *Optimal stochastic control: Numerical methods*. *Systems & Control Encyclopedia* édité par M. G. Singh. Pergamon Press.
- [41] O. HERNANDEZ-LERMA (1989) *Adaptive Markov control processes*. Springer Verlag.
- [42] R. A. HOWARD (1960) *Dynamic programming and Markov process*. J.Wiley.
- [43] P. R. KUMAR, P. VARAIYA (1986) *Stochastic systems: estimation, identification and adaptive control*. Prentice-hall.
- [44] H. KUSHNER (1971) *Introduction to stochastic control*. Holt, Rinehart and Winston.
- [45] S. M. ROSS (1983) *Introduction to dynamic programming*. Academic Press.
- [46] P. L. LIONS (1982) *Generalized solutions of Hamilton Jacobi equations*. Notes in Math. Pitman.
- [47] J.-P. QUADRAT, M. VIOT (1973) *Méthodes de simulation en programmation dynamique stochastique RAIRO, avril*.

- [48] J.-P. QUADRAT (1980) *Existence de solution et algorithme de résolution numérique de problème de contrôle optimal de diffusion stochastique dégénérée ou non*. SIAM Journal of Control, Mars.
- [49] D. J. WHITE (1969) *Dynamic programming*. Holden-Day.
- [50] P. WHITTLE (1982-83) *Optimization over time v.1 and v.2* J. Wiley and Sons.
  
- [51] K.J. ASTROM , B. WITTENMARK (1984) *Computer controlled systems: theory and design*, Pentice Hall.
- [52] B.D.O. ANDERSON, J.B. MOORE (1990) *Optimal control: linear quadratic methods*, Prentice Hall.
- [53] B.D.O. ANDERSON, J.B. MOORE (1979) *Optimal filtering*, Prentice Hall.
- [54] J.J. D'AZZO, C.H. HOUPIS (1966) *Feedback control system analysis and synthesis*, McGraw-Hill.
- [55] J.C. DOYLE, B.A. FRANCIS, A.R. TANNENBAUM (1992) *Feedback Control Theory*, MacMillan Publishing Company.
- [56] P. FAURRE, M. DEPEYROT (1974) *Eléments d'automatique*, Dunod.
- [57] P. FAURRE, M. CLERGET, F. GERMAIN (1979) *Opérateurs rationnels positifs*, Dunod.
- [58] F.R. GANTMACHER (1966) *Théorie des matrices*, tome 2, Dunod.
- [59] M.J. GRIMBLE, M.A. JOHNSON (1988) *Optimal Control and Stochastic Estimation v.1 and v.2*, J. Wiley and Sons.
- [60] T. KAILATH (1980) *Linear systems*, Prentice Hall.
- [61] M. MORARI, E. ZAFIROU (1989) *Robust process control*, Prentice Hall.
- [62] J.R. PARTINGTON (1988) *An introduction to Hankel operators*, Cambridge University Press.
- [63] M.G. SAFONOV (1980) *Stability and robustness of multivariable feedback systems*, MIT Press.
- [64] P. WHITTLE (1990) *Risk-Sensitive optimal control* J. Wiley and Sons.



## Index

- équation
  - affine, 31
- états
  - espace, 9
- action, 149
- algorithme
  - de Howard, 51
  - de relaxation, 83
- algèbre idempotente, 33
- algèbre max-plus, 29
- applications coordonnées, 9
  
- boucle fermée, 47
- boucle ouverte, 48
  
- centrée, 40
- chaîne
  - Markov, 9
- chaînes
  - Bellman, 43
- chaîne
  - agrégée, 65
  - agrégée contrôlée, 70
  - lente, 62
  - Markov
    - perturbée, 61
    - perturbée et commandée, 66
  - rapide, 62
  - réversible, 78
- citère
  - Routh, 116
- classe
  - finale, 23
  - transitoire, 23
- clos, 23
- clown, 149
- combinatoire, 23
- commande, 47
- commande optimale, 47
- convergence
  - faible, 42
  - sensibilité, 42
- coût
  - conditionnel, 39
  - densité, 39
  - mesure, 39
  - non actualisé, 48
- critère
  - cercle, 109
  - Nyquist, 113
  - positivité, 106
- cycle moyen
  - maximum, 36
  
