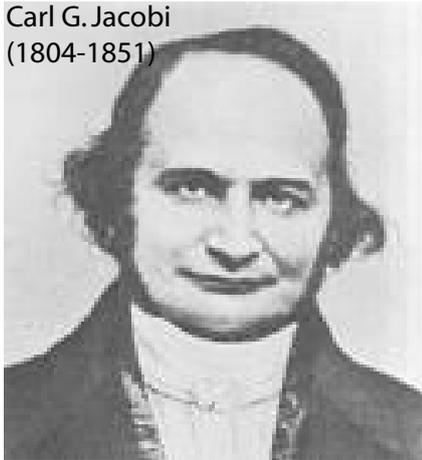


Notes de cours de commande optimale stochastique

Jean-Pierre Quadrat
rédigées par
Pierre Girardeau

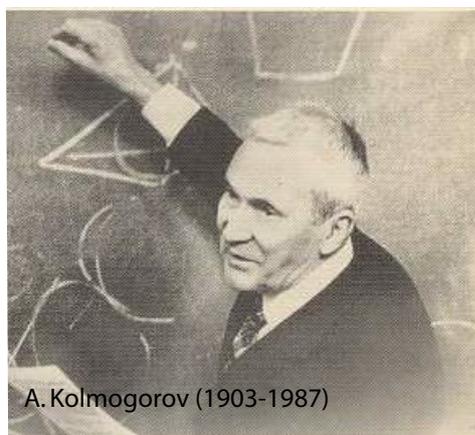
Carl G. Jacobi
(1804-1851)



Andrei A. Markov
(1856-1922)



Richard Bellman
(1920-1984)



A. Kolmogorov (1903-1987)

Table des matières

partie 1. Commande optimale des chaînes de Markov	5
Chapitre 1. Chaînes de Markov - Equations de Kolmogorov	9
1. Chaînes de Markov avec espace d'état fini	9
2. L'équation de Kolmogorov en avant	11
3. L'équation de Kolmogorov en arrière	12
4. L'équation de Kolmogorov actualisée	13
Chapitre 2. Commande optimale des chaînes de Markov en observation complète	17
1. Le problème de commande optimale	17
2. Programmation dynamique en horizon fini	17
3. Programmation dynamique avec coût actualisé	19
Chapitre 3. Propriétés combinatoires spectrales et asymptotiques des chaînes de Markov	23
1. Propriétés spectrales des matrices stochastiques	23
2. Propriétés combinatoires	24
3. Classification des matrices stochastiques	25
4. Calcul du projecteur spectral P	26
Chapitre 4. Commande optimale des chaînes de Markov en observation incomplète	29
1. Chaînes de Markov contrôlées et observées	29
2. Filtre optimal d'une chaîne de Markov observée	29
3. Équation de la programmation dynamique dans le cas d'une observation incomplète	31
Chapitre 5. Problème 1 : Gestion de réservoir	35
1. Énoncé	35
2. Corrigé	36
Chapitre 6. Problème 2 : Stratégies	39
1. Énoncé	39
2. Corrigé	40
partie 2. Algèbre maxplus et chaînes de Bellman	43
Chapitre 7. Semi-anneaux idempotents	45
1. Structures	45
2. Matrices sur des semi-anneaux	45
3. Exemples	46
Chapitre 8. Résolution des systèmes linéaires dans l'algèbre maxplus	47

1. Linéarité maxplus	47
2. Interprétation du produit matriciel	47
3. Equation affine à une inconnue dans \mathbb{R}_{\max}	48
4. Méthode du point fixe	50
5. Résiduation	51
Chapitre 9. Réseaux de Petri et Théorie Spectrale	53
1. Modèle	53
2. Quelques réseaux de Pétri élémentaires	54
3. Les graphes d'événements sont des systèmes maxplus linéaires	55
4. Théorie spectrale	56
Chapitre 10. Symétrisation de l'algèbre maxplus	59
Chapitre 11. Décision Optimale	63
1. Espace de décision, variables de décision	63
2. Transformée de Fenchel	65
3. Coûts stables	65
4. Les théorèmes limites pour les variables de décision	66
5. Transformée de Cramer	66
Chapitre 12. Chaînes de Bellman	69
Bibliographie	71
Annexe A. Cercles de Gershgorin	73
Annexe B. Test	75
1. Projecteur spectral	75
2. Programmation dynamique	75
3. Algorithme de Howard	75
4. Chemins les plus courts d'un graphe	75
5. Mise en équation d'un graphe d'événements	76
6. Calcul de valeur propre de matrice minplus	76

Première partie

Commande optimale des chaînes de
Markov

Cette partie présente une introduction au problème de commande optimale stochastique dans le cas où le temps, l'état et la commande sont discrets. De plus nous supposons que les nombres d'états et de commandes sont finis. Ces restrictions ne réduisent pas le domaine d'application des méthodes proposées puisque dans les cas plus généraux, après une discrétisation, on se ramène à la situation présentée ici.

Le but de la "commande optimale stochastique" est d'optimiser, au sens d'un critère donné, l'entrée d'un système dynamique perturbé par des aléas dont on connaît les statistiques. Le système est décrit par l'évolution de son état X_t en fonction du temps t . L'évolution de l'état dépend de la commande U_t et des aléas ω dont on connaît la loi de probabilité. Plusieurs cadres mathématiques sont possibles. Ils se ramènent à des choix divers de l'espace des états, des commandes, du temps, de la description de la dynamique et du critère à optimiser

Pour l'espace des états \mathcal{E} on peut distinguer les cas où : – l'espace est discret fini $\mathcal{E} = \{1, 2, \dots, E\}$, c'est le cadre des chaînes de Markov à espace d'état fini dans lequel nous nous placerons ; – l'espace des états est infini mais dénombrable $\mathcal{E} = \mathbb{N}$, c'est le cas des files d'attente ; – l'espace des états est un espace vectoriel de dimension finie $\mathcal{E} = \mathbb{R}^n$. C'est le cas des problèmes de commande Linéaire Quadratique Gaussien (LQG)¹, et des Équations Différentielles Stochastiques (EDS) souvent utilisées en finance, – L'espace des états est un espace de Sobolev, cadre naturel pour les systèmes régis par des Équations aux Dérivées Partielles Stochastiques (EDPS).

Le temps \mathcal{T} peut être entier \mathbb{N} ou réel \mathbb{R} .

L'espace de commandes \mathcal{U} peut être par exemple : – fini $\mathcal{U} = \{1, 2, \dots, F\}$ dans cas des chaînes de Markov que l'on étudiera ; – un espace vectoriel $\mathcal{U} = \mathbb{R}^n$ dans le cas LQG ou des EDS ; – un sous ensemble le plus souvent convexe compact d'un espace vectoriel, dans le cas de commandes contraintes.

Les aléas Ω sont probabilisés et la loi d'évolution de l'état est connue. On peut distinguer : – le cas des chaînes de Markov où une matrice, appelée matrice de transition $M_{xx'}$, donne la probabilité de la transition de x à x' ; – le cas LQG ou la dynamique est linéaire et les bruits V_t et W_t sont Gaussiens ou on n'observe qu'une combinaison linéaire des états Y_t :

$$\begin{cases} X_{t+1} &= AX_t + BU_t + V_t, \\ Y_t &= CX_t + W_t; \end{cases}$$

– le cas des EDS où la dynamique est donnée par :

$$dX_t = b(X_t, U_t) dt + \sigma(X_t) dW_t,$$

avec W_t mouvement Brownien.

Le but de la commande optimale est de trouver la commande minimisant un certain critère qui est fonctionnelle additive de la trajectoire par exemple : – dans le cas du temps discret avec un horizon fini :

$$\min_U \mathbb{E} \left\{ \sum_{t=0}^T c(X_t, U_t) + \Phi(X_T) \right\};$$

¹Dans un problème de commande Linéaire Quadratique Gaussien, on cherche à commander un état soumis à une dynamique linéaire, perturbée par un bruit gaussien, de façon à minimiser un critère quadratique.

– dans le cas du temps discret avec un horizon infini :

$$\min_U \mathbb{E} \left\{ \sum_{t=0}^{+\infty} \left(\frac{1}{1+\lambda} \right)^{n+1} c(X_t, U_t) \right\};$$

– dans le cas du temps continu avec un horizon fini :

$$\min_U \mathbb{E} \left\{ \int_{t=0}^T c(X_t, U_t) + \Phi(X_T) \right\};$$

– dans le cas du temps continu avec un horizon infini :

$$\min_U \mathbb{E} \left\{ \int_{t=0}^T e^{-\lambda t} c(X_t, U_t) \right\}.$$

CHAPITRE 1

Chaînes de Markov - Equations de Kolmogorov

Dans ce chapitre nous présentons les résultats relatifs aux chaînes de Markov ayant un nombre fini d'états. Après avoir introduit les définitions nécessaires, nous présentons les équation de Kolmogorov en avant, en arrière, et actualisées. On pourra consulter le livre de Feller [Fel57] pour plus d'information sur les processus de Markov.

1. Chaînes de Markov avec espace d'état fini

Pour toute fonction f définie sur $\mathcal{E} = \{1, \dots, E\}$ fini, on note f_x l'image par f de $x \in \mathcal{E}$.

Avant de définir ce qu'est une chaîne de Markov, rappelons deux définitions fondamentales.

Définition 1.1 (Probabilité, Matrice stochastique). *Une probabilité sur \mathcal{E} est une fonction p de \mathcal{E} dans \mathbb{R} telle que :*

$$p_x \geq 0, \forall x \in \mathcal{E}, \quad \sum_{x \in \mathcal{E}} p_x = 1.$$

Une matrice stochastique sur \mathcal{E} est une matrice M de dimension $E \times E$ telle que :

$$M_{xx'} \geq 0, \forall x, x' \in \mathcal{E}, \quad \sum_{x' \in \mathcal{E}} M_{xx'} = 1, \forall x \in \mathcal{E}.$$

$M_{xx'}$ représente la probabilité d'aller en x' sachant que l'on est en x . Dit avec des notations probabilistes, on a que :

$$M_{xx'} = \mathbb{P} \{ X^{n+1} = x' \mid X^n = x \}.$$

De plus, on peut remarquer que chaque ligne de M définit une loi de probabilité sur \mathcal{E} .

Nous avons maintenant tous les outils nécessaires à la construction des chaînes de Markov ayant un nombre fini d'états.

Définition 1.2 (Chaîne de Markov). *Une Chaîne de Markov est la donnée du quadruplet $(\mathcal{T}, \mathcal{E}, p^0, M)$, où :*

- $\mathcal{T} = \mathbb{N}$ est l'espace des temps $n \in \mathcal{T}$;
- $\mathcal{E} = \{1, 2, \dots, E\}$ est l'espace des états $x \in \mathcal{E}$;
- p^0 est une loi de probabilité appelée loi initiale de l'état ;
- M est une matrice stochastique appelée matrice de transition sur \mathcal{E} .

Introduisons maintenant l'espace de probabilité canonique de la chaîne de Markov.

Définition 1.3 (Probabilité d'une chaîne de Markov). *Une trajectoire est un élément ω de l'ensemble $\Omega = \mathcal{E}^{\mathcal{T}}$. De plus, on note $X^n(\omega) = \omega^n (\in \mathcal{E})$ l'état à*

l'instant n , pour tout n de \mathcal{T} . Une trajectoire ω est donc une suite définie sur \mathcal{T} à valeurs dans \mathcal{E} .

Sur Ω , on définit alors une loi de probabilité \mathbb{P} par sa valeur sur les cylindres de trajectoires de base finie :

$$\mathbb{P} \{X^0 = x^0, X^1 = x^1, \dots, X^n = x^n\} = p_{x^0}^0 M_{x^0 x^1} \dots M_{x^{n-1} x^n}.$$

Grâce à un théorème de Kolmogorov, on peut montrer que ce système compatible de lois marginales définit bien une loi de probabilité unique sur Ω qui est de dimension infinie.

Les chaînes de Markov que nous venons de définir sont donc des processus stochastiques en temps discret et à espace d'état discret et fini qui possèdent la propriété suivante, importante pour les calculs.

Proposition 1.4 (Propriété de Markov).

$$\mathbb{P} \{X^n = x^n \mid X^{n-1} = x^{n-1}, \dots, X^0 = x^0\} = M_{x^{n-1} x^n}$$

DÉMONSTRATION. Le résultat découle directement de la définition de \mathbb{P} et de la formule de Bayes :

$$\begin{aligned} \mathbb{P} \{X^n = x^n \mid X^{n-1} = x^{n-1}, \dots, X^0 = x^0\} &= \frac{\mathbb{P} \{X^n = x^n, \dots, X^0 = x^0\}}{\mathbb{P} \{X^{n-1} = x^{n-1}, \dots, X^0 = x^0\}}, \\ &= \frac{p_{x^0}^0 M_{x^0 x^1} \dots M_{x^{n-1} x^n}}{p_{x^0}^0 M_{x^0 x^1} \dots M_{x^{n-2} x^{n-1}}}, \\ &= M_{x^{n-1} x^n}. \end{aligned}$$

□

La propriété de Markov traduit le fait que l'état présent résume tout le passé. En effet, savoir que $X^{n-1} = x^{n-1}, \dots, X^0 = x^0$ équivaut à savoir que $X^{n-1} = x^{n-1}$.

Pour illustrer ces définitions, donnons un exemple simple.

Exemple 1.5. Reprenons un exemple proposé par A. Kaufman. Un vendeur de marionnettes mexicain vend un seul modèle de marionnette. A la fin de chaque mois, il fait ses comptes et constate le succès ou l'échec de son modèle du moment. Il a alors la possibilité d'en changer. Nous supposons ici qu'il applique la stratégie suivante :

– En cas de succès du modèle de marionnette, le vendeur choisit de conserver le même modèle le mois suivant. Il a alors une probabilité 0.5 de se retrouver en situation de succès, et une probabilité 0.5 de se retrouver en situation d'échec.

– En cas d'échec, il change de modèle. Il a alors une probabilité 0.7 de se retrouver en situation de succès, et une probabilité 0.3 de rester en situation d'échec.

Nous considérerons que le bénéfice du vendeur est de 500 pesos dans le cas d'un succès, et de 100 pesos dans le cas d'un échec.

Si on considère que l'espace d'état est $\{0, 1\}$, avec 0 signifiant l'échec et 1 la réussite, il découle directement de la définition que l'évolution de l'état est markovienne. On peut la décrire par le graphique de la figure 1.

La matrice de transition s'écrit ainsi :

	succès	échec
succès	0.5	0.5
échec	0.7	0.3

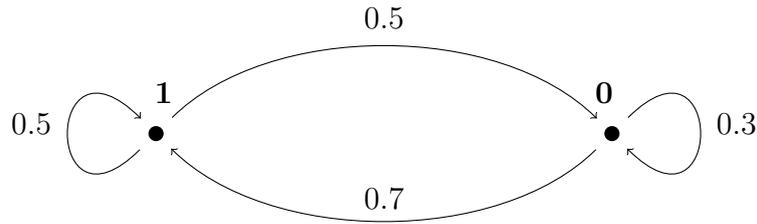


FIG. 1. Représentation graphique des transitions de la chaîne de Markov

On peut alors se poser la question suivante : en supposant que ce vendeur travaille pendant N mois, quelle est son espérance mathématique de gain ? Dans le but de répondre à ce genre de questions nous allons d'abord nous concentrer sur l'évaluation de probabilités de présence dans un état à un instant donné.

2. L'équation de Kolmogorov en avant

L'équation de Fokker-Planck, également connue sous le nom d'équation de Kolmogorov en avant, décrit l'évolution de la probabilité de présence, à l'instant n , de la chaîne de Markov, en un point $x \in \mathcal{E}$:

$$\mathbb{P}\{X^n(\omega) = x\} = p_x^n$$

On note : $p^n = [p_1^n \ \cdots \ p_E^n]$.

Proposition 1.6 (Equation de Fokker-Planck). *Supposons la loi initiale de l'état p^0 donnée. On a alors, pour tout $n \geq 0$:*

$$p^{n+1} = p^n M.$$

DÉMONSTRATION. Par définition de \mathbb{P} on a :

$$\mathbb{P}\{X^n(\omega) = x\} = \sum_{x^0, \dots, x^{n-1}} p_{x^0}^0 M_{x^0 x^1} \dots M_{x^{n-1} x} = \left[p^0 \underbrace{M \dots M}_{n \text{ fois}} \right]_x,$$

d'où le résultat. \square

Exemple 1.7. Calculons la probabilité d'avoir un succès ou un échec au bout de deux étapes pour le vendeur de marionnettes, sachant qu'à l'instant 0 il a eu un succès :

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$$

La probabilité d'avoir un succès au bout de deux étapes est de 0.6.

Si on itère ce procédé, afin de connaître la probabilité d'avoir un succès ou un échec au bout de trois, quatre mois, on trouve les distributions de probabilité suivante : $[0.58 \ 0.42]$ et $[0.584 \ 0.416]$ respectivement. Ces résultats nous incitent à penser qu'il existe ici une loi de probabilité p telle que :

$$p = pM,$$

que l'on appelle alors mesure invariante. Nous reviendrons longuement sur les mesures invariantes dans le chapitre 2, section 3.

3. L'équation de Kolmogorov en arrière

Comme annoncé au début de ce chapitre, nous voulons calculer des fonctionnelles additives des trajectoire de la chaîne de Markov. Dans le cadre d'un horizon fini, disons sur N étapes, nous voulons calculer :

$$V = \mathbb{E} \left\{ \sum_{k=0}^{N-1} C_{X^k} + \Phi_{X^N} \right\},$$

où C et Φ sont des vecteurs de \mathbb{R}^E , avec :

- C le coût instantané,
- Φ le coût en fin de période.

L'équation de Kolmogorov (en arrière) va nous permettre de calculer ce coût en procédant, d'une certaine manière, en marche arrière. Pour cela, nous introduisons la famille à deux paramètres n et x de fonctionnelles additives de la trajectoire :

$$V_x^n = \mathbb{E} \left\{ \sum_{k=n}^{N-1} C_{X^k} + \Phi_{X^N} \mid X^n = x \right\},$$

qui sont les espérances du coût futur sachant que nous sommes dans l'état x à l'instant n . Nous avons alors le résultat suivant.

Proposition 1.8 (Equation de Kolmogorov en arrière). *La fonctionnelle $V^n = [V_1^n \ \cdots \ V_E^n]$ satisfait à l'équation de récurrence arrière suivante :*

$$V^n = C + MV^{n+1}, \quad V^N = \Phi,$$

et donc $V = p_0 V_0$.

DÉMONSTRATION. Nous proposons deux démonstrations de ce résultat, l'une reposant sur des arguments probabilistes, et l'autre purement algébrique.

- (1) **Démonstration probabiliste :** Soit $n < N$. Par la propriété de l'espérance conditionnelle puis la propriété de Markov, on a :

$$\begin{aligned} V_x^n &= C_x + \mathbb{E} \left\{ \sum_{k=n+1}^{N-1} C_{X^k} + \Phi_{X^N} \mid X^n = x \right\}, \\ &= C_x + \mathbb{E} \left\{ \mathbb{E} \left\{ \sum_{k=n+1}^{N-1} C_{X^k} + \Phi_{X^N} \mid X^{n+1}, X^n = x \right\} \mid X^n = x \right\}, \\ &= C_x + \mathbb{E} \left\{ \mathbb{E} \left\{ \sum_{k=n+1}^{N-1} C_{X^k} + \Phi_{X^N} \mid X^{n+1} \right\} \mid X^n = x \right\}, \\ &= C_x + \mathbb{E} \{ V_{X^{n+1}}^{n+1} \mid X^n = x \}, \\ &= C_x + \sum_{y \in \mathcal{E}} M_{xy} V_y^{n+1}, \\ &= C_x + (MV^{n+1})_x. \end{aligned}$$

Enfin, pour $n = N$:

$$V_x^N = \mathbb{E} \{ \Phi_{X^N} \mid X^N = x \} = \Phi_x.$$

(2) **Démonstration algébrique :**

$$\begin{aligned}
V &= p_0 C + p_0 M C + p_0 M^2 C + \dots + p_0 M^N \Phi, \\
&= p_0 \left(C + M \left(C + \dots \left(C + M \underbrace{\left(C + M \underbrace{\Phi}_{V^N} \right)}_{V^{N-1}} \right) \right) \right)_{V^{N-2}} \Bigg).
\end{aligned}$$

□

Exemple 1.9. Revenons à notre vendeur de marionnettes. Supposons qu'au bout de deux mois il revende son commerce. Or, il sait que s'il s'il finit sur un succès, il vendra son commerce 2000 pesos. Au contraire, s'il termine sur un échec, il ne vendra que 1000 pesos. De plus, on a toujours qu'un succès lors d'un mois correspond à un bénéfice de 500 pesos, et un échec à 100 pesos.

Nous reprenons les mêmes probabilités de transition que décrites dans la figure 1.

L'équation de Kolmogorov s'écrit donc :

$$V^n = \begin{bmatrix} 500 \\ 100 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} V^{n+1}, n \in \{0, 1\}, \quad V^2 = \begin{bmatrix} 2000 \\ 1000 \end{bmatrix}.$$

Notre vendeur souhaiterait connaître son espérance de gain dès le début du premier mois, sachant qu'il part d'une situation de succès. On calcule donc V^1 :

$$V^1 = \begin{bmatrix} 500 \\ 100 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 2000 \\ 1000 \end{bmatrix} = \begin{bmatrix} 2000 \\ 1800 \end{bmatrix},$$

puis V^0 :

$$V^0 = \begin{bmatrix} 500 \\ 100 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 2000 \\ 1800 \end{bmatrix} = \begin{bmatrix} 2400 \\ 2040 \end{bmatrix}.$$

Son espérance de gain partant d'un succès est donc de 2400 pesos.

4. L'équation de Kolmogorov actualisée

Intéressons-nous maintenant au cas du coût actualisé (en horizon infini). La fonctionnelle que nous voulons alors calculer est de la forme :

$$(1) \quad V_x = \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X^n} \mid X^0 = x \right\},$$

avec $\lambda \in \mathbb{R}$ et $\lambda > 0$, afin d'assurer l'existence de la somme infinie. De plus, économiquement, $\lambda > 0$ traduit le fait qu'un gain est d'autant plus profitable qu'il est fait tôt. Il est donc naturel d'introduire cette "dévaluation".

Remarquons que :

$$(2) \quad V = \frac{1}{1+\lambda} \left(C + M \underbrace{\left(\frac{1}{1+\lambda} \right) \left(C + M \underbrace{\left(\frac{1}{1+\lambda} \right) (C + \dots)}_V \right)}_V \right).$$

Dès lors, le cas de l'horizon infini semble plus simple que le cas précédent, car nous n'avons pas à introduire de fonctions intermédiaires en temps. On observe directement que V est solution d'une certaine équation de point fixe. Citons donc ce résultat.

Proposition 1.10 (Equation de Kolmogorov avec coût actualisé). V défini par (1) est solution de :

$$A_\lambda V + C = 0,$$

avec $A_\lambda = M - (1 + \lambda)I$ le générateur de la chaîne de Markov.

On tire directement de l'équation (2) que $(1 + \lambda)V = C + MV$, ce qui nous donne le résultat. Cependant, afin de bien comprendre les mécanismes sous-jacents mis en jeu, nous proposons une démonstration reposant sur des arguments probabilistes. Ces derniers sont pratiquement identiques à ceux utilisés dans la démonstration de l'équation de Kolmogorov en horizon fini.

DÉMONSTRATION. Soit $x \in \mathcal{E}$. On utilise la définition de V , puis la propriété de l'espérance conditionnelle et la propriété de Markov pour montrer que :

$$\begin{aligned} V_x &= \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X^n} \mid X^0 = x \right\}, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X^{n+1}} \mid X^0 = x \right\}, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} \mathbb{E} \left\{ \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X^{n+1}} \mid X^1, X^0 = x \right\} \mid X^0 = x \right\}, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} \mathbb{E} \left\{ \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X^{n+1}} \mid X^1 \right\} \mid X^0 = x \right\}, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} \mathbb{E} \{ V_{X^1} \mid X^0 = x \}, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} \sum_{y \in \mathcal{E}} M_{xy} V_y, \\ &= \frac{1}{1+\lambda} C_x + \frac{1}{1+\lambda} (MV)_x, \end{aligned}$$

d'où : $(1 + \lambda)V = MV + C$. □

Exemple 1.11. Revenons à notre vendeur de marionnettes, qui a finalement décidé de ne pas vendre son commerce, mais de le léguer à son fils. Ce dernier, en début de carrière de vendeur, souhaiterait savoir quelle est son espérance de gain

à long terme. Il estime la dévaluation mensuelle à 10%. Son problème est donc d'estimer :

$$\mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+0.1)^{n+1}} C_{X^n} \mid X^0 = x \right\}.$$

Comme auparavant, la matrice de transition est :

$$M = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix}, \quad A_\lambda = M - (1 + \lambda)I = \begin{bmatrix} -0.6 & 0.5 \\ 0.7 & -0.8 \end{bmatrix}.$$

L'espérance de gain mensuel, représentée par V , est alors solution de :

$$A_\lambda \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} 500 \\ 100 \end{bmatrix} = 0.$$

A_λ est à diagonale dominante et donc inversible (voir annexe A), on obtient :

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = -\frac{1}{0.13} \begin{bmatrix} -0.8 & -0.5 \\ -0.7 & -0.6 \end{bmatrix} \begin{bmatrix} 500 \\ 100 \end{bmatrix} \simeq \begin{bmatrix} 3461 \\ 3153 \end{bmatrix}$$

L'espérance de gain du vendeur de marionnette sachant qu'il démarre sur un succès est de 3461 pesos, et de 3153 pesos s'il démarre sur un échec.

CHAPITRE 2

Commande optimale des chaînes de Markov en observation complète

Nous définissons les chaînes de Markov commandées. Nous optimisons la commande en résolvant des équations de la programmation dynamique dans le cas où on observe complètement l'état de la chaîne de Markov. La programmation dynamique a été introduite par R. Bellman [Bel57].

1. Le problème de commande optimale

Définition 2.1 (Chaîne de Markov commandée). *Une chaîne de Markov commandée est la donnée de $(\mathcal{T}, \mathcal{E}, \mathcal{F}, M^u, p^0)$, où : - $\mathcal{T} = \mathbb{N}$ est l'espace de temps, - $\mathcal{E} = \{1, \dots, E\}$ est l'espace des états discret et fini, - $\mathcal{F} = \{1, \dots, F\}$ est l'espace des commandes discret et fini, - $M_{xx'}^u$ est la probabilité de la transition de x à x' lorsqu'on la décision u est prise, - p^0 est la loi initiale de l'état.*

La différence importante avec le chapitre précédent est que la matrice de transition dépend maintenant de l'action de l'utilisateur. Il faut noter que, à u fixé, M^u est une matrice stochastique.

On appelle \mathcal{S} l'ensemble des *stratégies en boucle fermée sur l'état* (feedback en anglais), c'est-à-dire l'ensemble des suites $s = (s_n)_{n \in \mathcal{T}}$ de fonctions $s_n : \mathcal{E} \rightarrow \mathcal{F}$.

Définition 2.2 (Commande optimale en observation complète). *Etant donnés une chaîne de Markov commandée, le cout instantané $C : \mathcal{E} \times \mathcal{E} \mapsto \mathbb{R}^+$ et le coût sur l'état final $\Phi : \mathcal{E} \mapsto \mathbb{R}^+$, le problème de commande optimale stochastique en observation complète consiste à calculer la stratégie solution de :*

$$(3) \quad \min_{s \in \mathcal{S}} \mathbb{E} \left\{ \sum_{n=0}^{N-1} C_{X^n}^{U^n} + \Phi_{X^N} \right\},$$

où, pour simplifier l'écriture, on a noté $U^n = s^n(x^n)$.

De manière générale, on utilisera la notation U pour des commandes (éléments de \mathcal{F}) et s pour des stratégies (éléments de $\mathcal{E} \rightarrow \mathcal{F}$).

2. Programmation dynamique en horizon fini

L'équation de la programmation dynamique, également appelée équation de Bellman ou d'Hamilton-Jacobi-Bellman dans le contexte de la commande d'équations différentielles, nous permet de calculer ce minimum, ainsi que la stratégie s^* optimale, en procédant en marche arrière. Elle joue pour les problèmes de commande optimale le rôle de l'équation de Kolomogorov pour les chaînes de Markov.

Théorème 2.3 (Équation de la programmation dynamique). *Soit*

$$(4) \quad V_x^n = \min_{s \in \mathcal{S}} \mathbb{E} \left\{ \sum_{k=n}^{N-1} C_{X^k}^{U^k} + \Phi_{X^N} \mid X^n = x \right\}.$$

Alors V^n est solution de l'équation de la programmation dynamique :

$$V_x^n = \min_{u \in \mathcal{F}} (M^u V^{n+1} + C^u)_x, \quad V_x^N = \Phi_x.$$

et une stratégie optimale est donnée par :

$$s^n : x \mapsto u^* \in \operatorname{argmin}_{u \in \mathcal{F}} (M^u V^{n+1} + C^u)_x.$$

DÉMONSTRATION. Soit V^n une solution de (4) et σ une stratégie quelconque. Supposons que X^n évolue selon la loi induite par σ , on a :

$$\mathbb{E}^\sigma \{V_{X^{n+1}}^{n+1} - V_{X^n}^n \mid X^n = x\} = (M^{\sigma^n} V^{n+1} - V^n)_x.$$

Par l'optimalité de V^n , on a que $V_x^n \leq (M^u V^{n+1} + C^u)_x, \forall u \in \mathcal{F}$, et donc :

$$(M^{\sigma^n} V^{n+1} - V^n)_x \geq -C_{x^n}^{\sigma^n}.$$

Si on note $\mathbb{E}^{n,\sigma}$ l'espérance conditionnelle connaissant les états réalisés jusqu'à l'instant n , pour la loi induite par la stratégie σ , on obtient :

$$\mathbb{E}^{n,\sigma} \{V_{X^{n+1}}^{n+1} - V_{X^n}^n\} \geq -C_{X^n}^{\sigma^n}.$$

Alors :

$$\begin{aligned} \mathbb{E}^{0,\sigma} \{V_{X^N}^N - V_{X^0}^0\} &= \mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{N-1} V_{X^{n+1}}^{n+1} - V_{X^n}^n \right\}, \\ &= \mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{N-1} \mathbb{E}^{n,\sigma} \{V_{X^{n+1}}^{n+1} - V_{X^n}^n\} \right\}, \\ &\geq -\mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{N-1} C_{X^n}^{\sigma^n} \right\}. \end{aligned}$$

D'où :

$$V_{X^0}^0 \leq \mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{N-1} C_{X^n}^{\sigma^n} + \underbrace{V_{X^N}^N}_{\Phi_{X^N}} \right\}.$$

Et puisque l'on peut obtenir l'égalité en choisissant pour σ une stratégie optimale on obtient le résultat souhaité. \square

La programmation dynamique nous donne un procédé simple permettant de calculer des stratégies optimales, comme nous allons le voir sur un petit exemple.