- demi-corps, 29
  - commutatif, 29
  - idempotent, 29
- division, 30
- décision
  - espace, 39
  - variable, 39
  
- égalité des résolvantes, 16
- ensemble clos, 23
- entrée, 46
- equation affine
  - forme canonique, 33
- equation de Kolmogorov
  - actualisée, 13
  - arrière, 13
  - arrêté, 14
  - avant, 12
  - ergodique, 21
- équilibre local, 78
- ergodique, 22
- espace  $\mathbb{D}^p$ , 40
- etat, 46, 47
  
- factorisation, 74
- factorisation spectrale, 104
- feedback, 47
- feedback a priori, 48
- feedback locaux, 80
- file d'attente MM1, 75
- filtrage, 141
- filtrage optimal, 56
- filtre

- Kalman, 85
- Kalman asymptotique, 88
- filtre non normalisé, 56
- finale, 23
- fonction
  - affine, 31
  - caractéristique, 41
  - linéaire, 31
- forme de Brunowski, 152
- forme produit, 74, 79
- formule binomiale, 30
- gestion
  - de barrage, 10
  - de réservoir, 129
  - de stock, 117
- groupe, 29
- générateur, 13
- homogène, 9
- horizon
  - fini, 13, 48
  - infini, 13, 48
- identification, 59
- indice de Cauchy, 113
- indépendance, 40
- inf-convolution, 40
- information incomplète, 55
- irréductible, 23
- itération
  - sur les politiques, 51, 53
  - sur les politiques chaîne de Markov perturbée, 67
  - sur les valeurs, 50
- jeu de pile ou face, 133
- loi
  - de commande, 47
  - grands nombres, 42
  - initiale, 9
- maintenance, 45, 121, 142, 144
- martingale, 15
- matrice
  - de Hankel, 58
  - de routage, 74
  - de transition, 9
  - identité, 33
  - zéro, 33
- mesure
  - de probabilité invariante extrême, 25
  - invariante, 25
  - moduloïde, 32
  - nilpotent, 17
    - d'une matrice stochastique, 20
  - normalisée, 40
  - observation, 47
  - open loop, 48
  - optimum, 40
  - pile ou face, 10
  - placement de pôles, 100
  - point de Nash, 81
  - politique markovienne, 47
  - Pontriaguine, 153
  - population, 11
  - potentiel, 27
  - principe
    - du maximum, 24
    - du minimum, 24
  - probabilité markovienne, 9
  - problème
    - de Dirichlet, 14, 27
    - de jeux, 81
  - processus de calcul, 137
  - programmation dynamique
    - chaîne de markov perturbée, 67
    - coût actualisé, 50
    - coût ergodique, 53
    - en horizon fini, 48
    - information incomplète, 57
  - projecteur spectral, 17, 21
  - pseudoinverse, 17
  - relaxation, 83
  - réalisation, 58
  - régulateur
    - LQ, 99
    - LQ (approche fréquentielle), 103
    - LQG, 92
    - LQG information complète, 94
    - LQG information incomplète, 95
    - robustesse LQ, 110
  - résolvante, 16
  - sensibilité, 40
  - solution, 35
  - sortie, 46
  - sous-solution, 35
    - plus grande, 35
  - spectral, 36
  - stable
    - asymptotiquement au sens de Lyapounov, 110

- coûts, 41
- entrée bornée sortie bornée, 111
- exponentiellement, 111
- Lyapounov, 110
- stock d'eau, 45
- stratégie, 47
  - déterministe, 47
  - Nash optimal, 82
  - relaxée, 47
- stratégies, 146
- suite
  - de Sturm, 114
- système
  - d'équations linéaires, 33
  - fermé de files d'attente, 73
  - ouvert de files d'attente, 80
  - positif, 106
- série génératrice, 58
- tableau
  - de Routh, 115
- temps d'arrêt, 14
- théorie spectrale, 36
- théorème
  - grands nombres, 42
  - limite centrale, 42
  - Lyapounov, 111
  - Popov, 108
  - Routh, 115
  - Sturm, 114
- transformée
  - Cramer, 42
  - Fenchel, 41
- transitoire, 23
- transport, 126
- valeur propre
  - maximale, 38
- zéro
  - absorbant, 29