Exemple 2.4. Notre nouveau vendeur de marionnettes a défini deux types de politiques commerciales : une attitude habituelle et une attitude agressive (consistant par exemple à faire des ristournes et de la publicité), cette dernière lui coûte plus cher mais augmente ses chances de succès. On a donc maintenant deux matrices de transition selon l'attitude adoptée, représentées par le graphe de la figure 1. On a donc les matrices de transitions suivantes :

$$M^0 = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \text{ et } M^1 = \begin{bmatrix} 0.6 & 0.4 \\ 0.8 & 0.2 \end{bmatrix},$$

avec la convention : 0 signifie l'attitude normale et 1 l'attitude agressive.

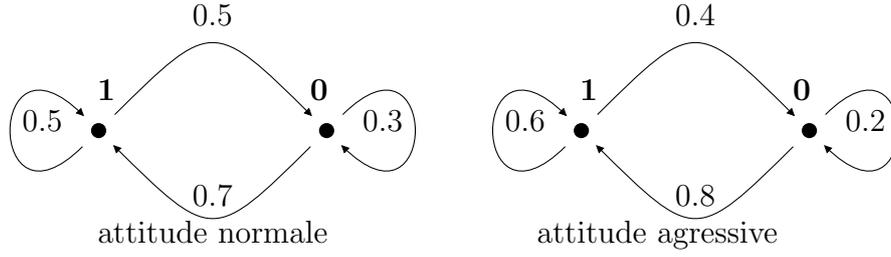


FIG. 1. Représentation graphique des transitions de la chaîne de Markov

Enfin, les coûts sont tels que :

$$C_x^u = \begin{cases} 20 & \text{pour } x = 0 \text{ et } u = 0 \\ 64 & \text{pour } x = 0 \text{ et } u = 1 \\ 325 & \text{pour } x = 1 \text{ et } u = 0 \\ 260 & \text{pour } x = 1 \text{ et } u = 1 \end{cases}$$

On peut maintenant facilement écrire l'équation de programmation dynamique, pour tout pas de temps n :

$$V_0^n = \max \left\{ \underbrace{0.3V_0^{n+1} + 0.7V_1^{n+1} + 20}_{u=0}, \underbrace{0.2V_0^{n+1} + 0.8V_1^{n+1} + 64}_{u=1} \right\},$$

$$V_1^n = \max \left\{ \underbrace{0.5V_0^{n+1} + 0.5V_1^{n+1} + 325}_{u=0}, \underbrace{0.4V_0^{n+1} + 0.6V_1^{n+1} + 260}_{u=1} \right\}.$$

3. Programmation dynamique avec coût actualisé

On s'intéresse maintenant au problème de commande optimale en horizon infini. Le facteur d'actualisation, λ nombre réel strictement positif, permet de gérer l'importance relative des événements à court et long terme et conduit à la résolution d'une équation de programmation dynamique stationnaire. Des méthodes itératives efficaces permettent de calculer la solution de l'équation de la programmation dynamique. Parmi ces méthodes, nous présentons l'*algorithme de Howard*, ou d'*itération sur les politiques*. La politique ainsi obtenue est stationnaire (indépendante du temps) et donc la résolution de ce problème actualisé conduit à l'obtention d'une stratégie plus simple que celle obtenue dans le cas de l'horizon fini.

Nous cherchons le coût optimal :

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

parmi les stratégies stationnaires $\mathcal{S} = \{s : \mathcal{E} \mapsto \mathcal{F}\}$.

Théorème 2.5 (Équation de la programmation dynamique actualisée). *Le coût optimal V est solution de l'équation :*

$$(5) \quad (1 + \lambda) V_x = \min_{u \in \mathcal{F}} (M^u V + C^u)_x.$$

L'argument du minimum donne la stratégie optimale s^ .*

EXISTENCE D'UNE SOLUTION À (5). L'existence d'une solution est donnée par l'algorithme de Howard [How60]. Étant donnée une stratégie $s^n : x \rightarrow u$, on lui associe un coût V^n en résolvant :

$$(1 + \lambda) V^n = M^{s^n} V^n + C^{s^n}.$$

Ensuite, étant donné le coût V^n , on lui associe la stratégie s^{n+1} :

$$s_x^{n+1} \in \operatorname{argmin}_{u \in \mathcal{F}} (M^u V^n + C^u)_x.$$

Nous allons montrer que la suite V ainsi obtenue est positive, décroissante, et qu'elle converge vers la solution de (5).

- La suite V^n est positive. C'est évident si l'on se souvient de l'interprétation stochastique de V^n comme solution d'une équation de Kolmogorov actualisé. En effet :

$$V_x^n = \mathbb{E} \left\{ \sum_{k=0}^{+\infty} \frac{1}{(1 + \lambda)^{k+1}} C_{X^k}^{s^n} \mid X^0 = x \right\},$$

et on a fait l'hypothèse que les coûts soient positifs.

- La suite V^n est décroissante. On note $A^{s^n} = M^{s^n} - (1 + \lambda) I$. Alors, l'algorithme impose :

$$A^{s^n} V^n + C^{s^n} = 0.$$

En soustrayant ces égalités à deux étapes successives, on trouve :

$$\underbrace{\left(A^{s^{n+1}} - A^{s^n} \right) V^n + \left(C^{s^{n+1}} - C^{s^n} \right)}_{\leq 0 \text{ par minimalité de } s^{n+1}} + A^{s^{n+1}} (V^{n+1} - V^n) = 0.$$

D'où $A^{s^{n+1}} (V^{n+1} - V^n) \geq 0$. Et en voyant $V^{n+1} - V^n$ comme la solution d'une équation de Kolmogorov actualisée avec un coût négatif on obtient :

$$V^{n+1} \leq V^n.$$

- Comme le nombre de politiques est fini, la suite V^n converge en un nombre fini d'étapes vers la solution de l'équation de la programmation dynamique. \square

UNICITÉ DE LA SOLUTION DE (5). Supposons qu'il existe deux solutions (V^1, s^1) et (V^2, s^2) à (5). Nous allons utiliser le même genre d'arguments que dans la preuve de l'existence. On a :

$$A^{s^1} V^1 + C^{s^1} = 0 \text{ et } A^{s^2} V^2 + C^{s^2} = 0.$$

En soustrayant ces deux équations, on obtient :

$$\underbrace{\left(A^{s^1} - A^{s^2} \right) V^1 + \left(C^{s^1} - C^{s^2} \right)}_{\leq 0 \text{ par minimalité de } s^1} + A^{s^2} (V^1 - V^2) = 0.$$

D'où $V^1 - V^2$ est solution d'une équation de Kolmogorov actualisée à coût négatif et donc $V^2 \geq V^1$. On obtient de même que $V^2 \leq V^1$. \square

V SOLUTION DE (5) EST SOLUTION DU PROBLÈME DE COMMANDE OPTIMALE. Soit V solution de (5) et σ une stratégie quelconque. Supposons que X^n évolue selon la loi induite par σ , on a :

$$\mathbb{E}^\sigma \{ V_{X^{n+1}} / (1 + \lambda) - V_{X^n} \mid X^n = x \} = 1 / (1 + \lambda) (M^{\sigma^n} V - (1 + \lambda) V)_x.$$

Par l'optimalité de V , on a que $(1 + \lambda)V_x \leq (M^u V + C^u)_x, \forall u \in \mathcal{F}$, et donc :

$$(M^{\sigma^n} V - (1 + \lambda)V^n)_x \geq -C_x^{\sigma^n}.$$

Si on note $\mathbb{E}^{n,\sigma}$ l'espérance conditionnelle connaissant les états réalisés jusqu'à l'instant n , pour la loi induite par la stratégie σ , on obtient :

$$\mathbb{E}^{n,\sigma} \left\{ 1/(1 + \lambda)^{n+1} V_{X^{n+1}} - 1/(1 + \lambda)^n V_{X^n} \right\} \geq -1/(1 + \lambda)^{n+1} C_{X^n}^{\sigma^n}.$$

Alors :

$$\begin{aligned} \mathbb{E}^{0,\sigma} \{-V_{X^0}\} &= \mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{\infty} 1/(1 + \lambda)^{n+1} V_{X^{n+1}} - 1/(1 + \lambda)^n V_{X^n} \right\}, \\ &\geq -\mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{\infty} 1/(1 + \lambda)^{n+1} C_{X^n}^{\sigma^n} \right\}. \end{aligned}$$

D'où :

$$V_{X^0}^0 \leq \mathbb{E}^{0,\sigma} \left\{ \sum_{n=0}^{\infty} 1/(1 + \lambda)^{n+1} C_{X^n}^{\sigma^n} \right\}.$$

On conclut en remarquant que la stratégie optimale conduit à l'égalité. \square

CHAPITRE 3

Propriétés combinatoires spectrales et asymptotiques des chaînes de Markov

Lorsque l'on doit étudier les chaînes de Markov sur un temps long ou des problèmes actualisés avec un facteur d'actualisation petit on doit comprendre la structure spectrale associée à la valeur propre 1 des matrices stochastiques. Ces structures sont complètement connues [Gan66, Pal65]. Nous rappelons ici ces résultats.

1. Propriétés spectrales des matrices stochastiques

Par la définition d'une matrice stochastique, on sait que chacune de ses lignes est une loi de probabilité. en particulier on a donc :

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

D'où le résultat :

Proposition 3.1. *M a la valeur propre 1, et le vecteur propre associé est $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.*

Remarque 3.2. *M peut bien sûr avoir d'autres valeurs propres de module 1. En effet, considérons la matrice stochastique $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Ses valeurs propres sont 1 et -1.*

Proposition 3.3. *M vérifie :*

$$|M|_\infty \leq 1.$$

DÉMONSTRATION.

$$|M|_\infty := \frac{\max_{1 \leq x \leq E} \left| \sum_{y \in \mathcal{E}} M_{xy} v_y \right|}{\max_{1 \leq x \leq E} |v_x|} \leq \frac{\max_{1 \leq x \leq E} \max_{1 \leq z \leq E} |v_z| \sum_{y \in \mathcal{E}} M_{xy}}{\max_{1 \leq x \leq E} |v_x|} \leq 1.$$

□

La propriété suivante est relative à la décomposition de Jordan d'une matrice stochastique.

Proposition 3.4. *Les valeurs propres de M de module 1 n'ont pas de nilpotent associé dans la décomposition de Jordan de la matrice. On dit alors que les valeurs propres sont semi-simples.*

DÉMONSTRATION. Raisonnons par l'absurde, pour un bloc de Jordan de dimension 2. Si les valeurs propres de module 1 avaient un nilpotent associé, on aurait un bloc de Jordan de la forme $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Or :

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

On aurait alors que la norme $|M^n|_\infty \rightarrow +\infty$, ce qui n'est pas possible, du fait que $|M^n|_\infty = |M|_\infty^n \leq 1$. \square

2. Propriétés combinatoires

Nous introduisons ici un ensemble de notions issus de la théorie des graphes qui sont liées aux propriétés asymptotiques des chaînes de Markov.

Soit une chaîne de Markov de matrice de transition M . On définit l'application *successeur* $\Gamma : \mathcal{E} \rightarrow \mathcal{P}(\mathcal{E})$ par :

$$\Gamma(x) = \{y \in \mathcal{E}, M_{xy} > 0\}.$$

Exemple 3.5. Considérons la chaîne de Markov décrite dans la figure 1.

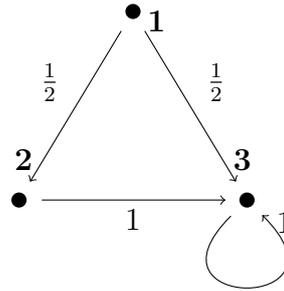


FIG. 1.

La matrice de transition associée est :

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

La fonction successeur est alors définie par : $\Gamma(1) = \{2, 3\}$, $\Gamma(2) = \{3\}$ et $\Gamma(3) = \{3\}$.

Définition 3.6 (Chemin, circuit). *Un chemin π de x^0 à x^n est une suite*

$$\pi = (x^0, x^1, \dots, x^n) \text{ avec } x^k \in \Gamma(x^{k-1}), \forall k \neq 0.$$

Un circuit est un chemin tel que $x^n = x^0$.

Les chemins de longueur 0 sont les états et sont des circuits. Ainsi, dans l'exemple 3.5, on a un seul circuit, correspondant à l'état 3 (il est de longueur 0).

Enfin, le pgcd de la longueur des circuits est appelé la *période* de M .

Revenons à l'exemple 3.5 et imaginons une particule se déplaçant selon cette loi de transition. Au vu des probabilités de transition choisies, il apparaît clairement qu'asymptotiquement la particule atteindra avec la probabilité 1 l'état 3, et y restera bloquée.

Nous allons maintenant formaliser cette idée. On introduit donc la relation d'équivalence \sim sur \mathcal{E} :

$$x \sim y \iff x \text{ et } y \text{ appartiennent au même circuit.}$$

On définit alors la relation d'ordre \succsim sur l'espace quotient \mathcal{E}/\sim par :

$$\bar{x} \succsim \bar{y} \iff \text{il existe un chemin de } x \text{ à } y, x \in \bar{x}, y \in \bar{y}$$

On peut maintenant introduire les notions de classe finale et de classe transitoire.

Définition 3.7 (Classe finale, classe transitoire). *Une classe d'équivalence est dite finale si elle est minimale pour la relation d'ordre \succsim . Une classe non finale est dite transitoire. Lorsque la chaîne a une seule classe, elle est dite irréductible.*

Illustrons un peu ces définitions par un exemple.

Exemple 3.8. Soit une chaîne de Markov dont la matrice de transition est représentée par le graphe suivant :

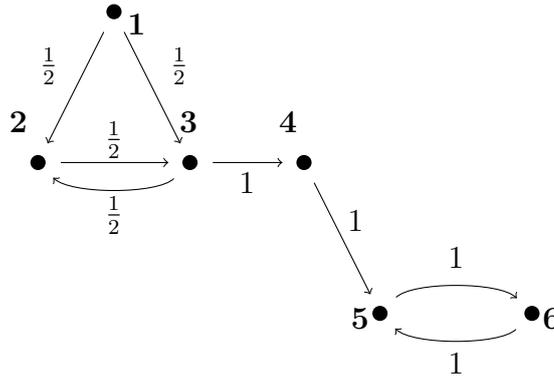


FIG. 2.

Les différentes classes sont : $\{1\}$, $\{2, 3\}$, $\{4\}$ et $\{5, 6\}$. Parmi elles, $\{1\}$, $\{2, 3\}$ et $\{4\}$ sont transitoires, et $\{5, 6\}$ est finale.

3. Classification des matrices stochastiques

Nous pouvons à présent faire le lien entre les propriétés introduites dans la section 2, les propriétés spectrales de la matrice de transition de la chaîne de Markov, et les propriétés asymptotiques de la chaîne elle-même. Pour une étude plus détaillée, le lecteur pourra se référer à [Pal65].

Soit une matrice stochastique M , on a puisque les valeurs propre de module 1 sont semi-simples que :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} (I + M + \dots + M^{n-1}) = P,$$

où P est le *projecteur spectral* associé à la valeur propre 1.

Le tableau 1 donne la classification des matrices stochastiques et rappelle leurs propriétés combinatoires (classes finales et transitoires), spectrales (valeurs et vecteurs propres) et asymptotiques (convergence vers les mesures invariantes).

Type	Combinatoire	Spectrale	Asymptotique $n \rightarrow \infty$ $p = pM$
Général	Plusieurs classes finales.	valeurs propres de module 1 semi-simples et multiples.	$\frac{1}{n} \sum_{i=0}^{n-1} M^i \rightarrow P$
Primitive	Toutes les classes finales sont de période 1.	1 est la seule valeur propre de module 1.	$M^n \rightarrow P$
Ergodique	Une seule classe finale.	1 est valeur propre simple.	$\frac{1}{n} \sum_{i=0}^{n-1} p^i \rightarrow p$
Régulière (Primitive et Ergodique)	Une seule classe finale de période 1.	1 est valeur propre simple et est la seule valeur propre de module 1.	$p^n \rightarrow p$
Irréductible	Une seule classe.	$p_x > 0 \quad \forall x \in \mathcal{E}$	$\frac{1}{n} \sum_{i=0}^{n-1} p^i \rightarrow p$

TAB. 1. Propriétés spectrales, combinatoires et asymptotiques des chaînes de Markov.

4. Calcul du projecteur spectral P

Pour calculer le projecteur spectral P on détermine des bases ayant une interprétation probabiliste des noyaux à droite et à gauche du générateur $A = M - I$ de la chaîne de Markov de matrice de transition M .

On peut montrer le résultat suivant :

Théorème 3.9 (Base des mesures de probabilités invariantes). *Il existe une base (p_1, \dots, p_m) de $\mathcal{N}(A')$ telle que :*

- Les p_i sont des mesures de probabilité de supports disjoints.
- Leurs supports sont les classes finales de la chaîne de Markov.
- Donc m est égal au nombre de classes finales de la chaîne.
- Le calcul des coefficients non nuls des p_i se ramène à calculer la mesure invariante unique de la classe finale i correspondante.

DÉMONSTRATION. Pour la démonstration de ce résultat on peut se référer par exemple à [QV91]. \square

Il existe une numérotation (les numéros des états d'une même classe finale sont connexes) des états telle que la matrice $A = M - I$ s'écrive :

$$A = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_m & 0 \\ A_{t1} & A_{t2} & \cdots & A_{tm} & A_{tt} \end{pmatrix}$$

Les blocs diagonaux A_1, \dots, A_m correspondent aux classes finales, A_{tt} correspond aux états transitoires.

Proposition 3.10. *La matrice A_{tt} est inversible.*

DÉMONSTRATION. La matrice A_{tt} est inversible car $M_{tt} = A_{tt} + I$ est de norme strictement plus petite que 1. En effet par définition d'une classe transitoire, de tout point de cette classe il existe un chemin allant à une classe finale, et tout chemin plus long reste dans cette classe finale. Ainsi, à partir d'un certain rang E , la probabilité de sortir de la classe transitoire est strictement positive, donc $|M_{tt}^E|_\infty < 1$. \square

Utilisons cette propriété pour construire une base de $\mathcal{N}(A)$ ayant une interprétation probabiliste.

Proposition 3.11. *Les vecteurs χ^j définis par $A\chi^j = 0$, $\chi^j = 1$ sur f_j et $\chi^j = 0$ ailleurs forment une base de $\mathcal{N}(A)$. Le nombre $\chi_x^j, x \in t$ est la probabilité de finir dans la classe finale i lorsque l'on part de l'état transitoire x .*

DÉMONSTRATION. Si on note f_j les classes finales de la chaîne de Markov, on a que :

$$A\chi^j = 0 \iff \begin{cases} A_j \mathbf{1}_{f_j} = 0 \\ A_{tj} \mathbf{1}_{f_j} + A_{tt} \chi^j = 0 \end{cases}$$

La première équation est vérifiée du fait que f_j est une classe finale.

A_{tt} étant inversible, du fait que la classe t est transitoire, la seconde équation est vérifiée pour $\chi^j = -A_{tt}^{-1} A_{tj} \mathbf{1}_{f_j}$. L'interprétation se déduit du fait que $A_{tt}^{-1} = I + M_{tt} + M_{tt}^2 + \dots$. \square

On vérifie alors facilement :

Théorème 3.12. *Soit $\{f_j, 1 \leq j \leq m\}$ l'ensemble des classes finales de la chaîne de Markov. Alors :*

$$P = \sum_{j=1}^m \chi_j \otimes p_j$$

est le projecteur spectral associé à la valeur propre 1.

Exemple 3.13. On considère la chaîne de Markov de matrice de transition :

$$M = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

On peut la représenter à l'aide du graphe de la figure 3.

On observe ainsi facilement deux classes finales $\{2, 3\}$ et $\{4, 5, 6\}$. On a donc deux mesures invariantes :

$$\left[0 \ \frac{1}{2} \ \frac{1}{2} \ 0 \ 0 \ 0\right] \text{ et } \left[0 \ 0 \ 0 \ \frac{1}{3} \ \frac{1}{3} \ \frac{1}{3}\right],$$

et deux vecteurs propres à droite :

$$\begin{bmatrix} \frac{1}{2} \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ et } \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

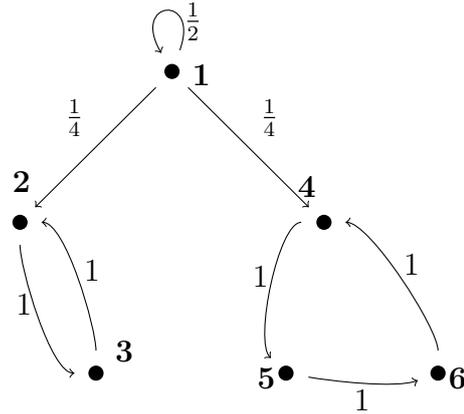


FIG. 3. Graphe de la chaîne de Markov de l'exemple 3.13.

Le projecteur spectral associé est donc :

$$P = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Commande optimale des chaînes de Markov en observation incomplète

Dans ce chapitre, nous étudions la commande optimale des chaînes de Markov partiellement observées. À la différence du cadre décrit dans le chapitre 2 nous n'observons plus l'état, mais seulement une observation dont la loi, discrète, est définie à partir de l'état grâce à une matrice de transition donnée. On reformule le problème de commande. Nous montrons que le problème de commande optimale en observation incomplète peut se ramener à la résolution de deux problèmes distincts :
 – estimer l'état à l'aide de la chaîne de Markov connaissant le passé des observations,
 – contrôler cet estimateur dont on montre que c'est un processus markovien contrôlé dont on peut déterminer l'équation de la programmation dynamique.

1. Chaînes de Markov contrôlées et observées

Définition 4.1. Une chaîne de Markov commandée et observée sera la donnée du 6-uple $(\mathcal{T}, \mathcal{E}, \mathcal{F}, \mathcal{G}, M^{uy}, p^o)$: – l'espace de temps est noté $\mathcal{T} = \mathbb{N}$, – l'espace d'état est $\mathcal{E} = \{1, \dots, E\}$, – l'espace des commandes est $\mathcal{F} = \{1, \dots, F\}$, – l'espace des observations $\mathcal{G} = \{1, \dots, G\}$ (une observation sera notée $y \in \mathcal{G}$), – une famille de matrices de transition : $M_{xx'}^{uy}$ est la probabilité de transiter de x à x' et d'observer y si la commande est u . – une loi de probabilité initiale p^o sur $\mathcal{E} \times \mathcal{G}$.

On note \mathcal{R} l'ensemble des politiques relaxées. Une *politique relaxée* est une suite $\rho = (\rho_n, n \in \mathcal{T})$ de probabilités sur \mathcal{F} conditionnellement au passé des observations :

$$\rho_n^{y_o \dots y_n u_n} = \mathbb{P} \{U_n = u_n \mid Y_o = y_o, \dots, Y_n = y_n\}.$$

Dans l'espace des trajectoires des triplets (état, commande, observation)

$$\Omega = (\mathcal{E} \times \mathcal{F} \times \mathcal{G})^{\mathcal{T}}$$

on définit la loi de probabilité grâce à sa valeur sur les cylindres de trajectoires :

$$\mathbb{P} \{(x, y, u)_o, (x, y, u)_1, (x, y, u)_2, \dots\} = p_{x_o y_o}^o \rho_o^{y_o u_o} M_{x_o x_1}^{u_o y_1} \rho_1^{y_1 u_1} M_{x_1 x_2}^{u_1 y_2} \dots$$

Etant donné un coût instantané $C \in \mathbb{R}^{\mathcal{E} \times \mathcal{F}}$ et un coût sur l'état final $\Phi \in \mathbb{R}^{\mathcal{E}}$ l'objectif est de minimiser un critère du type :

$$\min_{\rho \in \mathcal{R}} \mathbb{E} \left\{ \sum_{n=0}^{N-1} C_{X_n}^{U_n} + \Phi_{X_N} \right\} \text{ ou } \min_{\rho \in \mathcal{R}} \mathbb{E} \left\{ \sum_{n=0}^{+\infty} \frac{1}{(1+\lambda)^{n+1}} C_{X_n}^{U_n} \right\}.$$

2. Filtre optimal d'une chaîne de Markov observée

Nous considérons le cas particulier d'une chaîne de Markov observée qui correspond au cas où $F = 1$. Il n'est plus alors nécessaire de faire figurer u dans les indices. Nous donnons l'équation du *filtre optimal* qui donne l'évolution de loi de l'état connaissant les observations passées, c'est-à-dire :

$$q_x^n = \mathbb{P} \{X_n = x \mid Y_o = y_o, Y_1 = y_1, \dots, Y_n = y_n\}, \forall x \in \mathcal{E}, \forall n \in \mathcal{T}.$$

Dans la suite on notera $\mathbf{1}$ le vecteur de taille E constitué de 1.

Théorème 4.2. *Le filtre optimal de la chaîne de Markov q^n vérifie $q^n = r^n / r^n \mathbf{1}$ où on appelle r^n le filtre non normalisé qui satisfait l'équation suivante :*

$$r^n = r^{n-1} M^{y_n}, \quad r^0 = p^0.$$

DÉMONSTRATION. Par la définition de la probabilité \mathbb{P} :

$$r^n = \mathbb{P}\{y_0, y_1, \dots, y_n, x_n\} = \sum_{x_0, x_1, \dots, x_n} p_{x_0 y_0}^0 M_{x_0 x_1}^{y_1} \dots M_{x_{n-1} x_n}^{y_n} = p_{y_0}^0 M^{y_1} \dots M^{y_n}.$$

□

Exemple 4.3 (Filtrage). Considérons notre vendeur de marionnette qui ne connaît pas les probabilités de transition régissant son commerce et qui voudrait détecter si son commerce est régi avec les probabilités de transition du haut ou du bas du graphe de la figure 1. Il observe seulement ses succès $y = 1$ et ses échecs $y = 0$.

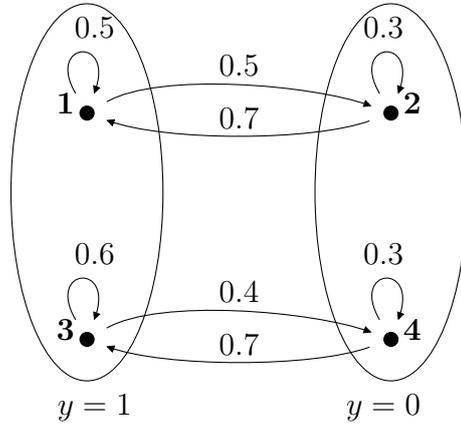


FIG. 1. Chaîne de Markov partiellement observée.

Nous avons donc les matrices de transitions suivantes selon que $y = 1$ ou $y = 0$

$$M^1 = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 \\ 0 & 0 & 0.7 & 0 \end{bmatrix}, \quad M^0 = \begin{bmatrix} 0 & 0.5 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0.3 \end{bmatrix}.$$

Nous considérons la suite d'observations suivante :

$$y = [1 \ 0 \ 0 \ 1 \ 1 \ 0].$$

Puisque la première observation est 1, on a directement les probabilités initiales :

$$p_0^1 = [\frac{1}{2} \ 0 \ \frac{1}{2} \ 0], \quad p_0^0 = [0 \ 0 \ 0 \ 0], \quad q^0 = p_0^1.$$

Alors :

$$q^1 = \frac{q^0 M^0}{q^0 M^0 \mathbf{1}} = \frac{1}{0.45} [0 \ 0.25 \ 0 \ 0.2] = [0 \ 0.56 \ 0 \ 0.44].$$

De même :

$$\begin{aligned} q^2 &= [0 \ 0.56 \ 0 \ 0.44], \\ q^3 &= [0.56 \ 0 \ 0.44 \ 0], \\ q^4 &= [0.51 \ 0 \ 0.49 \ 0]. \end{aligned}$$

En poursuivant ainsi sur un grand nombre d'itérations, cette technique permettra de déterminer si l'état reste "en haut" ou "en bas". Ainsi :

$$\begin{aligned} q^{100} &= [0.88 \ 0 \ 0.12 \ 0], \\ q^{500} &= [0 \ 0.99 \ 0 \ 0.01]. \end{aligned}$$

On a détecté que dans ce cas l'état était "en haut".

Nous voyons donc que le filtre est un processus stochastique vivant dans le simplexe des loi de probabilités sur \mathcal{E}

$$\mathcal{S}_E = \left\{ p \mid \sum_{x \in \mathcal{E}} p_x = 1, p_x \geq 0, x \in \mathcal{E} \right\}.$$

Nous schématisons l'évolution du filtre dans la figure 2 dans le cas $E = 3$.

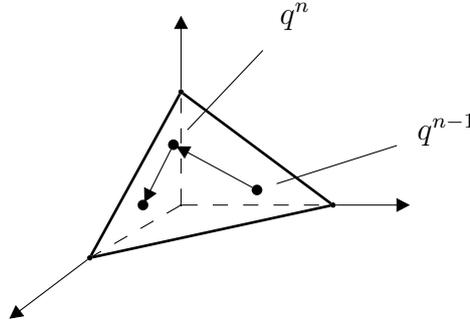


FIG. 2. Évolution du filtre dans le simplexe \mathcal{S}_E , dans le cas où \mathcal{E} est de cardinal 3.

Alors que l'état est discret le filtre vit dans un espace continu de taille le nombre d'état. Il y a là un énorme saut de complexité. Le filtre est un processus de Markov sur \mathcal{S}_E :

Proposition 4.4. *Le filtre q^n est une chaîne de Markov sur \mathcal{S}_E de transition :*

$$(6) \quad \mathbb{P} \left\{ q \longrightarrow \frac{qM^y}{qM^y \mathbf{1}} \right\} = qM^y \mathbf{1}.$$

DÉMONSTRATION.

$$\mathbb{P} \{ Y_n = y \mid Y_{n-1} = y_{n-1}, \dots, Y_o = y_o \} = \frac{\mathbb{P} \{ Y_n = y, Y_{n-1} = y_{n-1}, \dots, Y_o = y_o \}}{\mathbb{P} \{ Y_{n-1} = y_{n-1}, \dots, Y_o = y_o \}} = \frac{r^n \mathbf{1}}{r^{n-1} \mathbf{1}}.$$

□

3. Équation de la programmation dynamique dans le cas d'une observation incomplète

Nous allons maintenant utiliser la propriété de Markov du filtre pour résoudre le problème de commande optimale :

$$(7) \quad \min_{\rho \in \mathcal{R}} \mathbb{E} \left\{ \sum_{n=0}^N C_{X_n}^{U_n} \right\},$$

Ce problème (7) se réduit alors en utilisant les propriétés de l'espérance conditionnelle à :

$$(8) \quad \min_U \mathbb{E} \left\{ \sum_{n=0}^N \mathbb{E}^{\mathcal{G}_n} \{ C_{X_n}^{U_n} \} \right\} = \min_U \mathbb{E} \left\{ \sum_{n=0}^N q^n C^{U_n} \right\},$$

où \mathcal{G}_n désigne les observations antérieures à n .

Puisque q^n est un processus de Markov on peut lui appliquer les techniques de la programmation dynamique. On en déduit :

Théorème 4.5. *La fonction valeur :*

$$v^n(q) = \min_{\rho \in \mathcal{R}} \mathbb{E} \left\{ \sum_{k=n}^N C_{X_k}^{U_k} \mid q^n = q \right\}$$

vérifie l'équation de la programmation dynamique :

$$v^n(q) = \min_{u \in \mathcal{F}} \left\{ \sum_{y \in \mathcal{G}} v^{n+1} \left(\frac{qM^{uy}}{qM^{uy}\mathbf{1}} \right) qM^{uy}\mathbf{1} + qC^u \right\}, \quad v^{N+1}(q) = 0.$$

Illustrons cet algorithme par un exemple :

Exemple 4.6 (Maintenance). Nous considérons le problème de la maintenance d'un groupe électrogène, dont les caractéristiques sont les suivantes. L'état du groupe est $\mathcal{E} = \{0, 1\}$, où 0 signifie la panne et 1 le fonctionnement normal.

L'opérateur ne peut pas directement observer si le groupe est en panne. Le test sans remplacement coûte C_1^1 . Si lors du test on doit changer le matériel defectueux cela coûte au total C_0^1 . Si on ne teste pas et que le groupe est defectueux on prend un risque que l'on traduit par le coût C_0^0 . La commande appartient à l'espace $\{0, 1\}$ où 1 signifie que l'on décide de tester la machine, et 0 que l'on ne teste pas.

Enfin, l'espace des observations est $\mathcal{G} = \{\emptyset, 0, 1\}$, où \emptyset signifie que l'on n'a pas observé, 0 correspond à observer que la machine est en panne et 1 correspond à observer que la machine est en fonctionnement normal.

La figure 3 décrit le comportement du système.

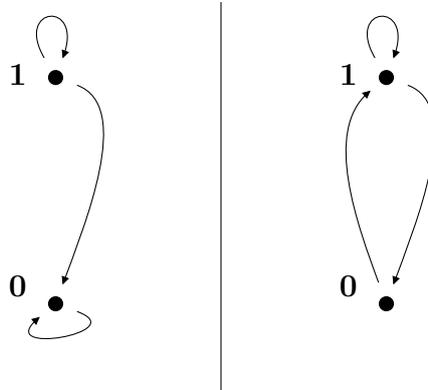


FIG. 3. Problème de maintenance.

On écrit donc les six matrices de transition correspondant aux différentes valeurs de la commande u et de l'observation y . Rappelons que nos notations sont telles que

$M_{xx'}^{uy}$ représente la probabilité de transiter de x à x' et d'observer y , sachant que l'on exerce la commande u . Ainsi :

$$M^{00} = \begin{bmatrix} 1 & 0 \\ 1-p & p \end{bmatrix}, \quad M^{01} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad M^{10} = \begin{bmatrix} 0 & 0 \\ p & 0 \end{bmatrix}, \quad M^{11} = \begin{bmatrix} 0 & 1 \\ 0 & 1-p \end{bmatrix},$$

et :

$$M^{10} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad M^{11} = \begin{bmatrix} 0 & 1 \\ p & 0 \end{bmatrix}, \quad M^{11} = \begin{bmatrix} 0 & 1 \\ 0 & 1-p \end{bmatrix}.$$

Il faut calculer les états possible du filtre sachant qu'il est dans l'état $q = [q_0, q_1]$. On obtient :

$$\frac{qM^{00}\mathbf{1}}{qM^{00}\mathbf{1}} = [1 - q_1p, q_1p], \quad \frac{qM^{10}\mathbf{1}}{qM^{10}\mathbf{1}} = [1 \ 0], \quad \frac{qM^{11}\mathbf{1}}{qM^{11}\mathbf{1}} = [0 \ 1].$$

Pour écrire l'équation de la programmation dynamique il faut aussi évaluer les quantités $qM^{uy}\mathbf{1}$ non nulles. Dans notre cas on en a trois seulement :

$$qM^{00}\mathbf{1} = 1, \quad qM^{11}\mathbf{1} = 1 - q_1p, \quad qM^{10}\mathbf{1} = q_1p.$$

D'où l'équation de la programmation dynamique en horizon fini :

$$v^n(q) = \min \left\{ v^{n+1}([1 - q_1p, q_1p]) + q_0C_0^0, \right. \\ \left. v^{n+1}([0 \ 1])(1 - q_1p) + q_0C_0^1 + v^{n+1}([1 \ 0])q_1p + q_1C_1^1 \right\}.$$

Pour résoudre ce genre d'équation numériquement, il est nécessaire de discrétiser le simplexe dans lequel vit q ou de paramétriser de façon fini les points du simplexe atteignable. Pour des problèmes de plus grande taille, il n'est pas possible de discrétiser le simplexe de façon général. Le lecteur intéressé pourra consulter [FQ04] pour voir des exemples de problèmes de commande optimale en observation incomplète de grande taille résolus numériquement en approximant le support de la loi du filtre.

CHAPITRE 5

Problème 1 : 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 ρ . 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.

1. Enoncé

1.0.1. *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 β deux constantes positives), dans le cas $u > 0$, on donnera une formule explicite pour p .
- (7) Etant donné un taux d'actualisation λ , donnez l'équation satisfaite par :

$$v_x^\lambda = E\left\{\sum_0^{\infty} 1/(1 + \lambda)^{n+1} X^n U^n | X^0 = x\right\},$$

où X^n désigne l'état de la chaîne de Markov et U^n sa commande à l'instant n .

- (8) Montrez que λv^λ reste borné lorsque λ tend vers 0.
- (9) On fait un développement de v^λ sous la forme $\nu/\lambda + w^0 + \lambda w^1 + \dots$. Donnez un système d'équations caractérisant ν .
- (10) En déduire que ν est un vecteur constant.
- (11) Quelle est l'interprétation probabiliste de ν ?
- (12) Calculez explicitement ν .

1.0.2. 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 λ .
- (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 α on ne turbine pas et qu'au dessus de ce seuil on turbine au maximum $u = 1 - \rho$.
- (4) On fait tendre à nouveau λ vers 0 et on fait à nouveau un développement en λ 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 α déterminant la stratégie optimale.

2. Corrigé

2.0.3. 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^λ est l'équation de Kolmogorov arrière :

$$(1 + \lambda)v_x^\lambda = (Mv^\lambda)_x + xu , \forall x .$$

(8) Le fait que λv^λ reste borné lorsque λ 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 λ et en conservant les deux premiers termes on obtient le système d'équations :

$$\nu = M\nu ,$$

$$\nu_x + w_x^0 = (Mw^0)_x + xu ,$$

qui caractérise complètement ν . En effet la première équation implique que ν 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 $p\nu = u \sum_x xp_x$ qui détermine complètement ν .

(10) Puisque ν est dans le noyau de $M - I$ et que $\sum_{x'} M_{xx'} = 1$, ν un vecteur constant. On appellera encore ν la constante correspondante

(11) Puisque $\nu = \sum_x xup_x$, ν s'interprète comme la moyenne du coût au sens de la mesure invariante p .

(12) Le calcul explicite de ν 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 :

$$\nu = u\beta S'(\beta)/S(\beta) ,$$

avec $\beta = \rho/u$ et $S(\beta) = (\beta^E - 1)/(\beta - 1)$.

2.0.4. 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 :

$$\begin{aligned} \nu &= \max_{0 \leq u \leq 1-\rho} M^u \nu, \\ \nu_x + w_x^0 &= \max_{u \in U_x^*} ((M^u w^0)_x + xu), \forall x, \end{aligned}$$

où U^* désigne l'ensemble des u réalisant le minimum dans la première équation. La première équation implique que ν est un vecteur constant et que

$$U_x^* = [0, 1 - \rho]$$

(la première équation indique seulement que ν est constant et ne donne pas d'information sur u). Ces 2 équations peuvent donc se réécrire en une seule : rechercher la constante ν et w tels que

$$\nu + w_x = \max_{0 \leq u \leq 1-\rho} ((M^u w)_x + xu), \forall x.$$

- (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.

CHAPITRE 6

Problème 2 : 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 α . Si on ne teste pas le système ($u = 0$) avant une étape, il a une probabilité ϵ 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.

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 ϵ est petit, beaucoup plus petit que α (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 λ .
- (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.

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 extrêmes 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 extrêmes associées à la stratégie S_{ij} . Elle vérifie $p^{ij} = p^{ij} M^{ij}$. Ces mesures sont extrêmes si et seulement si leurs supports sont des classes finales. On a :

$$p^{00} = (1 \ 0), \quad p^{01} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$p^{10} = (\epsilon/(1+\epsilon) \ 1/(1+\epsilon)), \quad p^{11} = (0 \ 1).$$

- (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 λ sont alors :

$$(1 + \lambda)v^{ij} = M^{ij}v^{ij} + c^{ij}.$$

- (4) Le premier terme du développement (lorsque λ 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 ν^{ij}/λ pour la politique S_{ij} avec ν^{ij} donné par $p^{ij}c^{ij}$ (dans le cas d'une seule classe finale les composantes de ν^{ij} sont identiques, dans le cas de deux classes finales le coefficient dépend de la classe finale considérée). On obtient donc :

$$\nu^{00} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nu^{01} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix},$$

$$\nu^{10} = \begin{pmatrix} (1 + \alpha)\epsilon/(1 + \epsilon) \\ (1 + \alpha)\epsilon/(1 + \epsilon) \end{pmatrix}, \quad \nu^{11} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}.$$

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 λ sous la forme

$$v = \nu/\lambda + w + \lambda w^1 \dots$$

On obtient en identifiant les termes de même puissance en λ

$$\nu_k = \min_u [M^u \nu]_k ,$$

$$\nu_k + w_k = \min_{u \in U_k^*} [M^u w + c^u]_k ,$$

avec

$$M^0 = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 - \epsilon \end{pmatrix}, \quad M^1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} ,$$

et $U_k^* = \operatorname{argmin}_u [M^u \nu]_k$. Dans le cas où la politique optimale conduit à une chaîne de Markov possédant une seule classe finale, ν_k ne dépend plus de k car le seul vecteur propre à droite de la chaîne est

$$\nu \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

D'autre part $U_k^* = \{0, 1\}$. La condition d'optimalité s'écrit alors

$$\nu + w_k = \min_{u \in U_k^*} [M^u w + c^u]_k ,$$

avec ν 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 (ν, w) satisfaisant

$$\nu = M^S \nu, \nu + w = M^S w + c^S .$$

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

$$\begin{aligned} S^{00} &\rightarrow \nu = 1, w_0 = 1/\epsilon, w_1 = 0 \rightarrow S^{11} \rightarrow \nu = \alpha, w_0 = 1, w_1 = 0 \rightarrow S^{10} \\ &\rightarrow \nu = (1 + \alpha)\epsilon/(1 + \epsilon), w_0 = (1 + \alpha)/(1 + \epsilon), w_1 = 0 \rightarrow S^{10} . \end{aligned}$$

Deuxième partie

Algèbre maxplus et chaînes de Bellman

Dans cette partie, nous introduisons la structure algébrique de semi-anneau idempotent. Le cas particulier de cette structure appelé maxplus, qui correspond aux nombres réels munis du “max” noté additivement et du “plus” noté multiplicativement, est la structure algébrique naturelle pour les problèmes d’optimisation. En particulier les problèmes d’optimisation sont vus comme des problèmes linéaires sur cette structure ou ses généralisations vectorielles. De plus un calcul matriciel peut être développé sur cette structure avec les mêmes gains de notation que dans l’algèbre ordinaire. Une sous classe des réseaux de Pétri appelé graphe d’événement, adaptée à la modélisation des problèmes de synchronisation, peut être vu comme l’analogie des systèmes dynamiques linéaires de la théorie standard avec sa représentation sous forme d’état par un triplet de matrice (A, B, C) . Enfin la recopie du vocabulaire des probabilités en substituant l’algèbre ordinaire par l’algèbre maxplus s’interprète en une théorie de la décision éclairant d’un jour nouveau la commande optimale. En particulier l’analogie des chaînes de Markov que l’on peut appeler chaînes de Bellman sont les problèmes de commande optimale. En particulier, les équations de la programmation dynamique déterministe (Hamilton Jacobi) ne sont que les équations de Kolmogorov à ce changement d’algèbre près. On donne dans cette partie une introduction à ces notions.

La référence la plus classique sur les structures ordonnées est le livre de Birkhof [Bir67]. On pourra consulter [BCOQ92] ou [Gau] disponibles sur le web pour un développement plus complet des notions introduites dans cette partie. Le site web maxplus.org est dédié à cette algèbre et à ses applications.

CHAPITRE 7

Semi-anneaux idempotents

1. Structures

Définition 7.1 (Semi-anneau, dioïde, semi-corps, semi-module). *Un semi-anneau (S, \oplus, \otimes) est un ensemble S muni de deux opérations vérifiant :*

- \oplus est associative, commutative, et admet un élément neutre noté ε ,
- \otimes est associative, distributive par rapport à \oplus , et admet un élément neutre noté e ,
- ε est absorbant pour \otimes . $\forall a \in S, \varepsilon \otimes a = a \otimes \varepsilon = \varepsilon$.

Lorsque le semi-anneau est idempotent, i.e. : $\forall a \in S, a \oplus a = a$, on dira que c'est un dioïde. Lorsque les éléments différents de ε ont un inverse on dira que la structure est semi-corps. La structure vérifiant les axiomes des modules sur des scalaires appartenant à un semi-anneau au lieu d'un anneau sera appelé semi-module.

Remarque 7.2 (Semi-anneau idempotent versus anneau). Il manque à la structure de semi-anneau l'existence d'opposés (inverse pour \oplus) pour en faire un anneau. Il y a une bifurcation au niveau des axiomes entre les semi-anneaux simplifiables $a \oplus b = a \oplus c \Rightarrow b = c$ qui sont symétrisables et les semi-anneaux idempotents. En effet supposons qu'un semi-anneau idempotent soit simplifiable et appliquons la règle de simplification à la définition de l'idempotence et de l'existence d'un élément neutre soit $a \oplus a = a = a \oplus \varepsilon$ cela implique $a = \varepsilon$. La structure est réduite à ε .

Autement dit anneau et semi-anneau idempotent sont deux structures aussi riches algébriquement l'une que l'autre mais incompatibles. Par contre l'idempotence est au moins aussi redoutable que la règle de simplifiabilité pour éviter l'explosion combinatoire de la taille des formules comme nous le montre, par exemple, la formule du binôme pour les semi-anneaux idempotents commutatifs :

$$(a \oplus b)^n = \bigoplus_{k=0}^n a^k \otimes b^{n-k} .$$

Remarque 7.3. Ces structures algébriques, bien que moins étudiées que les structures classiques (groupes, anneaux, corps, module, espaces vectoriels), jouent néanmoins un rôle essentiel dans la modélisation. Dans une application, il est souvent intéressant de se poser comme question quelle est la structure algébrique de base des objets considérés.

Le tableau 1 donne des exemples de semi-anneaux idempotents en indiquant leurs domaines naturels d'application.

2. Matrices sur des semi-anneaux

A partir de ces structures scalaires sont construites des structures vectorielles de semi-anneaux. En particulier on peut construire un semi-anneaux de matrices en

Idempotent

\mathcal{S}	\oplus	\otimes	ε	e	Application	Nom
$\mathbb{R} \cup \{+\infty\}$	min	+	$+\infty$	0	Plus court chemin	\mathbb{R}_{\min}
$\mathbb{R} \cup \{-\infty, +\infty\}$	min	+	$+\infty$	0	Plus court chemin	$\overline{\mathbb{R}}_{\min}$
$\mathbb{R} \cup \{-\infty\}$	max	+	$-\infty$	0	Plus long chemin	\mathbb{R}_{\max}
$\mathbb{R}^+ \cup \{+\infty\}$	max	min	0	$+\infty$	Logique floue	$\mathbb{R}_{\max, \min}^+$
$[0, 1]$	max	\times	0	1	Max. vraisemblance	$[0, 1]_{\max, \times}$
$\{0, 1\}$	ou	et	0	1	Logique	\mathbb{B}
$\mathcal{P}(\Sigma^*)$	\cup	Prod. Latin	\emptyset	-	Langage	\mathbb{L}

Non idempotent

\mathbb{R}^+	+	\times	0	1	Probabilité	
----------------	---	----------	---	---	-------------	--

TAB. 1. Exemples de semi-anneaux.

utilisant le produit matriciel :

$$[A \otimes B]_{ij} = \bigoplus_k A_{ik} \otimes B_{kj}.$$

Les polynômes ou les séries formels $(a_i)_{i \geq 0}$ héritent de la structure de semi-anneaux des scalaires en utilisant le produit :

$$(a_i)_{i \geq 0} \otimes (b_i)_{i \geq 0} = \left(\bigoplus_k (a_k \otimes b_{i-k}) \right)_{i \geq 0}.$$

3. Exemples

Illustrons les quelques notions que nous venons d'introduire.

– Le produit matriciel maxplus (on rappelle que $\varepsilon = -\infty$ et $e = 0$) :

$$\begin{bmatrix} 2 & \varepsilon \\ 2 & 5 \end{bmatrix} \otimes \begin{bmatrix} 5 & e \\ 3 & e \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 8 & 5 \end{bmatrix}$$

– Les polynômes formels maxplus :

$$\begin{aligned} (e \oplus 2x \oplus 3x^2) \otimes (5x \oplus 7x^3) &= 5x \oplus 7x^2 \oplus 8x^3 \oplus 7x^3 \oplus 9x^4 \oplus 10x^5 \\ &= 5x \oplus 7x^2 \oplus 8x^3 \oplus 9x^4 \oplus 10x^5 \end{aligned}$$

– Notations dans maxplus :

$$3^{\otimes 2} = 3 \otimes 3 = 3 + 3 = 6, \quad 3/4 = 3 - 4 = -1,$$

$$\sqrt{2} = 1/2, \quad 3^{1/5} = 3 \times \frac{1}{5} = \frac{3}{5}.$$

– Quelques formules maxplus :

$$(a \oplus b)^n = a^n \oplus b^n, \quad \min(a, b) = \frac{a \otimes b}{a \oplus b}.$$

Résolution des systèmes linéaires dans l'algèbre maxplus

1. Linéarité maxplus

Définition 8.1 (Linéarité maxplus). Une fonction $H : \mathbb{R}_{\max}^n \rightarrow \mathbb{R}_{\max}^n$ est linéaire au sens maxplus si et seulement si :

- (1) $H(\lambda + X) = \lambda + H(X)$,
- (2) $H(\max(X, Y)) = \max(H(X), H(Y))$.

On peut alors représenter l'application linéaire H par une matrice réelle, par exemple dans le cas $n = 2$:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \max(a + x_1, b + x_2) \\ \max(c + x_1, d + x_2) \end{bmatrix}.$$

Nous cherchons par la suite à résoudre des systèmes linéaires. Prenons un exemple.

Exemple 8.2. Soit le système :

$$\begin{cases} \max(3 + x_1, x_2) = 5 \\ \max(x_1, 2 + x_2) = 5. \end{cases}$$

De la première équation, nous savons que soit $x_2 = 5$ et $3 + x_1 \leq x_2$, soit $3 + x_1 = 5$ et $x_2 \leq 3 + x_1$. Or, si $x_2 = 5$, alors la deuxième équation ne peut être vérifiée. Donc $x_1 = 2$ et $x_2 \leq 3 + x_1$. Alors $2 + x_2 = 5$ et $x_1 \leq 2 + x_2$, donc $x_2 = 3$.

Cet exemple de résolution montre que sa généralisation à la dimension n conduirait à une complexité exponentielle rapidement inutilisable.

Remarque 8.3. Les applications maxplus linéaires sont rarement surjectives. Pour s'en convaincre on peut tracer l'image dans le plan d'une application linéaire maxplus de \mathbb{R}_{\max}^2 dans \mathbb{R}_{\max}^2 . On voit que de façon général c'est une bande parallèle à la première diagonale de largeur limitée. On peut voir cela facilement en pensant que l'on fait des combinaisons linéaires maxplus des vecteurs colonnes de la matrice représentant l'application linéaire.

2. Interprétation du produit matriciel

Soit A une matrice carrée réelle. Regardons de plus près le produit $A \otimes A$.

$$A \otimes A = \bigoplus_j (A_{ij} \otimes A_{jk}) = \max_j (A_{ij} + A_{jk}).$$

Ainsi, $(A \otimes A)_{ik}$ calcule le chemin le plus long pour aller de i à k en deux étapes si les coefficients de la matrice A_{ij} sont interprétés comme des longueurs de chemin pour aller de i à j .

Remarque 8.4. On voit apparaître l'intérêt de cette algèbre pour la commande optimale. L'équation de la programmation dynamique peut s'écrire $V^n = A \otimes V^{n-1}$ et sa solution est $V^n = A^{\otimes(N-n)} \otimes V^N$. Ce n'est donc que l'équation de Kolmogorov arrière dans l'algèbre maxplus.

3. Equation affine à une inconnue dans \mathbb{R}_{\max}

On considère l'équation affine générale :

$$(9) \quad a \otimes x \oplus b = a' \otimes x \oplus b',$$

que l'on notera lorsqu'il n'y a pas d'ambiguïté possible :

$$ax \oplus b = a'x \oplus b'.$$

Il faut bien comprendre que du fait qu'on ne peut pas facilement symétriser l'opération $\oplus = \max$, l'équation affine générale ne peut être réduite à : $ax \oplus b = \epsilon$.

Théorème 8.5 (Solutions de l'équation affine générale dans \mathbb{R}_{\max}). *Considérons l'équation (9).*

- si $a' < a$ et $b < b'$ alors il existe une unique solution :

$$x = \frac{b \oplus b'}{a \oplus a'},$$

la division étant à prendre au sens maxplus (c'est une soustraction au sens classique).

- si $a < a'$ et $b < b'$ alors il n'y a pas de solution.

- si $a = a'$ et $b \neq b'$ alors tout $x \geq \frac{b \oplus b'}{a}$ est solution.

- si $a \neq a'$ et $b = b'$ alors tout $x \leq \frac{b}{a \oplus a'}$ est solution.

- si $a = a'$ et $b = b'$ alors tout $x \in \mathbb{R}_{\max}$ est solution.

Nous allons vérifier ce théorème en s'aidant des représentations graphiques de ces fonctions affines. Dans un premier temps, nous traçons la représentation graphique de la fonction affine $x \rightarrow a \otimes x \oplus b$.

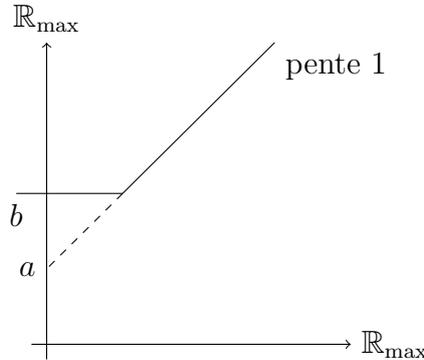


FIG. 1. Représentation graphique de la fonction affine $x \rightarrow a \otimes x \oplus b$.

Ainsi, à titre d'exemple, vérifions que si $a' < a$ et $b < b'$ alors il existe une unique solution $x = \frac{b \oplus b'}{a \oplus a'}$. La figure 2 représente ce cas.

Un cas différent de la théorie des équations affines dans l'algèbre classique, est celui où $a = a'$ et $b \neq b'$, représenté dans la figure 3. Ici, on a une infinité de solutions, qui est la demi-droite $x \geq \frac{b \oplus b'}{a} = \max(b, b') - a$.

Remarque 8.6 (Forme canonique). Même si l'on ne peut pas se ramener à $ax \oplus b = \epsilon$ à cause de l'absence d'opposés, on peut tout de même se ramener à une forme simplifiée que l'on qualifiera de canonique. Voyons l'exemple suivant :

$$3 \otimes x \oplus 2 = 2 \otimes x \oplus 3.$$

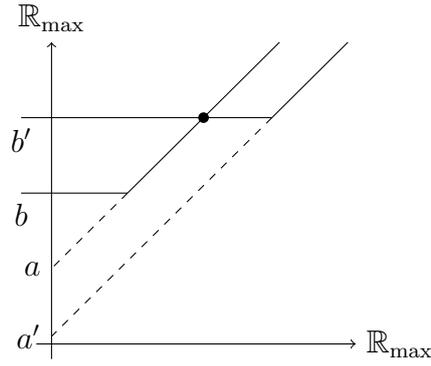


FIG. 2. Cas où l'équation affine générale a une unique solution ($a' < a$ et $b < b'$).

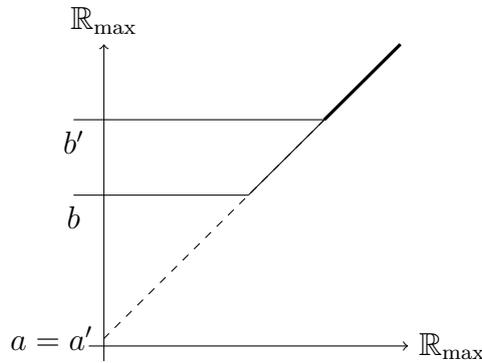


FIG. 3. Cas où l'équation affine générale a une demi-droite solution ($a = a'$ et $b < b'$).

Rappelons nous que cela signifie que l'on cherche x tel que :

$$\begin{aligned} \max(3+x, 2) = \max(2+x, 3) &\iff 3+x = \max(2+x, 3) \\ &\iff 3+x = \max(2+x, 3) \\ &\iff 3+x = 3 \end{aligned}$$

On s'est ramené par de simples règles de prépondérance à : $3 \otimes x = 3$.

De manière générale, on peut toujours se ramener à une des cinq formes suivantes présentées dans le tableau 1.

$$\begin{aligned} a \otimes x = b &\implies \text{une solution} \\ a \otimes x \oplus b = \varepsilon &\implies \text{pas de solution} \\ a \otimes x \oplus b = a \otimes x &\implies x \geq \frac{b}{a} = b - a \\ a \otimes x \oplus b = b &\implies x \leq \frac{b}{a} = b - a \\ a \otimes x \oplus b = a \otimes x \oplus b &\implies x \in \mathbb{R}_{\max} \end{aligned}$$

TAB. 1. Formes canoniques associées à l'équation affine de \mathbb{R}_{\max} .

Ainsi, si $a > a'$, il suffit de conserver a et de remplacer a' par ε . De même, si $b > b'$, il suffit de conserver b et de remplacer b' par ε .

Exemple 8.7. De même, l'équation générale de \mathbb{R}_{\max}^n peut être mise sous forme canonique. Donnons un exemple en dimension 2 :

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

Pour résoudre les systèmes linéaires de taille n nous allons présenter trois types d'approche : – la résolution de point fixe, – la résiduation dans $\overline{\mathbb{R}}_{\max}$, qui permet de résoudre les équations du type $Ax \leq b$, – les formules de Cramer maxplus qui seront présentés dans le chapitre suivant après avoir introduit une symétrisation incomplète de \mathbb{R}_{\max} .

4. Méthode du point fixe

Soit à résoudre \mathbb{R}_{\max}^n l'équation de point fixe :

$$(10) \quad x = A \otimes x \oplus b,$$

qui a des interprétations en terme de graphe.

Définition 8.8 (Graphe de précédence). *Soit une matrice $A \in (\mathbb{R}_{\max})^{n \times n}$. On appelle graphe de précédence $\mathcal{G}(A)$ associé à A le graphe à n noeuds numérotés $\{1, \dots, n\}$ tel que l'arc (i, j) existe si et seulement si $A_{ij} \neq \varepsilon$, et dans ce cas le poids associé à l'arc est A_{ij} .*

Théorème 8.9 (Solutions à l'équation de point fixe). *Si, dans $\mathcal{G}(A)$, tous les circuits ont de poids négatif ou nuls, alors l'équation (10) admet une solution $x = A^*b$, avec :*

$$A^* \triangleq \bigoplus_{i=0}^{+\infty} A^i = \bigoplus_{i=0}^{n-1} A^i.$$

Si, de plus, tous les circuits de $\mathcal{G}(A)$ sont de poids strictement négatif, alors la solution est unique.

EXISTENCE D'UNE SOLUTION. Si A^*b existe, alors :

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

Donc A^*b est solution. Or $(A^*)_{ij}$ s'interprète le poids maximal des chemins composés d'un nombre quelconque d'arc (longueur) allant i à j . Ainsi, une condition nécessaire et suffisante d'existence de A^* , et donc d'une solution, est que les circuits soient de poids négatifs ou nuls. \square

UNICITÉ DE LA SOLUTION. Si x est solution de (10), alors :

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

Et donc en utilisant la définition de A^* , on a que pour tout $k \geq n$:

$$x = A^*b \oplus A^kx.$$

Alors, les poids des circuits étant tous strictement négatifs, le poids associé au chemin allant de i à j en exactement k coups tend vers $\varepsilon = -\infty$ lorsque k tend vers ∞ . Donc $A^kx \rightarrow \varepsilon$ lorsque k tend vers ∞ . Et donc $x = A^*b$. \square

5. Résiduation

Soit A une matrice de $\overline{\mathbb{R}}_{\max}^{n \times n}$, où $\overline{\mathbb{R}}_{\max} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$, et b un vecteur de $\overline{\mathbb{R}}_{\max}^n$. Nous étudions la résolution de l'équation $Ax = b$. Ce problème n'a pas toujours de solution puisqu'en général A n'est pas surjective. Par contre il existe toujours une *plus grande sous-solution*

$$(11) \quad \max \underbrace{\{x \mid Ax \leq b\}}_{\mathcal{S}},$$

que nous noterons $A \setminus b$ qui est facile à calculer. Il suffit alors de faire le calcul $A(A \setminus b)$ et de le comparer à b pour savoir s'il y a une solution ou non à $Ax = b$.

Remarque 8.10 (Opération \setminus dans $\overline{\mathbb{R}}_{\max}$). Soient a et b des éléments de \mathbb{R} , alors :

$$a \setminus b \triangleq \max_x \{ax \leq b\} = b - a.$$

De plus :

$$\varepsilon \setminus \varepsilon = \top \triangleq +\infty, \quad \top \setminus \varepsilon = \varepsilon, \quad \varepsilon \setminus \top = \top, \quad \top \setminus \top = \top.$$

Théorème 8.11 (Résiduation). *Soit A une matrice de $\overline{\mathbb{R}}_{\max}^{n \times n}$ et b un vecteur de $\overline{\mathbb{R}}_{\max}^n$, la plus grande solution de $Ax \leq b$ existe. Elle est notée $A \setminus b$. Elle vaut :*

$$(A \setminus b)_j = \min_i [A_{ij} \setminus b_i],$$

où l'opération \setminus pour les scalaires est explicitée dans la remarque 8.10.

DÉMONSTRATION.

$$\begin{aligned} \{Ax \leq b\} &\iff \left\{ \bigoplus_{j=1}^n A_{ij} \otimes x_j \leq b_i, \forall i = 1, \dots, n \right\}, \\ &\iff \{x_j \leq b_i - A_{ij}, \forall i, j = 1, \dots, n\}, \\ &\iff \left\{ x_j \leq \min_i \{b_i - A_{ij}\}, \forall j = 1, \dots, n \right\}. \end{aligned}$$

Or $x_j = \min_i [A_{ij} \setminus b_i]$ définit un vecteur de $\overline{\mathbb{R}}_{\max}^n$ avec $x \in \mathcal{S}$. C'est donc bien la plus grande solution de $Ax \leq b$. \square

Réseaux de Petri et Théorie Spectrale

Dans cette partie, on montre que les graphes d'événements (sous classe des réseaux de Pétri) se modélise linéairement dans l'algèbre maxplus. Un réseau de Petri [Pet62] est un graphe permettant de spécifier de façon claire la dynamique d'automates dans lesquels apparaissent des synchronisations. Ces réseaux ou sa variante appelée GRAFCET sont beaucoup utilisés en informatique et productique.

1. Modèle

Rappelons le fonctionnement des réseaux de Petri.

Définition 9.1 (Réseau de Petri). *Un réseau de Petri est défini un 6-uplet $(\mathcal{P}, \mathcal{Q}, M, N, m^0, \tau)$ où :*

- \mathcal{P} désigne un ensemble fini de places,
- \mathcal{Q} désigne un ensemble fini de transitions,
- M est une matrice de taille $\mathcal{Q} \times \mathcal{P}$ à coefficients dans \mathbb{N} donnant la multiplicité des arcs sortant des transitions,
- N est une matrice de taille $\mathcal{P} \times \mathcal{Q}$ à coefficients dans \mathbb{N} donnant la multiplicité des arcs entrants,
- $m^0 \in \mathbb{N}^{\mathcal{P}}$ le marquage initial,
- $\tau \in \mathbb{N}^{\mathcal{Q}}$ est la temporisation.

On illustre la définition par la figure 1.

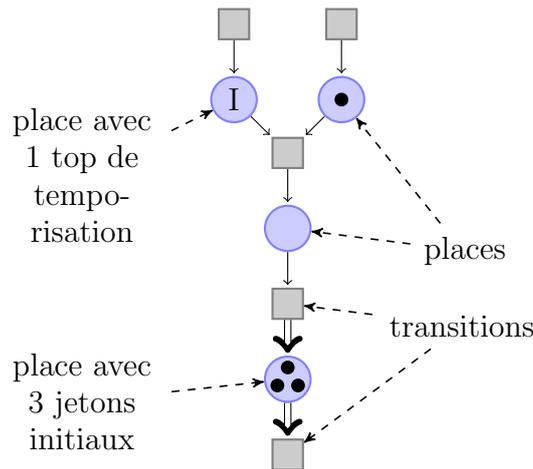


FIG. 1. Exemple de réseau de Petri. La multiplicité est représentée par le nombre de flèches.

La dynamique est régie par *les brûlages des transitions*. Lorsque toutes les places en amont de la transition ont un nombre de jetons (ayant séjourné un temps supérieur à la temporisation de la place) plus grand ou égal à la multiplicité de l'arc les joignant à la transition, celle-ci peut brûler. Lorsque la transition brûle un nombre de

jetons égal à la multiplicité des arcs (en amont) sont enlevés de toutes les places en amont et un nombre de jetons égal à la multiplicité des arcs (en aval) est placé dans chacune des places en aval de la transition. L'étape de brûlage est illustrée dans la figure 2.

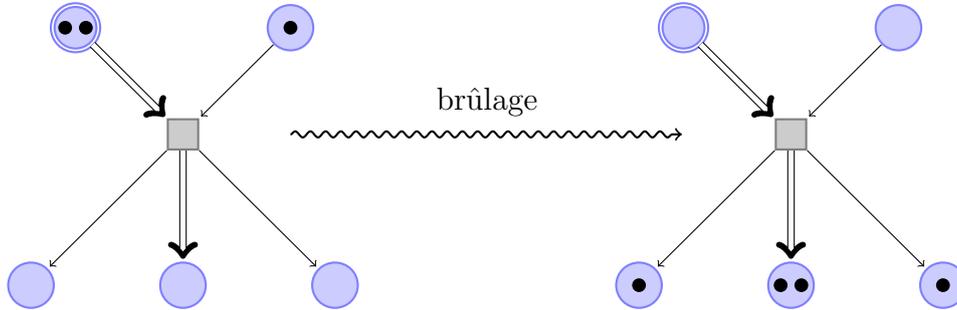


FIG. 2. Exemple de brûlage.

Définition 9.2. Lorsque les temporisations sont nulles et les multiplicités des arcs sont toutes égales à 1, on parle de réseau de Petri.

En revanche, lorsque les temporisations ne sont pas toutes nulles ($\tau \neq 0$), on parle de réseau de Petri temporisé, et lorsque les multiplicités ne sont pas toutes égales à 1, on parle de réseau de Petri avec multiplicateur.

Enfin, on parle de graphe d'évènements lorsque le réseau est composé de place où arrive un seul arc et d'où part un seul arc.

2. Quelques réseaux de Pétri élémentaires

Nous montrons quelques exemples de systèmes dynamiques facilement modélisables avec des réseaux de Pétri.

Exemple 9.3 (Tâche nécessitant plusieurs ressources). Considérons une pièce à usiner sur une machine. Pour pouvoir usiner il faut que la pièce et la machine soit disponible. Les deux ressources (pièce et machine) sont représentés par un jeton. La transition représente le tâche d'usinage. Ce système est modélisé dans la figure 3.

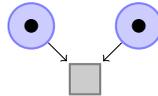


FIG. 3. Tâche nécessitant plusieurs outils.

Exemple 9.4 (Processus d'exclusion).

Nous souhaitons modéliser le fait que deux véhicules (ici des jetons) ne peuvent occuper simultanément la même place. Ceci se fait en ajoutant une place fictive qui donne l'autorisation d'entrer dans la place si elle est vide. Cet exemple est donné dans la figure 4.

Initialement la place p_2 a un jeton, et il n'y pas de jeton dans la place p_h , le jeton de la place p_1 ne peut donc pas entrer dans p_2 , car la brûlage de la transition t_2 ne peut avoir lieu que si les deux places p_1 et p_h ont un jeton. A $t = 1$, la place p_2 n'a plus de jeton et surtout la place p_h en a un. Le brûlage de la transition t_2 peut donc avoir lieu, et le jeton de la place p_1 peut donc entrer dans p_2 .

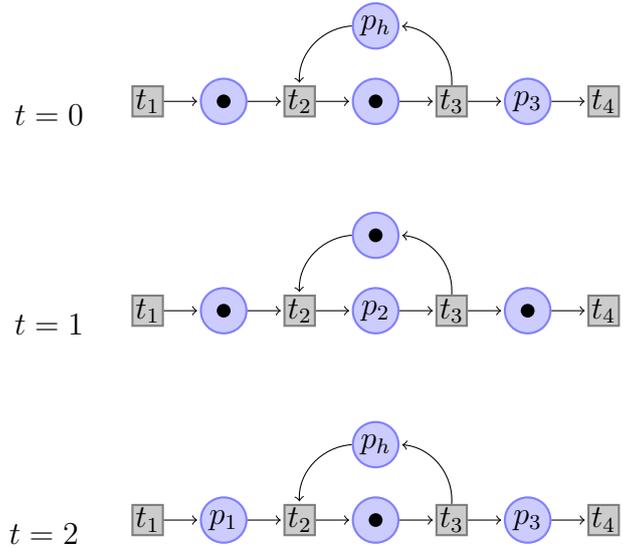


FIG. 4. File de véhicules.

3. Les graphes d'événements sont des systèmes maxplus linéaires

Montrons sur exemple que les graphes d'événements temporisés sont des systèmes dynamiques maxplus linéaires Prenons l'exemple du graphe d'événements temporisés, représenté dans la figure 5.

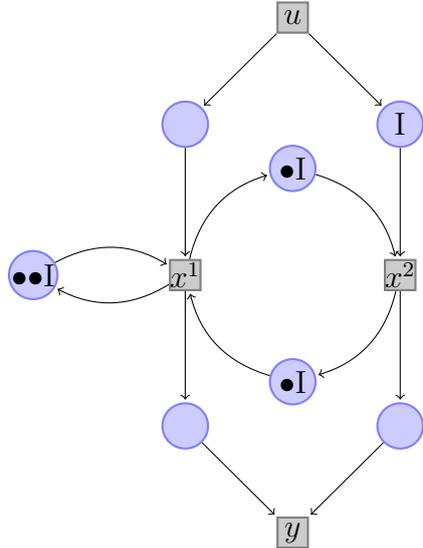


FIG. 5. Exemple de graphe d'événements temporisé.

On appelle dateur une fonction associée a une transition qui indique la date du brûlage de cette transition en fonction de son numéro de brûlage. Par exemple u_n donne la date du n -ième brûlage de la transition u . On cherche l'équation d'évolution des dateurs des transitions du réseau de Pétri. De façon plus générale on s'intéresse à la relation entre les dates d'entrée de sortie de jetons dans le système.

On a, pour tout entier $k \geq 2$, les équations d'évolution suivantes en exprimant les contraintes existantes entre les instants de brûlage des transitions explicitées par

le réseau de Pétri :

$$\begin{cases} x_k^1 = \max(1 + x_{k-1}^2, 1 + x_{k-2}^1, u_k), \\ x_k^2 = \max(1 + x_{k-1}^1, 1 + u_k), \\ y_k = \max(x_k^1, x_k^2). \end{cases}$$

Avec les notations maxplus, on peut écrire :

$$\begin{cases} x_k = A_1 x_{k-1} \oplus A_2 x_{k-2} \oplus B u_k, \\ y_k = C x_k, \end{cases}$$

avec :

$$A_1 = \begin{bmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & \varepsilon \end{bmatrix}, \quad B = \begin{bmatrix} e \\ 1 \end{bmatrix}, \quad C = [e \quad e].$$

Théorème 9.5. *Tout graphe d'événement temporisé admet une réalisation dans l'algèbre maxplus de la forme :*

$$\begin{cases} D x_k = A x_{k-1} \oplus B u_k, \\ y_k = C x_k. \end{cases}$$

Ce résultat peut être prouvé en rajoutant des transitions supplémentaires de telle sorte que chaque place contienne au plus un jeton.

4. Théorie spectrale

Un problème important pour les réseaux de Pétri est la détermination du taux de sortie (le nombre d'événements par unité de temps d'une sortie). Considérons le cas particulier $B = \varepsilon$, $D = e$, $C = e$ et supposons que le réseau soit fortement connexe dans ce cas la dynamique se ramène à $x_{k+1} = A x_k$ et on peut montrer que le système croît linéairement avec un comportement périodique autour de l'asymptote :

$$\exists K, T : \forall k \geq K, x_{k+T} = \lambda^{\otimes T} \otimes x_k$$

où λ est solution du problème de valeur propre maxplus

$$(12) \quad \lambda \otimes x = A \otimes x$$

On se reportera à [BCOQ92] pour avoir une démonstration de ce résultat. Ici nous étudierons uniquement le problème spectral.

Théorème 9.6. *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, ζ parcourt l'ensemble des circuits de $\mathcal{G}(A)$, on note $|\zeta|_w$ le poids du circuit ζ et $|\zeta|_1$ sa longueur.

DÉMONSTRATION. EXISTENCE DE x ET DE λ . Considérons la matrice $B = A \phi \lambda \stackrel{\text{def}}{=} (e \phi \lambda) A$, où $\lambda = \max_{\zeta} |\zeta|_w / |\zeta|_1$. Le poids maximum des circuits de $\mathcal{G}(B)$ est e . Par conséquent B^* and $B^+ = B B^*$ existent. La matrice B^+ a une colonne possédant un e sur la diagonale. Pour prouver cette affirmation choisissons un noeud k d'un circuit ξ tel que $\xi \in \arg \max_{\zeta} |\zeta|_w / |\zeta|_1$. Le poids maximum des chemins allant

de k à k est e . Et donc nous avons $e = B_{kk}^+$. Soit $B_{.k}$ 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

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

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

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

INTERPRÉTATION DE λ EN TERME DE GRAPHE. Si λ satisfait (??), 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 $A_{i_1i_2}x_{i_2} = \lambda x_{i_1}$, 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 λ^{m-l+1} est le poids du circuit de longueur $m-l+1$, ou, dit autrement, λ est le poids moyen du circuit β . 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 (9.6).

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 λ est plus grand que le poids moyen de γ . Par conséquent λ est le poids moyen maximum des circuits et donc λ 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 9.9 ci dessous).

Remarque 9.7. 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 9.8. Sous les hypothèses du théorème 9.6, l'unicité du vecteur propre n'est pas garantie comme le montre l'exemple suivant

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

et

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

Les deux vecteurs propres ne sont pas “proportionnels”.

Exemple 9.9. – L'exemple suivant est un contre exemple trivial à l'unicité dans le cas non fortement connexe

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

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

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

mais

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

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 :

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

Symétrisation de l'algèbre maxplus

Dans cette section on prolonge \mathbb{R}_{\max} en un ensemble plus grand appelé \mathbb{S} en faisant une construction analogue à la construction de \mathbb{Z} à partir de \mathbb{N} . Nous avons déjà vu qu'il faut choisir entre l'axiome d'idempotence et de simplifiabilité. Il s'en suit qu'il n'est pas possible de symétriser complètement \mathbb{R}_{\max} par contre grâce aux nouveaux nombres ainsi construit on peut énoncer de façon claire l'analogue maxplus des formules de Cramer donnant les solutions aux systèmes linéaires d'équations généraux.

Théorème 10.1. *Tout groupe idempotent est réduit à l'élément neutre ε .*

DÉMONSTRATION. Soit a un élément arbitraire d'un groupe idempotent (G, \oplus) dont l'élément neutre est noté ε . Soit b l'opposé de a On a :

$$a = a \oplus \varepsilon = a \oplus (a \oplus b) = (a \oplus a) \oplus b = a \oplus b = \varepsilon .$$

Il s'en suit que $G = \{\varepsilon\}$. □

Considérons l'ensemble \mathbb{R}_{\max}^2 muni des deux opérations :

$$\begin{cases} (x_1, x_2) \oplus (y_1, y_2) &= (x_1 \oplus y_1, x_2 \oplus y_2) , \\ (x_1, x_2) \otimes (y_1, y_2) &= (x_1 y_1 \oplus x_2 y_2, x_1 y_2 \oplus x_2 y_1) , \end{cases}$$

C'est un dioïde (semi-anneau idempotent) avec pour élément nul $(\varepsilon, \varepsilon)$ et unité (e, e) .

Si $x = (x_1, x_2)$ on définit : $- \ominus x = (x_2, x_1)$, $- |x| = x_1 \oplus x_2$, $- \dot{x} = x \ominus x = x \oplus (\ominus x) = (|x|, |x|)$ et sur \mathbb{R}_{\max}^2 la relation d'équivalence :

$$(x_1, x_2) \mathcal{R} (y_1, y_2) \Leftrightarrow \begin{cases} x_1 \oplus y_2 = x_2 \oplus y_1 & \text{si } x_1 \neq x_2 \text{ et } y_1 \neq y_2 , \\ (x_1, x_2) = (y_1, y_2) & \text{si } x_1 = x_2 \text{ ou } y_1 = y_2 . \end{cases}$$

Cette relation possède trois types de classes d'équivalence :

$$\begin{cases} \overline{(t, -\infty)} = \{(t, x) \mid x < t\} , & \text{les nombres positifs dont l'ensemble est noté } \mathbb{S}^{\oplus} , \\ \overline{(-\infty, t)} = \{(x, t) \mid x < t\} , & \text{les nombres négatifs } \mathbb{S}^{\ominus} , \\ \overline{(t, t)} = \{(t, t)\} , & \text{les nombres équilibrés } \mathbb{S}^{\bullet} . \end{cases}$$

La relation d'équilibre définie par :

$$(x_1, x_2) \nabla (y_1, y_2) \Leftrightarrow x_1 \oplus y_2 = x_2 \oplus y_1 ,$$

n'est pas une relation d'équivalence car elle n'est pas transitive comme le montre l'exemple suivant $(\varepsilon, 1) \nabla (1, 1) \nabla (1, \varepsilon)$ mais $(\varepsilon, 1)$ n'est pas en équilibre avec $(1, \varepsilon)$.

La relation d'équivalence \mathcal{R} est compatible avec la somme et la multiplication de \mathbb{R}_{\max}^2 , avec la relation d'équilibre ∇ et avec les opérations \ominus , $|\cdot|$ et $\dot{\cdot}$.

On appelle $\mathbb{S} = \mathbb{S}^{\oplus} \cup \mathbb{S}^{\ominus} \cup \mathbb{S}^{\bullet}$ l'algèbre symétrisée $\mathbb{R}_{\max}^2 / \mathcal{R}$ de \mathbb{R}_{\max} . L'ensemble \mathbb{S} est un dioïde avec la somme et la multiplication héritées des opérations de \mathbb{R}_{\max}^2 . L'ensemble des éléments signés est défini par $\mathbb{S}^{\vee} = \mathbb{S}^{\oplus} \cup \mathbb{S}^{\ominus}$.

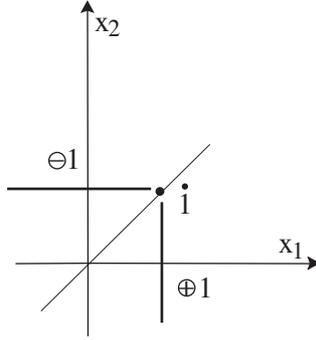


FIG. 1. Les trois types de classes d'équivalence.

Pour la manipulation des équations la relation \mathcal{R} est moins pratique que ∇ car : $a = b \Leftrightarrow (a \ominus b) \nabla \varepsilon, \forall a, b \in \mathbb{S}^\vee$, alors que $(a \ominus b) \mathcal{R} \varepsilon \Leftrightarrow a = b = \varepsilon$.

Les résultats suivant peuvent être démontrés facilement.

Théorème 10.2. *La relation ∇ vérifie :*

- (1) $a \nabla a$.
- (2) $a \nabla b \Leftrightarrow b \nabla a$.
- (3) $a \nabla x, x \nabla b$ et $x \in \mathbb{S}^\vee \Rightarrow a \nabla b$.
- (4) $a \nabla b, y \nabla a, b \in \mathbb{S}^\vee \Rightarrow a = b$.

Théorème 10.3. *Si $a \in \mathbb{S}_*^\vee = \mathbb{S}^\vee - \{\varepsilon\}$ et $b \in \mathbb{S}^\vee$, alors $x = \ominus b/a$ est l'unique solution de l'équilibre*

$$ax \oplus b \nabla \varepsilon$$

appartenant à \mathbb{S}^\vee .

Ce résultat peut s'étendre au cas vectoriel avec des vecteurs x, b et une matrice a de dimensions appropriés. Considérons alors une solution $x \in (\mathbb{R}_{\max})^n$ à l'équation :

$$(13) \quad Ax \oplus b = Cx \oplus d.$$

De la définition de l'équilibre on vérifie que :

$$(14) \quad (A \ominus C)x \oplus (b \ominus d) \nabla \varepsilon.$$

Inversement supposons que x est une solution positive de l'équilibre (14), alors

$$Ax \oplus b \nabla Cx \oplus d.$$

avec $Ax \oplus b \in (\mathbb{S}^\oplus)^n, Cx \oplus d \in (\mathbb{S}^\oplus)^n$. En utilisant le théorème 10.2 nous obtenons :

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

Puisque, l'ensemble des solutions du système (13) dans $(\mathbb{R}_{\max})^n$ est égal à l'ensemble des solutions positives de l'équilibre 14 dans \mathbb{S}^n , étudier les systèmes d'équations (13) se ramène à résoudre les équilibres linéaires (14), pour lesquels on dispose du théorème suivant.

Définition 10.4. Si $A \in \mathbb{S}^{n \times n}$, on définit :

$$\det(A) = \bigoplus_{\sigma} \operatorname{sgn}(\sigma) \bigotimes_{i=1}^n A_{i\sigma(i)},$$

$$A_{ij}^{\#} = \operatorname{cofactor}_{ji}(A),$$

et le signe de la permutation σ par :

$$\operatorname{sgn}(\sigma) = \begin{cases} e & \text{si } \sigma \text{ est paire,} \\ \ominus e & \text{si } \sigma \text{ est impaire.} \end{cases}$$

Théorème 10.5. Etant donné A une matrice $n \times n$ à coefficients dans \mathbb{S} , on a :

- (1) $AA^{\#} \nabla \det(A) e$.
- (2) Si $p(\lambda) \triangleq \det(A \ominus \lambda e)$, alors $p(A) \nabla \varepsilon$ (version maxplus du théorème de Cayley-Hamilton).
- (3) Si $\det(A) \in \mathbb{S}_*^{\vee}$ et $A^{\#}b \in (\mathbb{S}^{\vee})^n$, alors il existe une unique solution à $Ax \nabla b$ en $(\mathbb{S}^{\vee})^n$ donnée par : $x = A^{\#}b / \det(A)$.

Exemple 10.6. (1) $3 \ominus 2 = 3$; $\ominus 3 \oplus 2 = \ominus 3$; $\overset{\bullet}{2} \oplus 3 = 3$; $\overset{\bullet}{2} \ominus 3 = \ominus 3$; $\overset{\bullet}{2} \oplus 1 = \overset{\bullet}{2}$; $\overset{\bullet}{2} \ominus 1 = \overset{\bullet}{2}$.

(2) Considérons le système

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix},$$

alors

$$\det(A) = 3 \ominus 7 = \ominus 7,$$

$$\det \begin{bmatrix} 5 & 3 \\ 6 & 2 \end{bmatrix} = 7 \ominus 9 = \ominus 9 \Rightarrow x_1 = \ominus 9 / \ominus 7 = 2,$$

$$\det \begin{bmatrix} 1 & 5 \\ 4 & 6 \end{bmatrix} = 7 \ominus 9 = \ominus 9 \Rightarrow x_2 = \ominus 9 / \ominus 7 = 2.$$

On vérifie :

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

(3)

$$\det \begin{bmatrix} 1 \ominus \lambda & 3 \\ 4 & 2 \ominus \lambda \end{bmatrix} = (1 \ominus \lambda)(2 \ominus \lambda) \ominus 7 = \lambda^2 \ominus 2\lambda \ominus 7.$$

$$A^2 \ominus 2A \ominus 7e = \begin{bmatrix} 7 & 5 \\ 6 & 7 \end{bmatrix} \ominus \begin{bmatrix} 3 & 5 \\ 6 & 4 \end{bmatrix} \ominus \begin{bmatrix} 7 & \varepsilon \\ \varepsilon & 7 \end{bmatrix} = \begin{bmatrix} \overset{\bullet}{7} & \overset{\bullet}{5} \\ \overset{\bullet}{6} & \overset{\bullet}{7} \end{bmatrix}.$$

Décision Optimale

A partir de l'algèbre mmaxplus ou minplus on peut construire un formalisme analogue au calcul des probabilités. On préfère utiliser ici l'algèbre minplus qui est le dual de l'algèbre mmaxplus (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.

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 11.1. (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 11.2. 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 (de M. Akian) 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 11.3. (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. $d(x, y) = |e^{-x} - e^{-y}|$] 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'analogue 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 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 $c_{X+Y} = c_X \star c_Y$.

¹ \mathcal{O} représente l'ensemble des ouverts de E .

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 11.4. 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 11.5. 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 11.6. 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-compacité.

Définition 11.7. 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 11.8. 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 Γ désigne la fonction spéciale Gamma.

Théorème 11.9. 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. 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

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

4. Les théorèmes 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 11.10. *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 11.11. *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 11.12 (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 $\mathcal{M}_{0, \sigma^p}^p(X^0)$.

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 :

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

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.

TAB. 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	$\mathcal{M}_{m,\sigma}^p$

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 12.1. Une chaîne de Bellman homogène en temps, prenant un nombre fini de valeurs, est définie à partir du triplet (\mathcal{E}, C, ϕ) 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) ϕ 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 12.2. 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 12.3. 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 12.4. 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

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

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

Bibliographie

- [BCOQ92] F. Baccelli, G. Cohen, G. Y. Olsder, and J.-P. Quadrat, *Synchronization and linearity*, J. Wiley, 1992, maxplus.org.
- [Bel57] R. Bellman, *Dynamic programming*, Princeton University Press, 1957.
- [Bir67] G. Birkhof, *Lattice theory*, American Mathematical Society, 1967.
- [Fel57] W. Feller, *An introduction to probability theory and its application.*, J. Wiley, 1957.
- [FQ04] N. Farhi and J.-P. Quadrat, *Résolution numérique de problèmes de commande optimale de chaînes de markov observées imparfaitement*, Tech. Report 5348, INRIA, Octobre 2004.
- [Gan66] F. R. Gantmacher, *Théorie des matrices*, Dunod, 1966.
- [Gau] S. Gaubert, *Maxplus and discrete event systems*, amadeus.inria.fr/gaubert/.
- [How60] R. A. Howard, *Dynamic programming and markov process*, J. Wiley, 1960.
- [Pal65] R. Pallu de la Barrière, *Cours d'automatique théorique*, Dunod, 1965.
- [Pet62] C. A. Petri, *Kommunikation mit automaten*, Ph.D. thesis, University of Bonn, 1962.
- [QV91] J.-P. Quadrat and M. Viot, *Introduction à la commande stochastique*, www-rocq.inria.fr/metalau/quadrat/, 1991.

ANNEXE A

Cercles de Gershgorin

Le théorème de Gershgorin permet, à partir des coefficients d'une matrice complexe, de caractériser une zone du plan complexe à l'intérieur de laquelle se trouvent nécessairement ses valeurs propres. Il s'énonce ainsi :

Théorème A.1 (Théorème de Gershgorin). *Soit A une matrice complexe de taille $n \times n$, de terme général a_{ij} . Toute valeur propre de A appartient à l'un au moins des cercles de centre a_{ii} et de rayon $\sum_{j \neq i} |a_{ij}|$.*

Autrement dit, les valeurs propres de la matrice A appartiennent à l'ensemble du plan complexe :

$$\bigcup_{i=1}^n \mathcal{C} \left(a_{ii}, \sum_{j \neq i} |a_{ij}| \right).$$

DÉMONSTRATION. Soit λ une valeur propre de A . Alors il existe $x \in \mathbb{C}^*$ tel que :

$$Ax = \lambda x.$$

Alors, pour tout $i \in \{1, \dots, n\}$:

$$(a_{ii} - \lambda) = - \sum_{j \neq i} a_{ij} x_j,$$

et :

$$|a_{ii} - \lambda| |x_i| \leq \sum_{j \neq i} |a_{ij}| |x_j|.$$

Choisissons $i = \operatorname{argmax}_j |x_j|$. Alors :

$$|a_{ii} - \lambda| \leq \sum_{j \neq i} |a_{ij}| \left| \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |a_{ij}|.$$

□

Théorème A.2 (Corollaire). *Une matrice à diagonale dominante est inversible.*

ANNEXE B

Test

1. Projecteur spectral

Calculer le projecteur spectral associé à la valeur propre 1 de la matrice de transition markovienne :

$$M = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 0 & 1/4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 3/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

2. Programmation dynamique

Donner explicitement l'équation de la programmation dynamique du problème de commande stochastique suivant :

$$\max_{(U_n)} E \sum_{n=0}^{\infty} (1/2)^{n+1} (X_n U_n)$$

où X_n est une chaîne de Markov contrôlée à valeur 0 ou 1 dont la probabilité de transition est

$$\begin{bmatrix} 1/2 & 1/2 \\ u & 1 - u \end{bmatrix}$$

u peut prendre les valeurs $1/4$ et $3/4$.

3. Algorithme de Howard

Résoudre par l'algorithme de Howard l'équation :

$$\begin{cases} 2v_1 = \max\{v_1 + 1, v_2\} \\ 2v_2 = \max\{3v_1/4 + 1/4v_2 + 2, v_1/4 + 3v_2/4 + 1\} \end{cases}$$

4. Chemins les plus courts d'un graphe

On considère la matrice minplus donnant les longueurs des routes entre les noeuds représentés par les lignes et les colonnes de la matrice :

$$M = \begin{bmatrix} 0 & 2 & 3 \\ 3 & 0 & 2 \\ 5 & 1 & 0 \end{bmatrix}$$

Calculer la matrice donnant les longueurs des chemins minimaux entre tous les noeuds en utilisant l'algèbre minplus.

5. Mise en équation d'un graphe d'événements

On sait que les équations donnant les dates des brûlages des transitions d'un graphe d'événements sans entrées ni sorties est de la forme

$$X_{n+1} = A \otimes X_n$$

où \otimes désigne le produit matriciel maxplus. Donner la matrice A associée au graphe d'événements suivant :

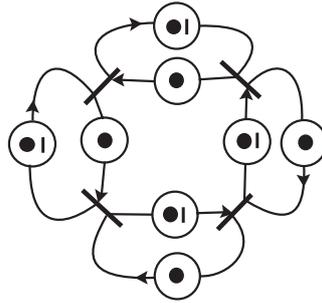


FIG. 1. Le graphe d'événements.

6. Calcul de valeur propre de matrice minplus

Etant donnée la matrice minplus suivante

$$A = \begin{bmatrix} \varepsilon & a_1 & \varepsilon & \varepsilon & \bar{a}_5 \\ \bar{a}_1 & \varepsilon & a_2 & \varepsilon & \varepsilon \\ \varepsilon & \bar{a}_2 & \varepsilon & a_3\varepsilon & \\ \varepsilon & \varepsilon & \bar{a}_3 & \varepsilon & a_4\varepsilon \\ a_5 & \varepsilon & \varepsilon & \bar{a}_4 & \varepsilon \end{bmatrix}$$

où les a_i désignent des nombres prenant les valeurs 0 ou 1 et $\bar{a} = 1 - a$ donner la valeur propre minplus de A comme fonction des a_i (attention le cours a été fait dans le cadre maxplus). Pour quelles configurations des a_i la valeur propre atteint un maximum.